Eur. Phys. J. D **46**, 227–240 (2008)
DOI: 10.1140/epjd/e2007-00327-x

THE EUROPEAN PHYSICAL JOURNAL D

α-helix↔random coil phase transition: analysis of ab initio theory predictions

I.A. Solov'yov[a], A.V. Yakubovich[a], A.V. Solov'yov[a,b], and W. Greiner

Frankfurt Institute for Advanced Studies, Ruth-Moufang Str. 1, 60438 Frankfurt am Main, Germany

Received 11 May 2007 / Received in final form 8 October 2007

Abstract. In the present paper we present results of calculations obtained with the use of the theoretical method described in our preceding paper [Eur. Phys. J. D, DOI: 10.1140/epjd/e2007-00328-9] and perform detail analysis of α-helix↔random coil transition in alanine polypeptides of different length. We have calculated the potential energy surfaces of polypeptides with respect to their twisting degrees of freedom and construct a parameter–free partition function of the polypeptide using the suggested method [Eur. Phys. J. D, DOI: 10.1140/epjd/e2007-00328-9]. From the build up partition function we derive various thermodynamical characteristics for alanine polypeptides of different length as a function of temperature. Thus, we analyze the temperature dependence of the heat capacity, latent heat and helicity for alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. Alternatively, we have obtained same thermodynamical characteristics from the use of molecular dynamics simulations and compared them with the results of the new statistical mechanics approach. The comparison proves the validity of the statistical mechanic approach and establishes its accuracy.

PACS. 82.60.Fa Heat capacities and heats of phase transitions – 87.15.He Dynamics and conformational changes – 64.70.Nd Structural transitions in nanoscale materials – 64.60.-i General studies of phase transitions

1 Introduction

In our preceding paper [1], we introduced a novel and general theoretical method for the description of phase transitions in finite complex molecular systems. In particular, we have demonstrated that for polypeptide chains, i.e. chains of amino acids, one can identify specific twisting degrees of freedom that are responsible for the folding dynamics of these amino acid chains. In other words, these degrees of freedom characterize the transition from a chain in a random coil state, to one in an α-helix structure and vice versa.

The essential domains of the potential energy surface (PES) of polypeptides with respect to these twisting degrees of freedom have been calculated and thoroughly analyzed on the basis of both classical molecular dynamics (MD) simulations, and ab initio methods such as density functional theory (DFT) and the Hartree-Fock approach. In references [1–3], it was shown that with the PES, one can construct a partition function of a polypeptide chain from which it is then possible to extract all essential thermodynamical variables and properties, such as the heat capacity, phase transition temperature, free energy, etc.

In this paper, we explore this further using a formalism we introduced previously [1] and apply it to a detailed analysis of the α-helix↔random coil phase transition in alanine polypeptides of different lengths. We have chosen this system because it has been widely investigated both theoretically [4–20] and experimentally [21–24] during the last five decades (for review see, e.g. [25–28]) and thus is perfect system for testing a novel theoretical approach.

The theoretical studies of the helix-coil transition in polypeptides have been performed both with the use of statistical mechanics methods [4–9,16–20,27] and of MD [12–16]. Previous attempts to describe the helix-coil transition in polypeptide chains using the principles of of statistical mechanics were based on the models suggested in sixties [4–7]. These models were based on the construction of the polypeptide partition function depending on several parameters and were widely used in references [16–20,25–27] for the description of the helix-coil transition in polypeptides.

For a comprehensive overview of the relevant work we refer to recent reviews [25,27,28] and the book [26], as well as to our preceding paper [1].

Experimentally, extensive studies of the helix-coil transition in polypeptides have been conducted [21–24].

---

[a] On leave from the A.F. Ioffe Institute, St. Petersburg, Russia.

[b] e-mail: `solovyov@fias.uni-frankfurt.de`

In reference [21], the enthalpy change of an α-helix to random coil transition for the Ac-Y(AEAAKA)$_8$F-NH$_2$ peptide in water was determined calorimetrically. The dependence of the heat capacity of the polypeptide on temperature was measured using differential scanning calorimetry. In references [22,23], UV resonance Raman spectroscopy was performed on the MABA-[A]$_5$-[AAARA]$_3$-ANH$_2$ peptide. Using circular dichroism methods, the dependence of helicity on temperature was measured. While in reference [24], the kinetics of the helix-coil transition of the 21-residue alanine polypeptide was investigated by means of infrared spectroscopy.

In this work, we have calculated the PES of polyalanines of different lengths with respect to their twisting degrees of freedom. This was done within the framework of classical molecular mechanics. However, to scrutinize the accuracy of these calculations, we compared the resultant molecular mechanics potential energy landscapes with those obtained using ab initio density functional theory (DFT). The comparison was only performed for alanine tripeptide and hexapeptide, since for larger polypeptides, the DFT calculation becomes increasingly computationally demanding. Hence for these larger systems, only molecular mechanics simulations have been used in this present work.

The calculated PES was then used to construct a parameter–free partition function of the polypeptide using the statistical method we had outlined in our preceding paper [1]. This partition function was then used to derive various thermodynamical characteristics of alanine polypeptides as a function of temperature and polypeptide length. We have calculated and analyzed the temperature dependence of the heat capacity, latent heat and helicity of alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. We have also established a correspondence between our ab initio method with the results of the semiempirical approach of Zimm and Bragg [4]. Thus, on the basis of our approach, we have determined the key parameters of the Zimm-Bragg theory that itself utilizes principles of statistical mechanics.

Finally, we have calculated the heat capacity, latent heat and helicity of alanine polypeptides using molecular dynamics and have compared the obtained results with those using our statistical approach. Comparison between the two methods allows us to establish the accuracy of our statistical method for relatively small molecular systems, and lets us gauge the feasibility of extending the description to larger molecular objects for which it is especially essential in those cases where MD simulations are hardly possible due to computational limitations.

Our paper is organized as follows. In Section 2 we present the final expressions obtained within the formalism described in our preceding paper [1] and introduce basic equations and the set of parameters which have been used in MD calculations. In Section 3 we present and discuss the results of computer simulations obtained with the use of developed theoretical method and compare then with results of MD simulations. In Section 4, we draw a conclusion to this paper.

2 Theoretical methods

2.1 Statistical model for the α-helix$\leftrightarrow$random coil phase transition

Our calculations have been performed using the statistical formalism we described previously [1]. Here, we will only outline the basic ideas of this method and present the final expressions that were used in our investigation.

Let us consider a polypeptide, consisting of n amino acids. The polypeptide can be found in one of its numerous isomeric states with different energies. A group of isomeric states with similar characteristic physical properties is called a *phase state* of the polypeptide. Thus, a regular bounded α-helix state corresponds to one phase state of the polypeptide, while all possible unbounded random conformations can be denoted as the random coil phase state.

The *phase transition* is then a transformation of the polypeptide from one phase state to another, i.e. the transition from a regular α-helix to a random coil conformation.

All thermodynamical properties of a molecular system are described by its partition function. The partition function of a polypeptide can be expressed as [1]:

$$\begin{aligned}\mathbb{Z} = AB(kT)(kT)^{3N-3-\frac{l_s}{2}}\Big[\beta Z_b^{n-1} Z_u \\ +\beta\sum_{i=4}^{n-2}(n-i)Z_b^i Z_u^{n-i} + Z_u^n \\ + \sum_{i=2}^{(n-3)/2}\beta^i \sum_{k=i}^{n-i-3}\frac{(k-1)!(n-k-3)!}{i!(i-1)!(k-i)!(n-k-i-3)!} \\ \times Z_b^{k+3i} Z_u^{n-k-3i}\Big]. \qquad (1)\end{aligned}$$

Here the first and the third terms in the square brackets describe, respectively, the partition function of the polypeptide in the α-helix and the random coil phases. The second term in the square brackets accounts for the situation of phase co-existence. The summation in this term is performed from $i = 4$ as the shortest α-helix has only 4 amino acids. The final term in the square brackets accounts for the polypeptide conformations in which a number of amino acids in the α-helix conformation are separated by amino acids in the random coil conformation. The first summation in this term goes over the separated helical fragments of the polypeptide, while the second summation goes over individual amino acids in the corresponding fragment. Polypeptide conformations with two or more helical fragments are energetically unfavorable. This fact will be discussed in detail further on in this paper. Therefore, the fourth term in the square brackets equation (1) can be omitted in the construction of the partition function.

A in equation (1) is a factor that is determined by the specific volume, momenta of inertia and frequencies of normal vibration modes of the polypeptide in different conformations [1], l_s is the total number of the “soft”

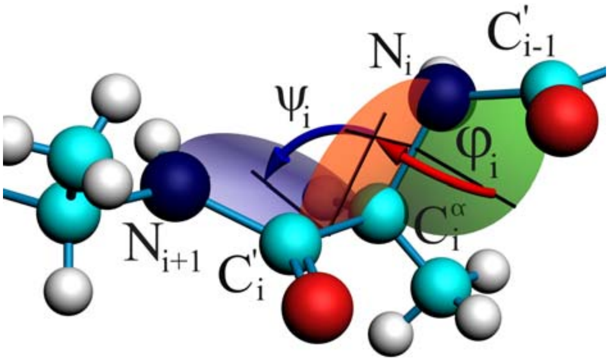

Fig. 1. (Color online) Dihedral angles φ and ψ used for characterization of the secondary structure of a polypeptide chain.

degrees of freedom in the system. $B(kT)$ is a function defined in our preceding paper [1], which describes the rotation of the side radicals in the polypeptide. Z_b and Z_u are the contributions to the partition function from a single amino acid being in the bounded or unbounded states respectively. They can be written as:

$$Z_b = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \exp\left(-\frac{\epsilon^{(b)}(\varphi,\psi)}{kT}\right) \mathrm{d}\varphi \mathrm{d}\psi \tag{2}$$

$$Z_u = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \exp\left(-\frac{\epsilon^{(u)}(\varphi,\psi)}{kT}\right) \mathrm{d}\varphi \mathrm{d}\psi, \tag{3}$$

where k and T are the Boltzmann constant and the temperature respectively, while N is the total number of atoms in the system. $\epsilon^{(b)}(\varphi,\psi)$ and $\epsilon^{(u)}(\varphi,\psi)$ in equations (2) and (3) are the potential energies of a single amino acid in the bounded and unbounded conformations calculated respectively versus the twisting degrees of freedom φ and ψ. These degrees of freedom are defined for each amino acid of the polypeptide except for the boundary ones and are described by two dihedral angels φ_i and ψ_i (see Fig. 1). β in equation (1) is the helix initiation factor, which is defined as:

$$\beta = \exp\left(-\frac{3E_{HB}}{kT}\right), \tag{4}$$

where E_{HB} is the energy of a single hydrogen bond.

Both angles, φ and ψ, are defined by four neighboring atoms in the polypeptide chain. The angle φ_i is defined as the dihedral angle between the planes formed by the atoms ($\mathrm{C}'_{i-1}-\mathrm{N}_i-\mathrm{C}^{\alpha}_i$) and ($\mathrm{N}_i-\mathrm{C}^{\alpha}_i-\mathrm{C}'_i$). While the angle ψ_i is defined as the dihedral angle between the ($\mathrm{N}_i-\mathrm{C}^{\alpha}_i-\mathrm{C}'_i$) and ($\mathrm{C}^{\alpha}_i-\mathrm{C}'_i-\mathrm{N}_{i+1}$) planes. The atoms are numbered from the NH_2– terminal of the polypeptide and φ_i and ψ_i take all possible values within the interval [−180°;180°]. For an unambiguous definition most commonly used [29–33], φ_i and ψ_i are counted clockwise if one looks on the molecule from its NH_2– terminal (see Fig. 1).

By substituting equations (2) and (3) into equation (1), one obtains the final expression for the partition function of a polypeptide experiencing an α-helix↔random coil phase transition. This is the expression which we then use to evaluate all thermodynamical characteristics of our polypeptide system.

2.2 Molecular dynamics

Molecular dynamics (MD) is an alternative approach which can be used for the study of phase transitions in macromolecular systems. Within the framework of MD, one tries to solve the equations of motion for all particles in the system interacting via a given potential. Since the technique of MD is well-known and described in numerous textbooks [34–36], we will only present the basic equations and ideas underlying this method.

MD simulations usually imply the numerical solution of the Langevin equation [36–38]:

$$m_i\mathbf{a}_i = m_i\ddot{\mathbf{r}}_i = -\frac{\partial U(\mathbf{R})}{\partial \mathbf{r}_i} - \beta_i\mathbf{v}_i + \boldsymbol{\eta}(t). \tag{5}$$

Here m_i, $\mathbf{r}_i$, $\mathbf{v}_i$ and $\mathbf{a}_i$ are the mass, radius vector, velocity and acceleration of the atom i. $U(\mathbf{R})$ is the potential energy of the system. The second term describes the viscous force which is proportional to the particle velocity. The proportionality constant $\beta_i = m_i\gamma$, where γ is the damping coefficient. The third term is the noise term that represents the effect of a continuous series of collisions of the molecule with the atoms in the medium. To study the time-evaluation of the system, the Langevin equations of motion, equation (5), are integrated for each particle.

In this paper, we use the MD approach to study the α-helix↔random coil phase transition in alanine polypeptides and compare the results with those obtained using the statistical mechanics approach. For the simulations, we use the CHARMM27 force field [39] to describe the interactions between atoms. This is a common empirical field for treating polypeptides, proteins and lipids [39–43].

MD simulations allow one to study the α-helix↔random coil phase transition of alanine polypeptide as this process occurs on the nanosecond time scale. From these simulations, one can obtain the important characteristics of the phase transition, such as the transition temperature, maximal heat capacity, the temperature range of the transition and the latent heat.

We perform MD simulations of alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. For this study it is necessary to specify the initial conditions for the system, i.e. to define the initial positions of all atoms and set their initial velocities. We assume the initial structure of the polypeptides as an ideal α-helix [26,44,45] and assign the particle velocities randomly according to the Maxwell distribution at a given temperature.

The MD simulations of the polypeptides were performed at different temperatures. For an alanine polypeptide consisting of 21 amino acids, 71 simulations were performed for the temperatures in the region of 300 K to 1000 K. For polypeptides consisting of 30, 40, 50 and 100 amino acids, 31 simulations were performed for each polypeptide in the temperature region of 300 K to 900 K. The simulations were carried out within a time interval of 100 ns and an integration step of 2 fs. The first 25 ns of the simulation were used to equilibrate the system, while the next 75 ns were used for obtaining data about the energy and structure of the system at a given temperature.

The set of the parameters used in our simulations can be found in references [34–36]. All simulations were performed using the NAMD molecular dynamics program [35], while visualization of the results was done with VMD [46]. The covalent bonds involving hydrogen atoms were considered as rigid. The damping coefficient γ was set to 5 ps^{-1}. The simulations were performed in the NVT canonical ensemble using a Langevin thermostat with no cutoff for electrostatic interactions.

3 Results and discussion

In this section we present the results of calculations obtained using our statistical mechanics approach and those from our MD simulations. In Section 3.1 we discuss the accuracy of this force field as applied to alanine polypeptides. In Section 3.2 we present the PESs for different amino acids in alanine polypeptide calculated versus the twisting degrees of freedom φ and ψ (see Fig. 1). In Section 3.3, the statistical mechanics approach is used for the description of the α-helix↔random coil phase transition. Here, the results of the statistical mechanics approach are compared to those obtained from MD simulations. In Section 3.4 the statistical independence of amino acids in the polypeptide is discussed.

3.1 Accuracy of the molecular mechanics potential

The PES of alanine polypeptides was calculated using the CHARMM27 force field [39] that has been parameterized for the description of proteins, in particular polypeptides, and lipids. Nevertheless, the level of its accuracy when applied to alanine polypeptides cannot be taken for granted and has to be investigated. Therefore, we compare the PESs for alanine tri- and hexapeptide calculated using the CHARMM27 force field with those calculated using ab initio density functional theory (DFT). In the DFT approach, the PES of alanine tri- and hexapeptides were calculated as a function of the twisting degrees of freedom, φ and ψ (see Fig. 1), in the central amino acid of the polypeptide [32]. All other degrees of freedom were frozen.

To establish the accuracy of the CHARMM27 force field, we have calculated the PESs of alanine polypeptides in its β-sheet conformation. The geometry of alanine tri- and hexapeptide used in the calculations are shown in Figures 2a and 2b respectively. The ab initio calculations were performed [32] using B3LYP, Becke's three-parameter gradient-corrected exchange functional [47] with the gradient-corrected correlation functional of Lee et al. [48]. The wave function of all electrons in the system was expanded using a standard basis set B3LYP/6-31G(2d,p). The PESs calculated within the DFT approach have been analyzed in reference [32].

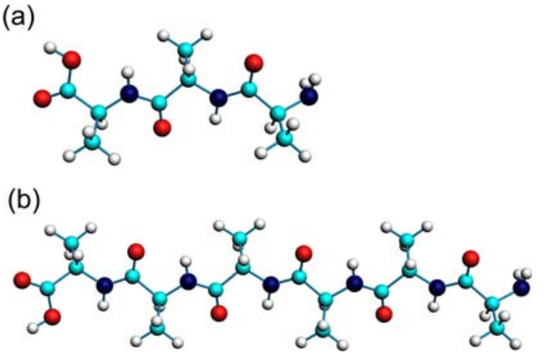

Fig. 2. (Color online) Optimized geometries of alanine polypeptide chains: (a) Alanine tripeptide; (b) Alanine hexapeptide in the β-sheet conformation.

The difference between the PESs calculated with the CHARMM27 force field and with the B3LYP functional is shown in Figure 3 for the alanine tripeptide (left plot) and for the alanine hexapeptide (right plot).

From Figure 3, we can see that the energy difference between the PESs calculated with the CHARMM27 force field and with the B3LYP functional is less than 0.15 eV. To describe the relative deviation of the PESs, we introduce the relative error of the two methods as follows:

$$\eta = \frac{2\int |E_{\mathrm{B3L}}(\varphi,\psi) - E_{\mathrm{CH27}}(\varphi,\psi)|\mathrm{d}\varphi\mathrm{d}\psi}{\int |E_{\mathrm{B3L}}(\varphi,\psi) + E_{\mathrm{CH27}}(\varphi,\psi)|\mathrm{d}\varphi\mathrm{d}\psi} 100\%, \qquad (6)$$

where $E_{\mathrm{B3L}}(\varphi,\psi)$ and $E_{\mathrm{CH27}}(\varphi,\psi)$ are the potential energies calculated within the DFT and molecular mechanics methods respectively. Calculating η for alanine tri- and hexapeptide, one obtains: $\eta_{3\times\mathrm{Ala}} = 27.6\%$ and $\eta_{6\times\mathrm{Ala}} = 23.4\%$ respectively. These values show that the molecular mechanics approach is reasonable for a qualitative description of the alanine polypeptide. Note however, that the PES obtained for alanine hexapeptide within the molecular mechanics method is closer to the PES calculated within the DFT approach. This occurs because the PESs $E_{\mathrm{CH27}}(\varphi,\psi)$ and $E_{\mathrm{B3L}}(\varphi,\psi)$ of alanine hexapeptide were calculated for the structure optimized within the DFT approach, while the PESs E_{CH27} and E_{B3L} of alanine tripeptide were calculated for the structure optimized within the molecular mechanics method and the DFT approach respectively.

Our analysis shows that the molecular mechanics potential can be used to describe qualitatively the structural and dynamical properties of alanine polypeptides with an error of about 20%. In the present paper, we have calculated the thermodynamical properties of alanine polypeptides with the use of MD method and compared the obtained results with those attained from the statistical approach. However, ab initio MD calculations of alanine polypeptides are hardly possible on the time scales when the α-helix↔random coil phase transition occurs, even for systems consisting of only 4–5 amino acids [30–33,49]. Therefore, we have performed MD simulations for alanine polypeptides using molecular mechanics forcefield. In order to establish the accuracy of the statistical mechanics approach, the PES used for the construction

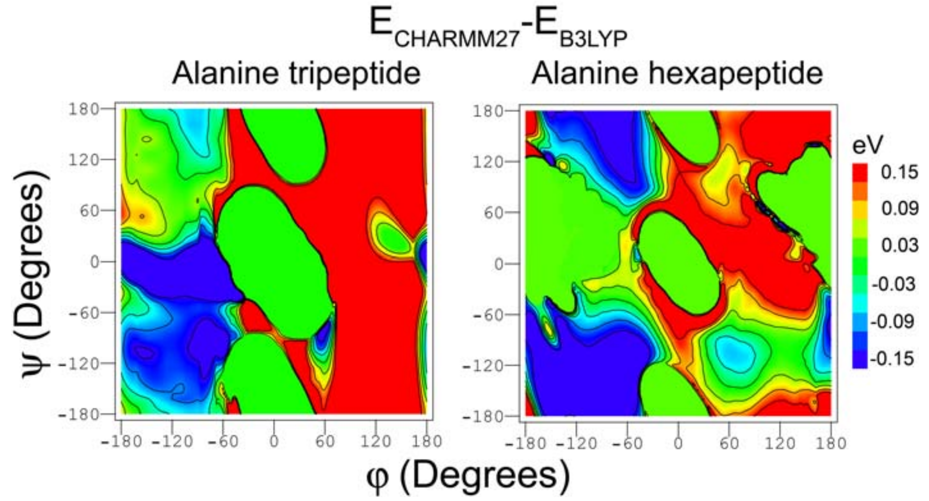

Fig. 3. (Color online) Difference between the PESs calculated with the CHARMM27 force field and with the B3LYP functional [32] for the alanine tripeptide (left) and the alanine hexapeptide (right). The relative energies are given in eV. The equipotential lines are shown for the energies –0.10, –0.05, 0, 0.05 and 0.1 eV.

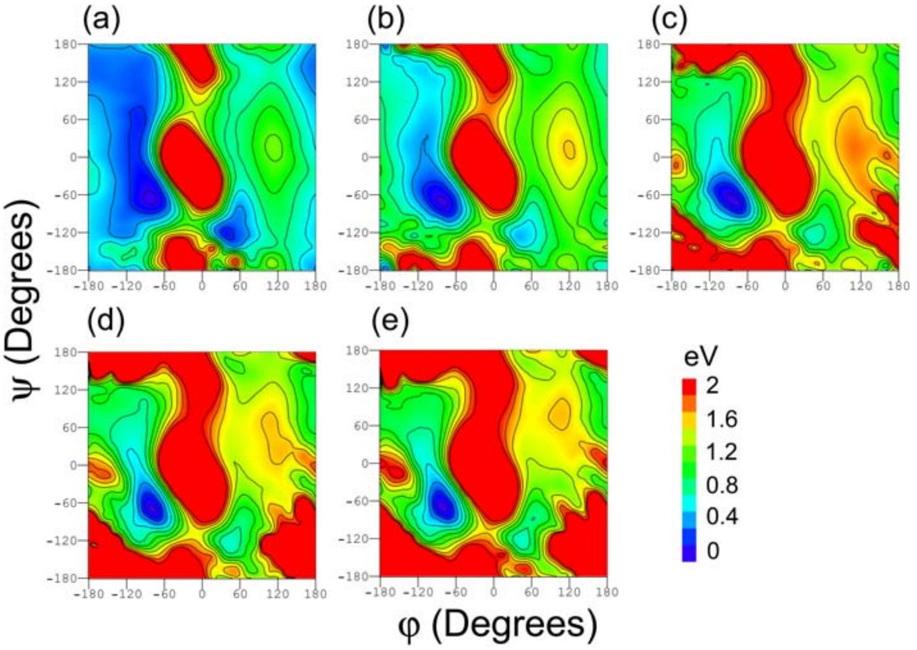

Fig. 4. (Color online) PESs for different amino acids of alanine polypeptide consisting of 21 amino acids calculated as the function of twisting dihedral angles φ and ψ in: (a) second alanine, (b) third alanine, (c) fourth alanine (d) fifth alanine and (e) tenth alanine. Amino acids are numbered starting from the NH$_2$ terminal of the polypeptide. Energies are given with respect to the lowest energy minimum of the PES in eV. The equipotential lines are shown for the energies 1.8, 1.6, 1.4, 1.2, 1.0, 0.8, 0.6, 0.4 and 0.2 eV.

of the partition function was also calculated with the same method.

3.2 Potential energy surface of alanine polypeptide

To construct the partition function equation (1), one needs to calculate the PES of a single amino acid in the bounded, $\epsilon^{(b)}(\varphi,\psi)$, and unbounded, $\epsilon^{(u)}(\varphi,\psi)$, conformations versus the twisting degrees of freedom φ and ψ (see Fig. 1). The potential energies of alanine in different conformations determine the Z_b and Z_u contributions to the partition function, defined in equations (2) and (3).

The PES of an alanine depends both on the conformation of the polypeptide and on the amino acid index in the chain. The PES for different amino acids of the 21-residue alanine polypeptide calculated as a function of twisting dihedral angles φ and ψ are shown in Figure 4. These surfaces were calculated with the use of the CHARMM27 forcefield for a polypeptide in the α-helix conformation. The PESs (a), (b), (c), (d) and (e) in Figure 4 correspond to the variation of the twisting angles in the second, third,

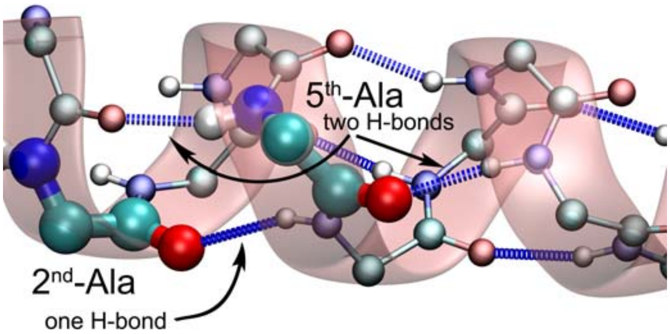

Fig. 5. (Color online) Alanine polypeptide in the α-helix conformation. Dashed lines show the hydrogen bonds in the system. Figure shows that the second alanine forms only one hydrogen bond, while the fifth alanine forms two hydrogen bonds with the neighboring amino acids.

fourth, fifth and tenth amino acids of the polypeptide respectively. Amino acids are numbered starting from the NH_2 terminal of the polypeptide. We do not present the PES for the amino acids at boundary because the angle φ is not defined for it.

On the PES corresponding to the tenth amino acid in the polypeptide (see Fig. 4e), one can identify a prominent minimum at $\varphi = -81°$ and $\psi = -71°$. This minimum corresponds to the α-helix conformation of the corresponding amino acid, and energetically, the most favorable amino acid configuration. In the α-helix conformation the tenth amino acid is stabilized by two hydrogen bonds (see Fig. 5). With the change of the twisting angles φ and ψ, these hydrogen bonds become broken and the energy of the system increases. The tenth alanine can form hydrogen bonds with the neighboring amino acids only in the α-helix conformation, because all other amino acids in the polypeptide are in this particular conformation. This fact is clearly seen from the corresponding PES Figure 4e, where all local minima have energies significantly higher than the energy of the global minima (the energy difference between the global minimum and a local minimum with the closest energy is $\Delta E = 0.736$ eV, which is found at $\varphi = 44°$ and $\psi = -124°$).

The PES depends on the amino acid index in the polypeptide. This fact is clearly seen from Figure 4. The three boundary amino acids in the polypeptide form a single hydrogen bond with their neighbors (see Fig. 5) and therefore are more weakly bounded than the amino acids inside the polypeptide. The change in the twisting angles φ and ψ in the corresponding amino acids leads to the breaking of these bonds, hence increasing the energy of the system. However, the boundary amino acids are more flexible then those inside the polypeptide chain, and therefore their PES is smoother.

Figure 4 shows that the PESs calculated for the fourth, fifth and the tenth amino acids are very close and have minor deviations from each other. Therefore, the PESs for all amino acids in the polypeptide, except the boundary ones can be considered identical.

Each amino acid inside the polypeptide forms two hydrogen bonds. However since these bonds are shared by two amino acids, there is only effectively one hydrogen bond per amino acid (see Fig. 5). Therefore, to determine the potential energy surface of a single amino acid in the bounded, $\epsilon^{(b)}(\varphi,\psi)$, and unbounded, $\epsilon^{(u)}(\varphi,\psi)$, conformations, we use the potential energy surface calculated for the second amino acid of the alanine polypeptide (see Fig. 4a), because only this amino acid forms single hydrogen bond with its neighbors (see Fig. 5).

The PES of the second amino acid Figure 4a has a global minima at $\varphi = -81°$ and $\psi = -66°$, and corresponds to the bounded conformation of the alanine. Therefore the part of the PES in the vicinity of this minima corresponds to the PES of the bounded state of the polypeptide, $\epsilon^{(b)}(\varphi,\psi)$. The potential energy of the bounded state is determined by the energy of the hydrogen bond, which for an alanine is equal to $E_{HB} = 0.142$ eV. This value is obtained from the difference between the energy of the global minima and the energy of the plateaus at $\varphi \in (-90°... - 100°)$ and $\psi \in (0°...60°)$ (see Fig. 4a). Thus, the part of the potential energy surface which has an energy less then E_{HB} corresponds to the bounded state of alanine, while the part with energy greater then E_{HB} corresponds to the unbounded state.

In Figure 6 we present the potential energy surfaces for alanine in both the bounded (plot a) and unbounded (plot b) conformations. Both PESs were calculated from the PES for the second amino acid in the polypeptide, which is shown in plot (c) of Figure 6.

3.3 α-helix↔random coil phase transition in alanine polypeptide

3.3.1 Internal energy of alanine polypeptide

Knowing the PES for all amino acids in the polypeptide, one can construct the partition function of the system using from equation (1). Plots (a) and (b) in Figure 6 show the dependence of $\epsilon^{(b)}(\varphi,\psi)$ and $\epsilon^{(u)}(\varphi,\psi)$ on the twisting angles φ and ψ, while $\epsilon^{(b)}$ and $\epsilon^{(u)}$ define the contributions of the bounded and unbounded states of the polypeptide to the partition function of the system (see Eqs. (2) and (3)). The expressions for Z_b and Z_u are integrated numerically and the partition function of the polypeptide is evaluated according to equation (1). The partition function defines all essential thermodynamical characteristics of the system as discussed in reference [1].

The first order phase transition is characterized by an abrupt change of the internal energy of the system with respect to its temperature. In the first order phase transition the system either absorbs or releases a fixed amount of energy while the heat capacity as a function of temperature has a pronounced peak [26,28,29,50]. We study the manifestation of these peculiarities for alanine polypeptide chains of different lengths.

Figure 7 shows the dependencies of the internal energy on temperature calculated for alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. The thick solid lines correspond to the results obtained using the statistical approach, while the dots show the results of

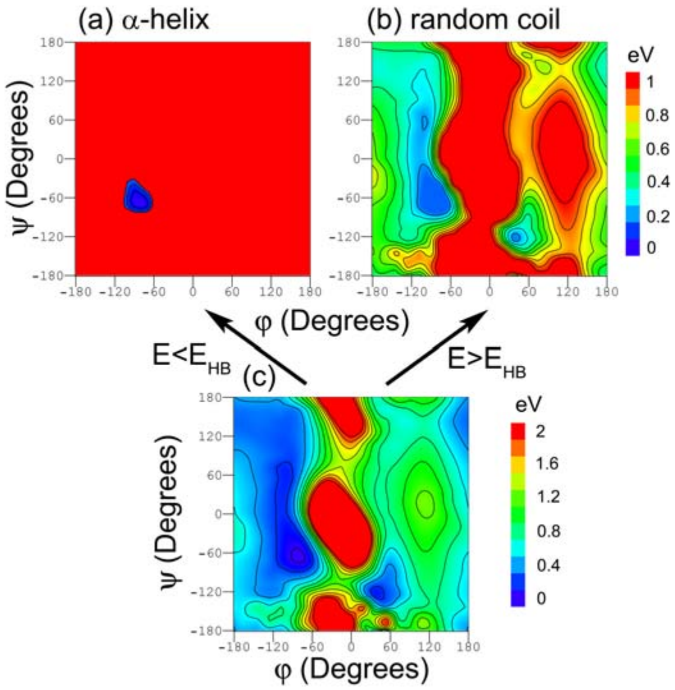

Fig. 6. (Color online) PESs for alanine in α-helix (plot a) and random coil conformation (plot b). The potential energy surface for the second amino acid of the polypeptide is shown in plot (c) and is used to determine the PESs for alanine in α-helix and random coil conformations. The part of the PES shown in plot (c), with energy less then E_{HB} corresponds to the α-helix conformation (bounded state) of the alanine, while the part of the potential energy surface with energy greater then E_{HB} corresponds to the random coil conformation (unbounded state). The energies are given in eV. The equipotential lines in plot (a) are shown for the energies 0.05 and 0.1 and 0.15 eV; in plot (b) for the energies 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9 eV; in plot (c) for the energies 1.8, 1.6, 1.4, 1.2, 1.0, 0.8, 0.6, 0.4 and 0.2 eV.

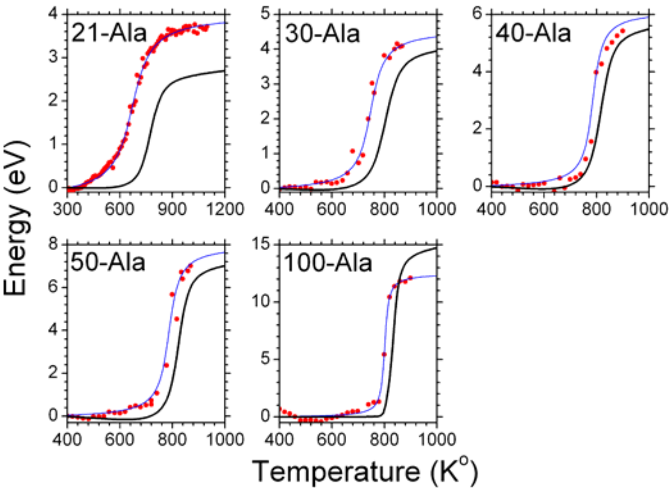

Fig. 7. (Color online) Dependencies of the internal energy on temperature calculated for the alanine polypeptide chains consisting of 21, 30, 40, 50 and 100 amino acids. Thick solid lines correspond to the results obtained within the framework of the statistical model. Dots correspond to the results of MD simulations, which are fitted using equation (7). The fitting functions are shown with thin solid lines. The fitting parameters are compiled in Table 1.

MD simulations. From Figure 7 it is seen that the internal energy of alanine polypeptide rapidly increases in the vicinity of a certain temperature corresponding to the temperature of the first order phase transition. The value of the step-like increase of the internal energy is usually referred as the the latent heat of the phase transition denoted as Q. The latent heat is the energy that the system absorbs at the phase transition. Figure 7 shows that the latent heat increases with the growth of the polypeptide length. This happens because in the α-helix state, long polypeptides have more hydrogen bonds than short ones and, for the formation of the random coil state, more energy is required.

The characteristic temperature region of the abrupt change in the internal energy (half-wight of the heat capacity peak) characterizes the temperature range of the phase transition. We denote this quantity as ΔT. With the increase of the polypeptide length the dependence of the internal energy on temperature becomes steeper and ΔT decreases. Therefore, the phase transition in longer polypeptides is more pronounced. In the following subsection we discuss in detail the dependence of ΔT on the polypeptide length.

With the molecular dynamics, one can evaluate the dependence of the total energy of the system on temperature, which is the sum of the potential, kinetic and vibrational energies. Then the heat capacity can be factorized into two terms: one, corresponding to the internal dynamics of the polypeptide and the other, to the potential energy of the polypeptide conformation. The conformation of the polypeptide influences only the term related to the potential energy and the term corresponding to the internal dynamics is assumed to be independent of the polypeptides conformation.

This factorization allows one to distinguish from the total energy the potential energy term corresponding to the structural changes of the polypeptide. The formalism of this factorization is discussed in detail in reference [1]. The energy term corresponding to the internal dynamics of the polypeptide neither influence the phase transition of the system, nor does it grow linearly with temperature. The term corresponding to the potential energy of the polypeptide conformation has a step-like dependence on temperature that occurs at the temperature of the phase transition. Since we are interested in the manifestation of the phase transition we have subtracted the linear term from the total energy of the system and consider only its non-linear part. The slope of the linear term was obtained from the dependencies of the total energy on temperature in the range of 300–450 K, which is far beyond the phase transition temperature (see Fig. 7). Note that the dependence shown in Figure 7 corresponds only to the non-linear potential energy terms.

Table 1. Parameters used in equation (7) to fit the results of MD simulations.

n	E_0 (eV)	ΔE (eV)	γ (K)	T_0 (K)	a (eV/K)
21	11.38 ± 0.23	4.30 ± 0.10	79.37 ± 7.62	670.0 ± 2.0	0.0471 ± 0.0003
30	13.61 ± 0.58	4.70 ± 0.16	37.92 ± 7.31	747.4 ± 3.3	0.0699 ± 0.0008
40	16.80 ± 0.39	6.261 ± 0.083	26.59 ± 2.25	785.7 ± 1.8	0.0939 ± 0.0005
50	19.94 ± 0.79	8.15 ± 0.21	29.36 ± 5.51	786.6 ± 2.9	0.1178 ± 0.0010
100	29.95 ± 0.67	12.58 ± 0.16	10.49 ± 2.00	801.1 ± 1.1	0.2437 ± 0.0009

Table 2. Parameters, characterizing the heat capacity peak in Figure 8 calculated using the statistical approach. Heat capacity at 300 K, C_{300}, the transition temperature T_0, the maximal value of the heat capacity C_0, the temperature range of the phase transition ΔT and the specific heat Q are shown as a function of polypeptide length, n.

n	C_{300} (meV/K)	T_0 (K)	C_0 (eV/K)	ΔT (K)	Q (eV)
21	2.034	770	0.020	134	2.70
30	2.840	805	0.037	96	4.15
40	3.736	815	0.060	86	5.77
50	4.631	825	0.080	66	7.38
100	9.111	835	0.308	35	15.47

The heat capacity of the system is defined as the derivative of the total energy on temperature. However, as seen from Figure 7 the MD data is scattered in the vicinity of a certain expectation line. Therefore, the direct differentiation of the energy obtained within this approach will lead to non-physical fluctuations of the heat capacity. To overcome this difficulty we define a fitting function for the total energy of the polypeptide:

$$E(T) = E_0 + \frac{\Delta E}{\pi} \arctan\left[\frac{T - T_0}{\gamma}\right] + aT, \qquad (7)$$

where E_0, ΔE, T_0, γ and a are the fitting parameters. The first and the second terms are related to the potential energy of the polypeptide conformation, while the last term describes the linear increase of the total energy with temperature. The fitting function equation (7) was used for the description of the total energy of polypeptides in earlier papers [15,51]. The results of fitting are shown in Figure 7 with the thin solid lines. The corresponding fitting parameters are compiled in Table 1.

Figure 7 shows that the results obtained using the MD approach are in a reasonable agreement with the results obtained from the the statistical mechanics formalism. The fitting parameter ΔE corresponds to the latent heat of the phase transition, while the temperature width of the phase transition is related to the parameter γ. With the increase of the polypeptides length, the temperature width of the phase transition decreases (see γ in Tab. 1), while the latent heat increases (see ΔE in Tab. 1). These features are correctly reproduced in MD and in our statistical mechanics approach.

Furthermore, MD simulations demonstrate that with an increase of the polypeptide length, the temperature of the phase transition shifts towards higher temperatures (see Fig. 7). The temperature of the phase transition is described by the fitting parameter T_0 in Table 1. Note also, that the increase of the phase transition temperature is reproduced correctly within the framework of the statistical mechanics approach, as seen from Figure 7.

Nonetheless, the results of MD simulations and the results obtained using the statistical mechanics formalism have several discrepancies. As seen from Figure 7 the latent heat of the phase transition for Ala_{100} polypeptide obtained within the framework of the statistical approach is higher than that obtained in MD simulations. This happens because within the statistical mechanics approach, the potential energy of the polypeptide is underestimated. Indeed, long polypeptides (consisting of more than 50 amino acids) tend to form short-living hydrogen bonds in the random coil conformation. These hydrogen bonds lower the potential energy of the polypeptide in the random coil conformation. However, the "dynamic" hydrogen-bonds are neglected in the present formalism of the partition function construction.

Additionally, the discrepancies between the two methods arise due to the limited MD simulation time and to the small number of different temperatures at which the simulations were performed. Indeed, for alanine polypeptide consisting of 100 amino acids 26 simulations were performed, while only 3–5 simulations correspond to the phase transition temperature region (see Fig. 7).

3.3.2 Heat capacity of alanine polypeptide

The dependence of the heat capacity on temperature for alanine polypeptides of different lengths is shown in Figure 8. The results obtained using the statistical approach are shown with the thick solid line, while the results of MD simulations are shown with the thin solid line. Since the classical heat capacity is constant at low temperatures, we subtract out this constant value of the for a better analysis of the phase transition in the system. We denote the constant contribution to the heat capacity as C_{300} and calculate it as the heat capacity value at 300 K. The C_{300} values for alanine polypeptides of different length are compiled in the second column of Table 2.

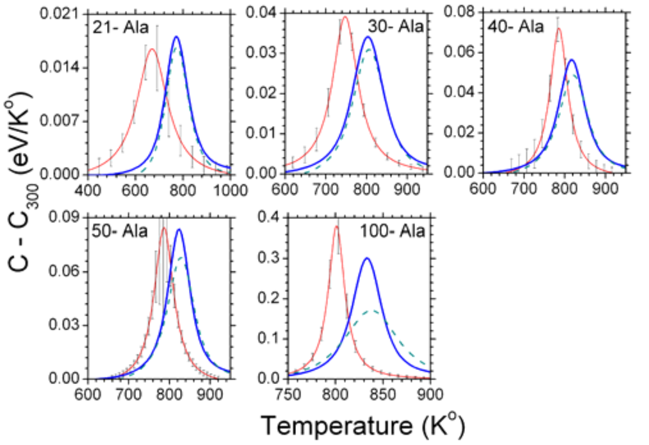

Fig. 8. (Color online) Dependencies of the heat capacity on temperature calculated for the alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. The results obtained using the statistical approach are shown with the thick solid line, while the results of MD simulations are shown with the thin solid line. Dashed lines show the heat capacity as a function of temperature calculated within the framework of the Zimm-Bragg theory [4]. C_{300} denotes the heat capacity at 300 K, which are compiled in Table 2.

As seen from Figure 8, the heat capacity of the system as a function of temperature acquires a sharp maximum at a certain temperature corresponding to the temperature of the phase transition. The peak in the heat capacity is characterized by the transition temperature T_0, the maximal value of the heat capacity C_0, the temperature range of the phase transition ΔT and the latent heat of the phase transition Q. These parameters have been extensively discussed in our preceding paper [1]. Within the framework of the two-energy level model describing the first order phase transition, it is shown that:

$$\begin{aligned} T_0 &\sim \frac{\Delta E}{\Delta S} = \text{const} \\ C_0 &\sim \Delta S^2 \sim n^2 \\ Q &\sim \Delta E \sim n \\ \Delta T &\sim \frac{\Delta E}{\Delta S^2} \sim \frac{1}{n}. \end{aligned} \tag{8}$$

Here ΔE and ΔS are the energy and the entropy changes between the α-helix and the random coil states of the polypeptide, while n is the number of amino acids in the polypeptide. Figure 9 shows the dependence of the α-helix↔random coil phase transition characteristics on the length of the alanine polypeptide. The maximal heat capacity C_0 and the temperature range of the phase transition ΔT are plotted against the squared number of amino acids (n^2) and the inverse number of amino acids ($\frac{1}{n}$) respectively, while the temperature of the phase transition T_0 and the latent heat of the phase transition Q are plotted against the number of amino acids (n). Squares and triangles represent the phase transition parameters calculated using the statistical approach and those obtained from the MD simulations respectively.

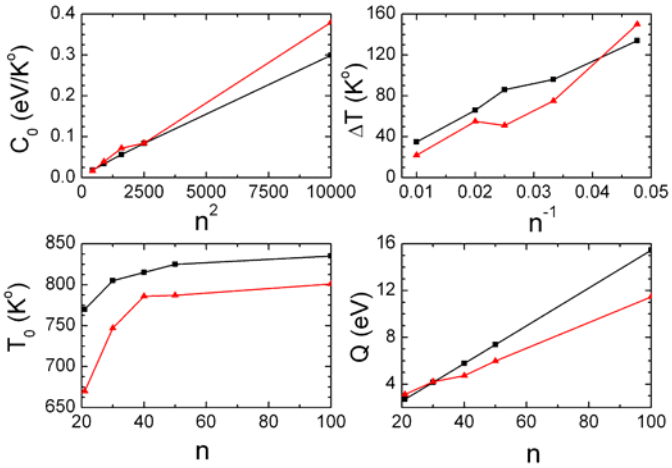

Fig. 9. (Color online) Phase transition parameters C_0, ΔT, T_0 and Q calculated as a function of polypeptide length. Squares and triangles represent the phase transition parameters calculated using the statistical approach and those obtained from the MD simulations respectively.

The results obtained within the framework of the statistical model are in a good agreement with the results obtained on the basis of MD simulations. However, since the MD simulations are computationally time demanding it is difficult to simulate phase transition in large polypeptides. The difficulties arise due to the large fluctuations which appear in the system at the phase transition temperature and to the large time scale of the phase transition process. The relative error of the phase transition temperature obtained on the basis of MD approach is in the order of 3–5%, while the relative error of the heat capacity is about 30% in the vicinity of the phase transition (see Fig. 8).

At present, there are no experiments devoted to the study of phase transition of alanine polypeptides in vacuo, but such experiments are feasible and are already planned[1]. In reference [19] the temperature of the α-helix↔random coil phase transition was calculated. Depending on the parameter set, the temperature of the transition ranges from 620 K to 650 K for right-handed α-helix, and from 730 K to 800 K for a left-handed α-helix.

In our previous work [2] on to the theoretical study of phase transitions in polypeptide chains, we have introduced the basic ideas of a theoretical method which we have described in detail in reference [1] and which we currently apply in this work. It was shown that the PES calculated as a function of twisting degrees of freedom φ and ψ determines the partition function of the system. To illustrate our method, we used the PES calculated for alanine hexapeptide within the framework of the ab initio density functional theory [2] and obtained the phase transition temperature equal to 300 K. On the other hand, in this paper we established that the phase transition temperature of alanine polypeptide in vacuo is 770 K. This is because in reference [2] the PES was calculated for alanine from the hexapeptide. The hydrogen bonds which stabilize

---

[1] Helmut Haberland, Private communication.

the α-helix structure of the hexapeptide are impaired and therefore the PES of a single alanine is smoother compared to a long polypeptide where every amino acid forms two hydrogen bonds. The smoothing of the potential energy surface results in lowering of the energy barriers and the phase transition temperature.

Nonetheless, smoothing of the PES of an alanine should happen in solution, as the effective number of hydrogen bonds in the polypeptide decreases. This fact was demonstrated previously [2], where we compared results of our calculation with available experimental data on alanine rich peptides in water solution and observed a good correspondence of the phase transition temperature.

The heat capacity peak is asymmetric. The heat capacity at higher temperatures, beyond the heat capacity peak, is not zero and forms a plateau (see Fig. 8). The plateau is formed due to the conformations of the amino acids with larger energies [2]. At $T = 1000$ K, the difference in the heat capacity of the polypeptide is 10.0×10^{-4}, 16.0×10^{-4}, 22.3×10^{-4}, 28.3×10^{-4} and 54.7×10^{-4} eV/K for the Ala_{21}, Ala_{30}, Ala_{40}, Ala_{50} and Ala_{100} peptides respectively. The magnitude of the plateau increases with the growth of the polypeptide length. This happens because the number of energy levels with high energies rapidly increases for longer polypeptide chains.

3.3.3 Calculation of the Zimm-Bragg parameters

An alternative theoretical approach for the study of α-helix$\leftrightarrow$random coil phase transition in polypeptides was introduced by Zimm and Bragg [4]. It is based on the construction of the partition function of a polypeptide involving two parameters s and σ, where s describes the contribution of a bounded amino acid relative to that of an unbounded one, and σ describes the entropy loss caused by the initiation of the α-helix formation.

The Zimm-Bragg theory [4] is semiempirical because it is parameter dependent. The theoretical method described in our preceding paper [1] and which we use in the present paper is different as it does not include any parameters and the construction of the partition function is based solely on the PES of a polypeptide. Therefore, the construction of our partition function is free of any parameters, and this is what makes it different from the models suggested previously. Assuming that the polypeptide has a single helical region, the partition function derived within the Zimm-Bragg theory, reads as:

$$Q = 1 + \sigma \sum_{k=1}^{n-4} (n-k-3) s^k, \tag{9}$$

where n is the number amino acids in the polypeptide, s and σ are the parameters of the Zimm-Bragg theory. The partition function, which we use in the present paper equation (1) can be rewritten in a similar form:

$$Z = \left[1 + \beta s(T)^3 \sum_{k=1}^{n-4} (n-k-3) s(T)^k\right] Z_u^n(T). \tag{10}$$

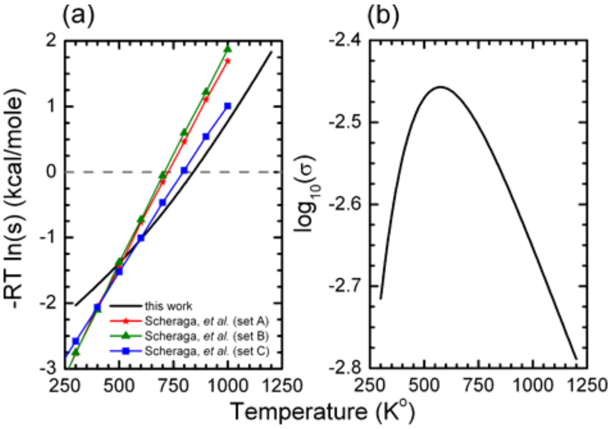

Fig. 10. (Color online) Dependence of the parameters of the Zimm-Bragg theory [4] s (plot a) and σ (plot b) on temperature. Parameter s describes the contribution to the partition function of a bounded amino acid relative to that of an unbounded one. The parameter σ describes the entropy loss caused by the initiation of the α-helix formation. Parameter s was also calculated in reference [19] using three different force fields, shown with stars, triangles and squares in plot (a).

Here n is the number of amino acids in the polypeptide and the function $s(T)$ is defined as:

$$s(T) = \frac{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \exp\left(-\frac{\epsilon^{(b)}(\varphi,\psi)}{kT}\right) \mathrm{d}\varphi \mathrm{d}\psi}{\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \exp\left(-\frac{\epsilon^{(u)}(\varphi,\psi)}{kT}\right) \mathrm{d}\varphi \mathrm{d}\psi}, \tag{11}$$

where $\epsilon^{(b)}(\varphi,\psi)$ and $\epsilon^{(u)}(\varphi,\psi)$ are the potential energies of a single amino acid in the bounded and unbounded conformations respectively calculated versus its twisting degrees of freedom φ and ψ. By comparing equations (9) and (10), one can evaluate the Zimm-Bragg parameter σ as:

$$\sigma(T) = \beta(T) s(T)^3, \tag{12}$$

where $\beta(T)$ is defined in equation (4).

The dependence of the Zimm-Bragg parameters s and σ on temperature is shown in Figures 10a and 10b respectively. The function $-RT\ln(s)$ grows monotonously with an increase in temperature, as seen in Figure 10a. The zero of this function corresponds to the temperature of the phase transition in an infinitely long polypeptide. In our calculation it is 840 K (see black line in Fig. 10a). Parameter σ is shown in the logarithmic scale and has a maximum at $T = 575$ K. Note, that this maximum does not correspond to the temperature of the phase transition.

The parameters of the Zimm-Bragg theory were considered in earlier papers [16,19,52]. In Figure 10a we present the dependence of parameter s on temperature calculated in [19] (see squares, triangles and stars in Fig. 10b) using a matrix approach described in reference [6]. The energies of different polypeptide conformations were calculated using the force field described in reference [53]. Squares, triangles and stars correspond to three different force field parameter sets used in reference [19], which are denoted as sets A, B and C. Figure 10a shows that the results of our calculations are closer to the results obtained using the parameter set C. This figure also illustrates that the Zimm-Bragg parameter s depends on the parameter set used. Therefore, the discrepancies between our calculation and the calculation performed in reference [19] arise due to the utilization of different force fields.

The Zimm-Bragg parameter σ was also calculated in reference [19]. However, it was not systematically studied for the broad range of temperatures, and therefore we do not plot it in Figure 10b. In reference [19] the parameter σ was calculated only for the temperature of the α-helix↔random coil phase transition ranging from 620 K to 800 K. In reference [19], it was also demonstrated that parameter σ is very sensitive to the force field parameters, being in the range $10^{-9.0}$–$10^{-3.6}$. In our calculation $\sigma = 10^{-2.6}$ at 860 K. The dependence of the parameter σ on the force field parameters was extensively discussed in reference [19], where it was demonstrated that this parameter does not have a strong influence on the thermodynamical characteristics of phase transition.

If the parameters s and σ are known, it is possible to construct the partition function of the polypeptide in the form suggested by Zimm and Bragg [4], and on its basis calculate all essential thermodynamic characteristics of the system. The dependence of the heat capacity calculated within the framework of the Zimm-Bragg theory is shown in Figure 8 by dashed lines for polypeptides of different length.

From Figure 8 it is seen that results obtained on the basis of the Zimm-Bragg theory are in a good agreement with the results of our statistical approach. The values of the phase transition temperature and of the maximal heat capacity in both cases are close. The comparison shows that the heat capacity obtained within the framework of the Zimm-Bragg model at temperatures beyond the phase transition window is slightly lower than the heat capacity calculated within the framework of our statistical model.

An important difference of the Zimm-Bragg theory from our theory arises due to the accounting for the states of the polypeptide with more than one α-helix fragment. These states are often referred to as multihelical states of the polypeptide. However, their statistical weight in the partition function is suppressed. The suppression arises because of entropy loss in the boundary amino acids of a helical fragment. The boundary amino acids have weaker hydrogen bonds than amino acids in the central part of the α-helix. At the same time the entropy of such amino acids is smaller than the entropy of an amino acids in the coil state. These two factors lead to the decrease of the statistical weight of the multihelical states.

The contribution of the multihelical states to the partition function leads to the broadening of the heat capacity peak while the maximal heat capacity decreases. The multihelical states become important in longer polypeptide chains that consist of more than 100 amino acids. As seen from Figure 8, the maximal heat capacity obtained within the framework of the Zimm-Bragg model for Ala$_{100}$ polypeptide is significantly lower than that obtained using our suggested statistical approach. For alanine polypeptide consisting of less than 50 amino acids the multihelical states of the polypeptide can be neglected as seen from the comparison performed in Figure 8. Omission of the multihelical states significantly simplifies the construction and evaluation of the partition function.

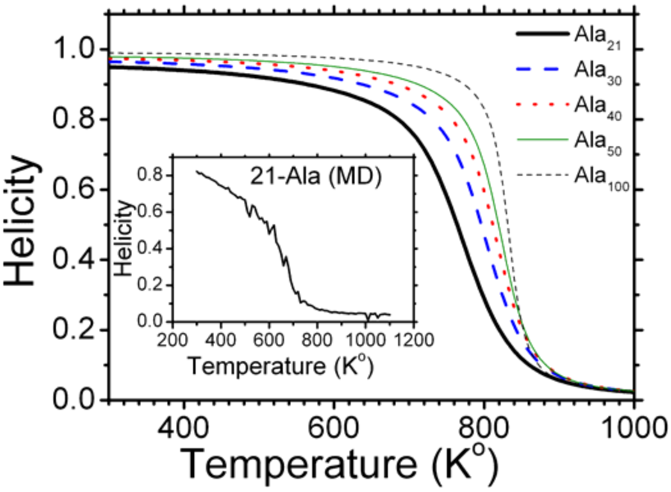

Fig. 11. (Color online) Dependency of the helicity on temperature obtained using the statistical approach for alanine polypeptide chains consisting of 21, 30, 40, 50 and 100 amino acids. The helicity for alanine polypeptide consisting of 21 amino acids obtained within a framework of MD approach is shown in the inset.

3.3.4 Helicity of alanine polypeptides

Helicity is an important characteristic of the polypeptide which can be measured experimentally [21–24]. It describes the fraction of amino acids in the polypeptide that are in the α-helix conformation. With the increase of temperature the fraction of amino acids being in the α-helix conformation decreases due to the α-helix↔random coil phase transition. In our approach, the helicity of a polypeptide is defined as follows:

$$f_\alpha = \frac{\beta \sum_{i=4}^{n} (n-i+1)(i-1) Z_b^{i-1} Z_u^{n-i+1}}{n\mathbb{Z}}, \qquad (13)$$

where n is the number of amino acids in the polypeptide, Z_b, Z_u are the contributions to the partition function from amino acids in the bounded and unbounded states defined in equations (2) and (3) respectively. $\mathbb{Z}$ is the partition function of the polypeptide defined in equation (1). The dependence of helicity on temperature obtained using the statistical approach for alanine polypeptides of different length are shown in Figure 11.

On the basis of MD simulations, it possible to evaluate the dependence of helicity on temperature. Helicity can be defined as the ratio of amino acids being in the α-helix conformation to the total number of amino acids in the polypeptide, averaged over the MD trajectory. The amino acid is considered to be in the conformation of an α-helix

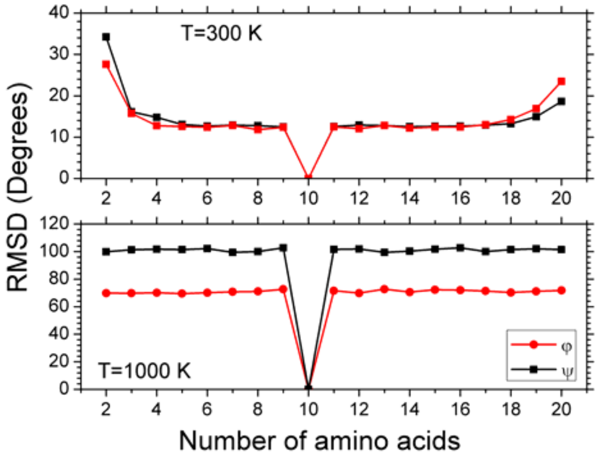

Fig. 12. (Color online) The root mean square deviation of angles φ and ψ calculated with the use of equation (14) for alanine polypeptide consisting of 21 amino acids. The calculations were done in respect to the tenth amino acid of the polypeptide for 300 K (top plot) and for 1000 K (bottom plot).

if the angles describing its twisting are within the range of $\varphi \in [-72°; -6°]$ and $\psi \in [0°; -82°]$. This region was chosen from the analysis of angles φ and ψ distribution at 300 K. The helicity for alanine polypeptide consisting of 21 amino acids obtained within the framework of MD approach is shown in the inset to Figure 11. From this plot it is seen that at $T \approx 300$ K, which is far beyond the temperature of the phase transition, the helicity of the Ala$_{21}$ polypeptide is 0.82. The fact that at low temperatures the helicity of the polypeptide obtained within the MD approach is smaller than unity arises due to the difficulty of defining the α-helix state of an amino acid. Thus, the helicity obtained within the MD approach rolls off at lower temperatures compared to the helicity of the polypeptide of the same length obtained using the statistical mechanics approach.

The kink in the helicity curve corresponds to the temperature of the phase transition of the system. As seen from Figure 11, with an increase of the polypeptide length, the helicity curve is becomes steeper as the phase transition is getting sharper. In the limiting case of an infinitely long polypeptide chain, the helicity should behave like a step function. This is yet another feature of a first-order phase transition.

3.4 Correlation of different amino acids in the polypeptide

An important question concerns the statistical independence of amino acids in the polypeptide at different temperatures. In the present section we analyze how a particular conformation of one amino acids influences the PES of other amino acids in the polypeptide. In Figure 12 we present the deviations of angles φ and ψ from the twisting angles φ_{10} and ψ_{10} in the 10th amino acid of alanine polypeptide. These results were obtained on the basis of MD simulations of the Ala$_{21}$ polypeptide at 300 K and at 1000 K. The deviation of angles φ and ψ is defined as follows:

$$RMSD(\varphi_i) = \sum_{j=1}^{j<=M} \sqrt{\frac{1}{M}(\varphi_i - \varphi_{10})^2}$$

$$RMSD(\psi_i) = \sum_{j=1}^{j<=M} \sqrt{\frac{1}{M}(\psi_i - \psi_{10})^2}, \qquad (14)$$

where i is the amino acid index in the polypeptide and M is the number of MD simulation steps. Note, that the plots shown in Figure 12 do not depend on the reference amino acid (we used the middle amino acid in the polypeptide).

The top plot in Figure 12 was obtained at 300 K. At this temperature, all amino acids in the polypeptide are in the α-helix conformation, and the deviation of angles φ and ψ is less than 16° for all amino acids except the boundary ones, where the relative deviation of the angles φ and ψ is 28° and 34° respectively. This happens because, while the boundary amino acids are loosely bounded, the central amino acids in the polypeptide are close to the minima that corresponds to an α-helix conformation. In the α-helix state, all central amino acids are stabilized by two hydrogen bonds, while the boundary amino acids form only one hydrogen bond.

At 1000 K the polypeptide is, to large extent, found in the random coil phase and therefore becomes more flexible. In the random coil phase, the stabilizing hydrogen bonds are broken, and the deviation of angles φ and ψ significantly increases. This fact is clearly seen in the bottom plot of Figure 12. However at 1000 K, the deviation of angles φ and ψ in the central and in the boundary amino acids is almost the same, confirming the assumption that in the random coil phase, short alanine polypeptides do not build hydrogen bonds.

Another important fact which is worth mentioning is that in the random coil phase (and in the central part of the α-helix), the deviation of angles φ and ψ does not depend on the distance between amino acids in the polypeptide chain. For instance, the deviation between angles in the 10th and in the 11th amino acid is almost the same as the deviation between angles in the 10th and in the 17th amino acid. This fact allows one to conclude that in a certain phase of the polypeptide (α-helix or random coil), amino acids can be treated as statistically independent.

4 Conclusion

In the present paper, we presented results of calculations obtained with the statistical method described in our preceding paper [1]. We have also performed a detail analysis of the α-helix↔random coil transition in alanine polypeptides of different lengths. We have calculated the potential energy surfaces of polypeptides with respect to their twisting degrees of freedom and constructed a parameter-free partition function of the polypeptide using our statistical formalism [1]. From this partition function, we derived and analyzed the temperature dependence of the heat capacity, latent heat and helicity of alanine polypeptides consisting of 21, 30, 40, 50 and 100 amino acids. Alternatively, we have obtained the same thermodynamical characteristics from the use of molecular dynamics simulations and compared them with the results of our statistical mechanics approach. The comparison proved the validity of our method and established its accuracy.

It was demonstrated that the heat capacity of alanine polypeptides has a peak at a certain temperature. The parameters of this peak (i.e. the maximal value of the heat capacity, the temperature of the peak, the width at half maximum, the area of the peak) were analyzed as a function of polypeptide length. Based on the predictions of the two energy-level model, it was demonstrated that the α-helix$\leftrightarrow$random coil transition in alanine polypeptide is a first order phase transition.

We have established a correspondence of our method with the results of the semiempirical approach suggested by Zimm and Bragg [4]. For this purpose we have determined the key parameters of the Zimm-Bragg semiempirical statistical theory. The calculated parameters of the Zimm-Bragg theory were compared with the results of earlier calculations from reference [19].

The final part of this paper deals with the statistical independence of amino acids in the polypeptide at different temperatures. It was shown that a particular conformation of one amino acids influences the PES of other amino acids in the polypeptide. We demonstrated that in a certain phase, amino acids can be treated as statistically independent.

In this paper, we demonstrated that the new statistical approach is applicable for the description of α-helix$\leftrightarrow$random coil phase transition in alanine polypeptides. However, this method is general and can be used to study similar processes in other complex molecular systems. For example, it would be interesting to apply the suggested formalism to the study of β-sheet$\leftrightarrow$random coil phase transition and to the study of non-homogeneous polypeptides (i.e. consisting of different amino acids). The suggested method can also be applied to the description of protein folding — an important question left open for further consideration. However, to study many of these problems it is important to extend the formalism further and to account for the long-range interactions [54], which might play an important role in complex molecular systems like proteins, DNA, RNA.

In this work, we have investigated α-helix$\leftrightarrow$random coil phase transition of alanine polypeptides in vacuo. So far there has been done no experimental work on α-helix$\leftrightarrow$random coil transition in gas phase. Nevertheless, it is important that such experiments are possible and can be performed using of the techniques MALDI [55–58] and the ESI mass spectroscopy [59,60]. We hope that our theoretical analysis of the α-helix$\leftrightarrow$random coil in alanine polypeptides in vacuo will stimulate experimentalists to verify our predictions.

This work is partially supported by the European Commission within the Network of Excellence project EXCELL, by INTAS under the grant 03-51-6170. We are grateful to Ms. Adilah Hussein for critical reading of the manuscript and several suggestions for improvement. The possibility to perform complex computer simulations at the Frankfurt Center for Scientific Computing is also gratefully acknowledged.

References

1. A. Yakubovich, I. Solov'yov, A.V. Solov'yov, W. Greiner, Eur. Phys. J. D, DOI: 10.1140/epjd/e2007-00328-9
2. A. Yakubovich, I. Solov'yov, A. Solov'yov, W. Greiner, Eur. Phys. J. D **40**, 363 (2006)
3. A. Yakubovich, I. Solov'yov, A. Solov'yov, W. Greiner, Europhys. News **38**, 10 (2007)
4. B. Zimm, J. Bragg, J. Chem. Phys. **31**, 526 (1959)
5. J. Gibbs, E. DiMarzio, J. Phys. Chem. **30**, 271 (1959)
6. S. Lifson, A. Roig, J. Chem. Phys. **34**, 1963 (1961)
7. J.A. Schellman, J. Phys. Chem. **62**, 1485 (1958)
8. S. Lifson, J. Chem. Phys. **40**, 3705 (1964)
9. D. Poland, H.A. Scheraga, J. Chem. Phys. **45**, 1456 (1966)
10. T. Ooi, M. Oobatake, Proc. Natl. Acad. Sci. USA **88**, 2859 (1991)
11. J. Gomez, V.J. Hilser, D. Xie, E. Freire, Proteins **22**, 404 (1995)
12. D.J. Tobias, C.L. Brooks, Biochem. **30**, 6059 (1991)
13. A.E. Garcia, K.Y. Sanbonmatsu, Proc. Natl. Acad. Sci. USA **99**, 2781 (2002)
14. H. Nymeyer, A.E. Garcia, Proc. Natl. Acad. Sci. USA **100**, 13934 (2003)
15. A. Irbäck, F. Sjunnesson, Proteins **56**, 110 (2004)
16. D. Shental-Bechor, S. Kirca, N. Ben-Tal, T. Haliloglu, Biophys. J. **88**, 2391 (2005)
17. R.A. Kromhout, B. Linder, J. Phys. Chem. B **105**, 4987 (2001)
18. A. Chakrabartty, T. Kortemme, R.L. Baldwin, Protein Sci. **3**, 843 (1994)
19. M. Go, N. Go, H.A. Scheraga, J. Chem. Phys. **52**, 2060 (1970)
20. H.A. Scheraga, J.A. Villa, D.R. Ripoll, Biophys. Chem. **01-102**, 255 (2002)
21. J.M. Scholtz, S. Marqusee, R.L. Baldwin, E.J. York, J.M. Stewart, M. Santoro, D.W. Bolen, Proc. Natl. Acad. Sci. USA **88**, 2854 (1991)
22. I.K. Lednev, A.S. Karnoup, M.C. Sparrow, S.A. Asher, J. Am. Chem. Soc. **123**, 2388 (2001)
23. P.A. Thompson, W.A. Eaton, J. Hofrichter, Biochem. **36**, 9200 (1997)
24. S. Williams, R.G. Thimothy, P. Causgrove, K.S. Fang, R.H. Callender, W.H. Woodruff, R.B. Dyer, Biochem. **35**, 691 (1996)
25. E. Shakhnovich, Chem. Rev. **106**, 1559 (2006)
26. A. Finkelstein, O. Ptitsyn, *Protein Physics. A Course of Lectures* (Elsevier Books, Oxford, 2002)
27. J.E. Shea, C.L. Brooks, Ann. Rev. Phys. Chem. **52**, 499 (2001)
28. N.V. Prabhu, K.A. Sharp, Ann. Rev. Phys. Chem. **56**, 521 (2005)
29. A. Rubin, *Biophysics: Theoretical Biophysics* (Moscow University Press, Nauka, 2004)

30. A. Yakubovich, I. Solov'yov, A. Solov'yov, W. Greiner, Eur. Phys. J. D **39**, 23 (2006)
31. A. Yakubovich, I. Solov'yov, A.V. Solov'yov, W. Greiner, Khimicheskaya Fizika (Chemical Physics) (in Russian) **25**, 11 (2006)
32. I. Solov'yov, A. Yakubovich, A. Solov'yov, W. Greiner, Phys. Rev. E **73**, 021916 (2006)
33. I. Solov'yov, A. Yakubovich, A.V.Solov'yov, W. Greiner, J. Exp. Theor. Phys. **102**, 314 (2006); original Russian text, published in Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki **129**, 356 (2006)
34. D. Rapaport, *The Art of Molecular Dynamics Simulation* (Cambridge University Press, 2004)
35. J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, J. Comp. Chem. **26**, 1781 (2005)
36. D. Frenkel, B.J. Smit, *Understanding Molecular Simulation* (Academic Press, 2001)
37. W. Coffey, Y. Kalmykov, J. Waldron, *The Langevin Equation*, Vol. 14 of *World Scientific in Contemporary Chemical Physics* (World Scientific Publishing Co., 2004)
38. F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw Hill, New York, 1965)
39. A. MacKerell et al., J. Phys. Chem. B **102**, 3586 (1998)
40. E. Henriques, A. Solov'yov, Abstract at the WE-Heraeus-Seminar *"Biomolecular Simulation: From Physical Principles to Biological Function"* (2006)
41. E. Henriques, A. Solov'yov, Eur. Phys. J. D, submitted, arXiv:0704.3193
42. M. Sotomayor, D.P. Corey, K. Schulten, Science **13**, 669 (2005)
43. J. Gullingsrud, K. Schulten, Biophys. J. **86**, 3496 (2004)
44. D.L. Nelson, M.M. Cox, *Principles of Biochemistry* (W.H. Freeman and Company, New York, 2005)
45. D. Voet, J. Voet, *Biochemistry* (John Willey and Sons, Inc., USA, 2004)
46. W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. **14**, 33 (1996)
47. A. Becke, Phys. Rev. A **38**, 3098 (1988)
48. C. Lee, W. Yang, R. Parr, Phys. Rev. B **37**, 785 (1988)
49. D. Wei, H. Guo, D. Salahub, Phys. Rev. E **64**, 011907 (2001)
50. L. Landau, E. Lifshitz, *Statistical Physics* (Pergamon Press, London-Paris, 1959)
51. A. Irbäck, B. Samuelsson, F. Sjunnesson, S. Wallin, Biophys. J. **85**, 1466 (2003)
52. C. Nowak, V.G. Rostiashvilli, T.A. Viglis, J. Chem. Phys. **46**, 4410 (1967)
53. T. Ooi, R.A. Scott, G. Vanderkooi, H.A. Scheraga, J. Chem. Phys. **46**, 4410 (1967)
54. M. Cubrovic, O. Obolensky, A. Solov'yov, in preparation (2007)
55. M. Karas, F. Hillenkamp, Anal. Chem. **60**, 2299 (1988)
56. F. Hillenkamp, M. Karas, Int. J. Mass Spectrom. **200**, 71 (2000)
57. M. Karas, U. Bahr, I. Fournier, M. Gluckmann, A. Pfenninger, J. Mass Spectrom. **226**, 239 (2003)
58. M. Wind, W. Lehmann, J. Anal. Atom. Spectrom. **19**, 20 (2004)
59. J. Fenn, M. Mann, C. Meng, S. Wong, C. Whitehouse, Science **246**, 64 (1989)
60. S. Bröndsted-Nielsen, J. Andersen, P. Hvelplund, B. Liu, S. Tomita, J. Phys. B-At. Mol. Opt. **37**, R25 (2004)