Ab initio study of alanine polypeptide chain twisting

Ilia A. Solov'yov,* Alexander V. Yakubovich, Andrey V. Solov'yov,† and Walter Greiner
Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Max von Laue Str. 1, 60438 Frankfurt am Main, Germany
(Received 4 November 2005; published 28 February 2006)

We have investigated the potential energy surfaces for alanine chains consisting of three and six amino acids. For these molecules we have calculated potential energy surfaces as a function of the Ramachandran angles φ and ψ, which are widely used for the characterization of the polypeptide chains. These particular degrees of freedom are essential for the characterization of the proteins folding process. Calculations have been carried out within the *ab initio* theoretical framework based on the density functional theory and accounting for all the electrons in the system. We have determined stable conformations and calculated the energy barriers for transitions between them. Using a thermodynamic approach, we have estimated the times of characteristic transitions between these conformations. The results of our calculations have been compared with those obtained by other theoretical methods and with the available experimental data extracted from the Protein Data Base. This comparison demonstrates a reasonable correspondence of the most prominent minima on the calculated potential energy surfaces to the experimentally measured angles φ and ψ for alanine chains appearing in native proteins. We have also investigated the influence of the secondary structure of polypeptide chains on the formation of the potential energy landscape. This analysis has been performed for the sheet and the helix conformations of chains of six amino acids.

DOI: 10.1103/PhysRevE.73.021916 PACS number(s): 87.15.−v, 31.15.Ar, 31.15.Ew, 31.50.Bc

I. INTRODUCTION

It is well known that proteins consist of amino acids whose number may vary in the range from hundreds up to tens of thousands. Small fragments of proteins are usually called polypeptide chains or polypeptides. This work is devoted to a study of the conformational properties of alanine polypeptide chains.

Since recently, it became possible to study experimentally small fragments of proteins and polypeptides in the gas phase with the use of the matrix assisted laser desorption ionization (MALDI) mass spectroscopy [1–4] and the electrospray ionization (ESI) mass spectroscopy [5,6]. From theoretical viewpoints, the investigation of small polypeptides is of significant interest because they can be treated by means of *ab initio* methods which allow accurate comparison of theoretical predictions with experiment. The results of *ab initio* calculations can be then utilized for the development of model approaches applicable for the description of larger and more complex protein structures.

Polypeptides are characterized by the primary and the secondary structure [7–10]. Different geometrical configurations of a polypeptide are often called as the conformations. The number of various conformations (isomeric states) grows rapidly with the growth of a system size. Thus, a search for the most stable conformations becomes an increasingly difficult problem for large molecules.

The sheet and the helix structures are the most abundant motifs in proteins. The study of the transition between these motifs and the evaluation of its characteristic duration is of significant interest, because it is closely related to one of the most intriguing problems of the protein physics—the protein folding. To study this transition it is necessary to investigate the potential energy surface of amino acid chains with respect to their twisting. In the present paper we have studied the potential energy surfaces for small alanine chains. These molecules were chosen because they are often present in native proteins as fragments, and also because they allow *ab initio* theoretical treatment due to their relatively small size.

Previously, only glycine and alanine dipeptides were studied in detail. Sometimes their analog was used to reduce the computational costs [for example, (S)-α-formylamino)propanamide]. In Refs. [11–13] alanine and glycine dipeptides were investigated within the Hartree-Fock theory. In these papers the potential energy surfaces were calculated versus the twisting angles of the molecules. Different stable states of the dipeptides, corresponding to different molecular conformations, were determined. In Refs. [14–19] different conformations and their energies were determined within the framework of the density functional theory. In Ref. [19] dynamics of the alanine dipeptide analog was discussed and the time of the transitions between the two conformations of the alanine dipeptide was found.

A number of papers were devoted to the study of tripeptides. In Refs. [20–24] dynamics of the alanine and glycine tripeptides was studied by means of classical molecular dynamics and with the use of semiempirical potentials (such as GROMOS [25], CHARMM [26], and AMBER [27]). In [28] within the framework of the Hartree-Fock theory several stable conformations of alanine and glycine tripeptides were found. In Ref. [29] the Raman and IR spectra for alanine and glycine tripeptides were measured in neutral, acidy, and alkali environments.

Polypeptides of greater length have been studied less. We are aware of only several related papers. In particular, stable

*On leave from the A.F. Ioffe Institute, St. Petersburg, Russia. E-mail: ilia@th.physik.uni-frankfurt.de
†On leave from the A.F. Ioffe Institute, St. Petersburg, Russia. E-mail: solovyov@fias.uni-frankfurt.de

conformations of neutral and charged alanine hexapeptides were obtained with the use of empirical potentials and discussed in Ref. [30]. Experimental NMR study of various conformations of alanine heptapeptides at different temperatures was carried out in Ref. [31]. In Ref. [32] with the use of empirical molecular dynamics based on Monte-Carlo methods, a polypeptide consisting of 21 amino acids was described.

In the present paper we have performed an *ab initio* calculation of the multidimensional potential energy surface for the alanine polypeptide chains consisting of three and six amino acids. The potential energy surface versus twisting degrees of freedom of the polypeptide chain has been calculated. The calculations have been performed within an *ab initio* theoretical framework based on the density functional theory (DFT) accounting for all the electrons in the system. Previously, these kinds of calculations were performed only for dipeptides (see, e.g., [11,12,19]). For larger molecules, only a few conformations were considered (see citations above). We have calculated the energy barriers for the transitions between different conformations and determined the energetically most favorable ones. Using a thermodynamic approach, we have estimated times of the characteristic transitions between the most energetically favorable conformations. The results of our calculation have been compared with other theoretical simulations and with the available experimental data. We have also analyzed how the secondary structure of polypeptide chains influences the potential energy landscapes. In particular, the role of the secondary structure in the formation of stable conformations of the chains of six amino acids being in the sheet and in the helix conformations has been elucidated. Some preliminary results of our work were published as electronic preprints [33,34].

Our paper is organized as follows. In Sec. II we give a brief overview of theoretical methods used in our work. In Sec. III we present and discuss the results of our computations. In Sec. IV we draw a conclusion to this paper. The atomic system of units, $|e|=m_e=\hbar=1$, is used throughout the paper unless other units are indicated.

II. THEORETICAL METHODS

In the present paper we study the multidimensional potential energy surfaces for alanine polypeptides within the framework of the density functional theory. The potential energy surfaces are multidimensional functions of atomic coordinates. In our work the potential energy surfaces are considered as a function of the dihedral angles formed by the atoms of the polypeptide chain. For this calculation the Born-Oppenheimer approximation allowing to separate the motion of the electronic and ionic subsystems is used.

DFT is a common tool for the calculating of properties of quantum many body systems in which many electron correlations play an important role. The DFT formalism is well known and can be found in many textbooks (see, e.g., [35,36]). Therefore in our work we present only the basic equations and ideas of this method.

Electronic wave functions and energy levels within the framework of DFT are obtained from the Kohn-Sham equations, which read as (see, e.g., [35,36])

$$\left(\frac{\hat{p}^2}{2}+U_{ions}+V_H+V_{xc}\right)\psi_i=\varepsilon_i\psi_i,\quad i=1,\ldots,N, \tag{1}$$

where the first term represents the kinetic energy of the ith electron with the wave function ψ_i and the energy ε_i, N is the number of electrons in the system. U_{ions} describes the electron attraction to the ionic centers, V_H is the Hartree part of the interelectronic interaction Ref. [37],

$$V_H(\mathbf{r})=\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}', \tag{2}$$

and $\rho(\mathbf{r}')$ is the electron density,

$$\rho(\mathbf{r})=\sum_{\nu=1}^{N}|\psi_i(\mathbf{r})|^2. \tag{3}$$

V_{xc} is the local exchange-correlation potential, which is defined as a functional derivative of the exchange-correlation energy functional:

$$V_{xc}=\frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}. \tag{4}$$

Equation (4) is exact and follows from the Hohenberg theory Ref. [38]. However, no unique potential E_{xc}, universally applicable for different systems and conditions, has been found so far.

Approximate functionals employed by the DFT usually partition the exchange-correlation energy into two parts, referred to as the *exchange* and the *correlation* terms:

$$E_{xc}[\rho]=E_x(\rho)+E_c(\rho). \tag{5}$$

Both terms are the functionals of the electron density, which can be of two distinctly different types: either a *local* functional depending on only the electron density ρ, or a *gradient-corrected* functional depending on both ρ and its gradient, $\nabla\rho$. A variety of exchange-correlation functionals can be found in the literature. Below, we refer only to those which are related to the calculation performed in this work.

The local exchange functional is virtually always defined as follows:

$$E_x^{LDA}=-\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\int\rho^{4/3}d^3\mathbf{r}. \tag{6}$$

This form was developed to reproduce the exchange energy of a uniform electron gas. By itself, however, it is not sufficient for the adequate description of a multiatomic system.

The gradient-corrected exchange functional introduced by Becke [39] and based on the LDA exchange functional reads as

$$E_x^{B88}=E_x^{LDA}-\gamma\int\frac{\rho^{4/3}x^2}{1+6\gamma\sinh^{-1}x}d^3\mathbf{r}, \tag{7}$$

where $x=\rho^{-4/3}|\nabla\rho|$ and $\gamma=0.0042$ is a parameter chosen to fit the known exchange energies of the noble gas atoms.

Analogously to the above written exchange functionals, there are local and gradient-corrected correlation functionals, for example, those introduced by Lee, Yang, and Parr [40].

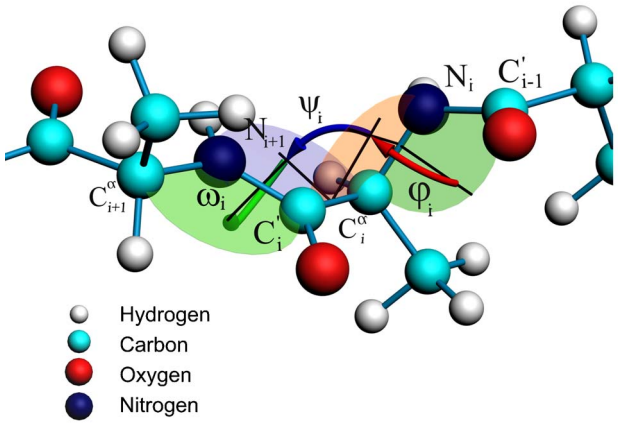

FIG. 1. (Color online) Dihedral angles φ and ψ used to characterize the potential energy surface of the polypeptide chain.

Its explicit expression is somewhat lengthy and thus we do not present it here and refer to the original papers.

In the pure DFT, an exchange functional usually pairs with a correlation functional. For example, the well-known BLYP functional pairs Becke's gradient-corrected exchange functional (7) with the gradient-corrected correlation functional of Lee, Yang, and Parr [40].

In spite of the success of the pure DFT theory in many cases, one has to admit that the Hartree-Fock theory accounts for the electron exchange more naturally and precisely. Thus, Becke has suggested [39] functionals which include a mixture of Hartree-Fock and DFT exchange along with DFT correlations, conceptually defining E_{xc} as

$$E_{xc}^{mix} = c_{HF}E_x^{HF} + c_{DFT}E_{xc}^{DFT}, \tag{8}$$

where c_{HF} and c_{DFT} are constants. Following this idea, a Becke-type three parameter functional can be defined as follows:

$$E_{xc}^{B3LYP} = E_x^{LDA} + c_0(E_x^{HF} - E_X^{LDA}) + c_x(E_x^{B88} - E_x^{LDA}) + E_c^{VWN3} + c_c(E_c^{LYP} - E_c^{VWN3}). \tag{9}$$

Here, c_0=0.2, c_x=0.72, and c_c=0.81 are constants, which were defined by fitting to the atomization energies, ionization potentials, proton affinities, and first-row atomic energies [41]. E_x^{LDA} and E_x^{B88} are defined in (6) and (7), respectively. E_x^{HF} is the functional corresponding to Hartree-Fock equations. The explicit form for the correlation functional E_c^{VWN3} as well as for gradient-corrected correlation functional of Lee, Yang, and Parr, E_c^{LYP}, can be found in [42,40], respectively.

III. RESULTS AND DISCUSSION

A. Determination of the polypeptide twisting degrees of freedom

In this section we present the potential energy surfaces for the glycine polypeptide chains calculated versus dihedral angles φ and ψ defined in Fig. 1. In particular, we focus on the chains consisting of three and six amino acids.

Both angles are defined by the four neighboring atoms in the polypeptide chain. The angle φ_i is defined as the dihedral angle between the planes formed by the atoms $(C'_{i-1}-N_i-C_i^\alpha)$ and $(N_i-C_i^\alpha-C'_i)$. The angle ψ_i is defined as the dihedral angle between the $(N_i-C_i^\alpha-C'_i)$ and $(C_i^\alpha-C'_i-N_{i+1})$ planes. Beside the angles φ_i and ψ_i there is an angle ω_i, which is defined as the dihedral angle between $(C_i^\alpha-C'_i-N_{i+1})$ and $(C'_i-N_{i+1}-C_{i+1}^\alpha)$ planes. The atoms are numbered from the NH_2 terminal of the polypeptide. The angles φ_i, ψ_i, and ω_i take all possible values within the interval $[-180°;180°]$. For the unambiguous definition we count the angles φ_i, ψ_i, and ω_i clockwise, if one looks at the molecule from its NH_2 terminal (see Fig. 1). This way of angle counting is the most commonly used [10].

The angles φ_i and ψ_i can be defined for any amino acid in the chain, except the first and the last ones. Below we omit the subscripts and consider angles φ and ψ for the middle amino acid of the polypeptide.

B. Optimized geometries of alanine polypeptides

In order to study the twisting of a polypeptide chain one needs first to define its initial structure. Although the number of its conformations increases with the growth of the molecule size, there are certain types of polypeptide structure, namely the sheet and the helix conformations, which are the most typical. In the present paper we have investigated the twisting of the polypeptide chains of the sheet and the helix conformations. By varying the angles φ and ψ in the central amino acid one can create the structure of the polypeptide differing significantly from the pure sheet or helix conformations. If the structure of a polypeptide can be transformed to a helix or a sheet one by a trivial variation of φ and ψ, such polypeptides for the sake of simplicity are referred below as belonging to the group of the helix or the sheet structure, respectively.

In Fig. 2 we present the optimized geometries of alanine polypeptide chains that have been used for the exploration of the potential energy surfaces. All geometries were optimized with the use of the B3LYP functional. Figure 2(a) shows the alanine tripeptide structure. In the present work we choose the sheet conformation, because the tripeptide is too short to

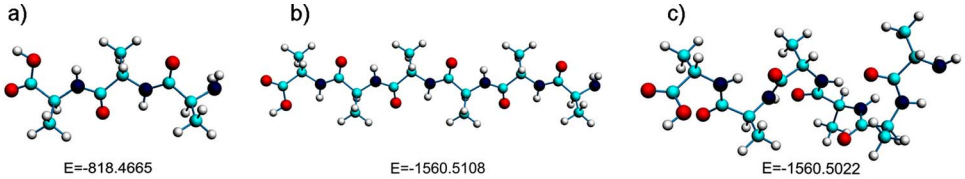

FIG. 2. (Color online) Optimized geometries of alanine polypeptide chains calculated by the B3LYP/6−31++G(d,p) method: (a) alanine tripeptide; (b) alanine hexapeptide (sheet conformation); (c) alanine hexapeptide (helix conformation).

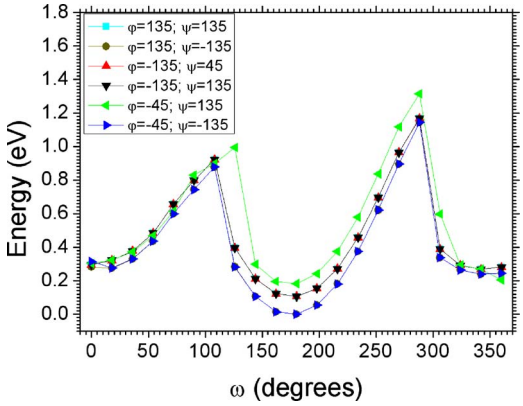

FIG. 3. (Color online) Dependence of alanine tripeptides energy on angle ω calculated by the B3LYP/6–31G(d) method at different values of angles φ and ψ.

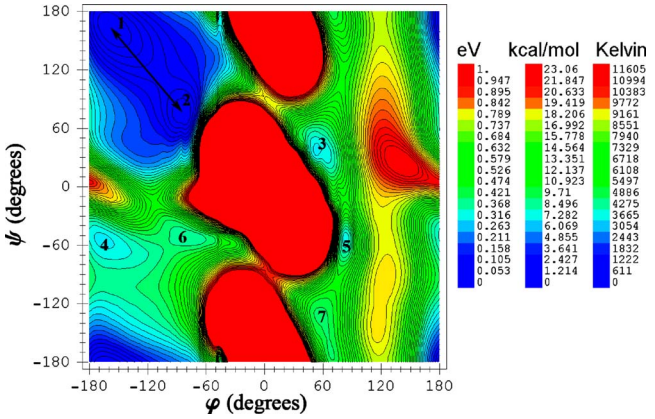

FIG. 4. (Color online) Potential energy surface for the alanine tripeptide calculated by the B3LYP/6–31G(2d,p) method. Energies are given in eV, kcal/mol, and kelvin. Numbers mark the energy minima on the potential energy surface. Arrows show transition paths between different conformations of the molecule.

form the helix conformation. Figures 2(b) and 2(c) show alanine hexapeptide in the sheet and the helix conformations, respectively. The total energies (in atomic units) of the molecules are given below the images.

C. Polypeptide energy dependance on the dihedral angle ω

For each amino acid there are only three dihedral angles formed by atoms of the polypeptide chain which describe its twisting. The angle ω (see Fig. 1) differs from the angles φ and ψ, because the C_i' atom has the sp^2 hybridization state, what leads to formation of a quasidouble bond between C_i' and N_{i+1} atoms. Therefore, the angle ω is often referred as a "stiff" degree of freedom, whose value depends only slightly on both the polypeptide constituent amino acids and the values of other degrees of freedom. To illustrate this fact, in Fig. 3 we present the energy dependencies on ω calculated for alanine dipeptide with different values of angles φ and ψ in the central amino acid.

From this figure it is clear that there are two stable states in the system with $\omega=0°$ and $\omega=180°$ which do not depend on the angles φ and ψ. The heights of the barriers between these states are weakly dependent on φ and ψ, being equal to $\sim$1 eV=23.06 kcal/mol.

The calculation shows that at temperatures close to the room temperature, the value of the angle ω changes insignificantly. The potential energy surface as a function of the angles φ and ψ appears to be much more complex as it is shown in the next section.

D. Potential energy surface for alanine tripeptide

In Fig. 4 we present the potential energy surface for the alanine tripeptide calculated by the B3LYP/6–31G(2d,p) method. The energy scale is given in eV, kcal/mol, and kelvin. Energies on the plot are measured from the lowest energy minimum of the potential energy surface.

From the figure follows that there are several minima on the potential energy surface. They are numbered according to the value of the corresponding energy value. Each minimum corresponds to a certain conformation of the molecule. These conformations differ significantly from each other. In the case of alanine tripeptide there are six conformations, shown in Fig. 5. Dashed lines show the strongest hydrogen bonds in the system, which arise when the distance between hydrogen and oxygen atoms becomes less then 2.9 Å.

To calculate the potential energy surface the following procedure was adopted. Once the stable structure of the molecule has been determined and optimized, all but two (these are the angles φ and ψ in the central amino acid) degrees of freedom were frozen. Then the energy of the molecule was calculated by varying φ and ψ. This procedure was used to calculate all potential energy surfaces presented below in this section. It allows one to find efficiently the minima on the energy surface and to determine the main stable conformations of the molecule. The absolute energy values of different conformations of the tripeptide found by this method are not too accurate, because the method does not account for the relaxation of other degrees of freedom in the system. To calculate the potential energy surface with accounting for the relaxation one needs 20–30 times more of the computer time. Therefore, calculations with accounting for the relaxation have not been performed in our work. Instead, we have performed a complete optimization of the molecular conformations, corresponding to all minima on the calculated potential energy surface.

In Fig. 5 we compare stable conformations of the alanine tripeptide calculated with and without accounting for the relaxation of all atoms in the system. As it is seen from this figure the angles φ and ψ differ by about 10 percent in the two cases. This difference arises due to the coupling of φ and ψ with other degrees of freedom. Note the change of the sign of the relative energies of some conformations. This effect is due to the rearrangement of side atoms (radicals) in the polypeptide chain which lowers the energies of different conformations differently.

In our work the potential energy surface has been calculated and interpolated on the grid with the step of 18°. This step size is an optimal one, because the interpolation error is about 9°, i.e., comparable with the angle deviations caused by the relaxation of all degrees of freedom in the system.

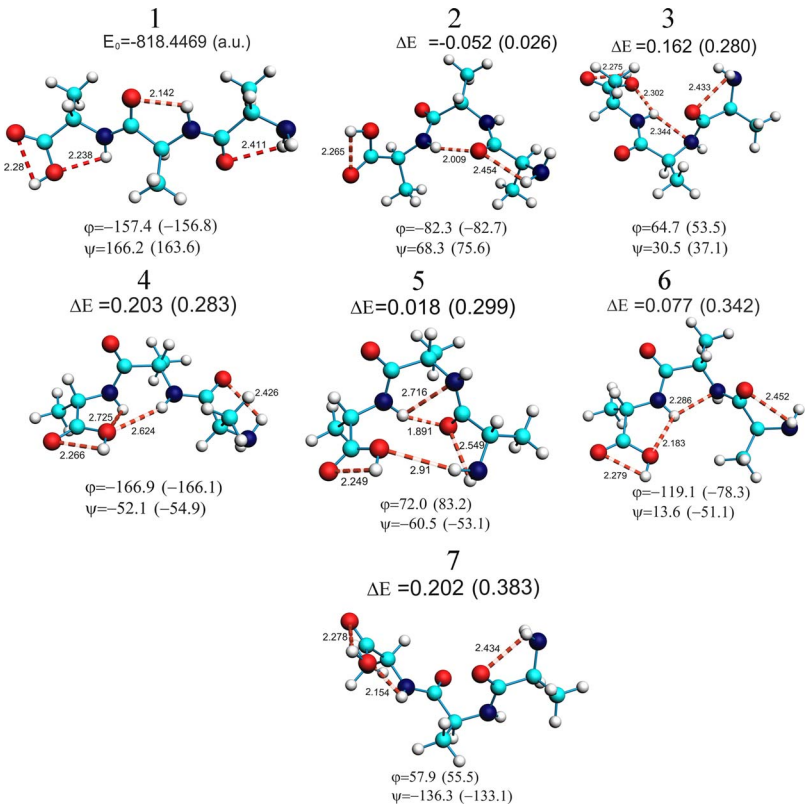

FIG. 5. (Color online) Optimized conformations of the alanine tripeptide. Different geometries correspond to different minima on the potential energy surface (see contour plot in Fig. 4). Below each image we present angles φ and ψ, which have been obtained with accounting for relaxation of all degrees of freedom in the system. Values in brackets give the angles calculated without accounting for relaxation. Above each image the energy of the corresponding conformation is given in eV. The energies are counted from the energy of conformation 1 (the energy of conformation 1 is given in a.u.). Values in parentheses correspond to the energies obtained without relaxation of all degrees of freedom in the system. Dashed lines show the strongest hydrogen bonds. Their lengths are given in angstroms.

Note that for the alanine tripeptide an additional maximum appears at $\varphi=120°\pm50°$, $\psi=30°\pm30°$, while it is absent on the potential energy surface for the glycine tripeptide [33]. This maximum is a result of the overlapping of the side CH_3-radicals, which are substituted in the case of the glycine polypeptide with the H atoms.

In Refs. [11,12] several stable conformations were found for alanine and glycine dipeptides. The values of angles φ and ψ for the stable conformations of dipeptide and tripeptide are close indicating that the third amino acid in tripeptide makes a relatively small influence on the values of dihedral angles of two other amino acids. In earlier papers Refs. [11,12] dipeptides were studied within the framework of the Hartree-Fock theory. In Ref. [11], values of φ and ψ were obtained by the HF/6-31+G* method, and in Ref. [12] by HF/6-31G**. In Table I we compare the results of our calculation for tripeptides with the corresponding data obtained for dipeptides. Some discrepancy between the values presented is due to the difference between the dipeptide and tripeptide (i.e., the third alanine in the tripeptide affects the values of angles φ and ψ). However, another source of discrepancy might arise due to the accounting for the many-electron correlations in the DFT and neglecting this effect in the Hartree-Fock theory used in Refs. [11,12].

Figure 4 shows that some domains of the potential energy surface, where the potential energy of the molecule increases significantly, appear to be unfavorable for the formation of a stable molecular configuration. The growth of energy takes place when some atoms in the polypeptide chain approach each other at small distances. Accounting for the molecule relaxation results in the decrease of the system energy in such cases, but the resulting molecular configurations remain unstable. We call such domains on the potential energy surface as forbidden ones. In Fig. 4 one can identify two forbidden regions in the vicinity of the points (0, 0) and (0, 180). At (0, 0) a pair of hydrogen and oxygen atoms ap-

TABLE I. Comparison of dihedral angles φ and ψ corresponding to different conformations of alanine tripeptide.

Conformation	φ [11]	ψ [11]	φ [12]	ψ [12]	φ	ψ
1	−168.4	170.5	−157.2	159.8	−157.4	166.2
2			−60.7	−40.7	−82.3	−68.3
3	63.8	32.7	67.0	30.2	64.7	30.5
4					−166.9	−52.1
5	74.1	−57.3	76.0	−55.4	72.0	−60.5
6	−128.0	29.7	−130.9	22.3	−119.1	13.6
7					57.9	−136.3

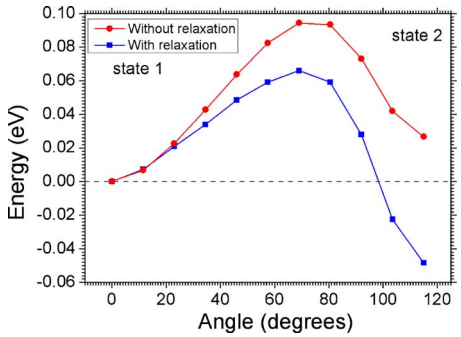

FIG. 6. (Color online) Transition barriers for between conformations $1 \leftrightarrow 2$ of the alanine tripeptide. Circles and squares correspond to the barriers calculated without and with relaxation of all degrees of freedom in the system.

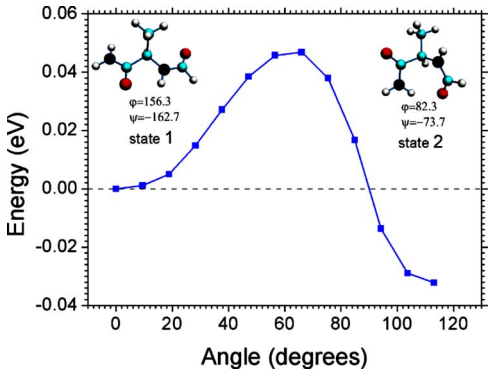

FIG. 7. (Color online) Transition barriers for between conformations $1 \leftrightarrow 2$ of the alanine dipeptide analog calculated by the B3LYP/6−31+G(2d,p) method accounting for the relaxation of all degrees of freedom in the system. Structure of the conformations 1 and 2 is shown near each minimum.

proach the distances much smaller than the characteristic H-O bondlength. This leads to a strong interatomic repulsion caused by the exchange interaction of electrons. At (0, 180) the Coulomb repulsion of a pair of oxygen atoms causes the similar effect.

Figure 4 shows that there are six minima on the potential energy surface for alanine tripeptide. The transition barrier between the conformations $1 \leftrightarrow 2$ is shown in Fig. 6. The barrier has been calculated with and without relaxation of the atoms in the system. The corresponding transition path is marked in Fig. 4 by an arrow. This comparison demonstrates that accounting for the relaxation significantly lowers the barrier height and influences the relative value of energy of the minima.

Let us now estimate the time needed for a system for the transition from one conformation to another. To do this we use the Arhenius equation, which reads as

$$\frac{1}{\tau} = \Omega e^{-\Delta E/kT} \tag{10}$$

where τ is the transition time, Ω is the factor, determining how frequently the system approaches the barrier, ΔE is the barrier height, T is the temperature of the system, k is the Boltzmann factor.

Figure 7 shows the transition barrier between two main conformations of the alanine dipeptide analog [(S)-α-(formylamino)propanamide]. It is seen that $\Delta E_{1\to 2} = 0.047$ eV for the transition $1 \to 2$, while $\Delta E_{2\to 1} = 0.079$ eV for the transition $2 \to 1$. The frequency Ω for this molecule is equal to 42.87 cm^{-1}. Thus, at $T = 300$ K, we obtain $\tau^{1\to 2}_{2\times Ala} \sim 5$ ps and $\tau^{2\to 1}_{2\times Ala} \sim 17$ ps. This result is in excellent agreement with the molecular dynamics simulations results obtained in Ref. [19] predicting $\tau \sim 7$ ps for the transition $1 \to 2$ and $\tau \sim 19$ ps for the transition $2 \to 1$. This comparison demonstrates that our method is reliable enough and it can be used for the estimation of transition times between various conformations of the polypeptides.

Using the B3LYP/6-31G(2d,p) method we have calculated the frequencies of normal vibration modes for the alanine tripeptide. The characteristic frequency corresponding to the twisting of the polypeptide chain is equal to 32.04 cm^{-1}. From Fig. 6 follows that $\Delta E_{1\to 2} = 0.066$ eV for the transition $1 \to 2$ and $\Delta E_{2\to 1} = 0.114$ eV for the transition $2 \to 1$. Thus, we obtain $\tau^{1\to 2}_{3\times Ala} \sim 13$ ps and $\tau^{2\to 1}_{3\times Ala} \sim 86$ ps. Let us note that these transition times can be measured experimentally by means of NMR Refs. [10,43].

E. Potential energy surface for alanine hexapeptide with the sheet and the helix secondary structure

In Fig. 8 we present contour plots of the potential energy surface for the alanine hexapeptide with the sheet [part (a)] and the helix [part (b)] secondary structure, respectively, versus dihedral angles φ and ψ. In both cases the forbidden regions arise because of the repulsion of oxygen and hydrogen atoms analogously to the alanine tripeptide case.

Minima 1–6 on the potential energy surface Fig. 8(a) correspond to different conformations of the alanine hexapeptide with the sheet secondary structure. Note that these minima are also present on the potential energy surface of the alanine tripeptide (see Fig. 4). Geometries of the conformations 1–6 are shown on the right-hand side of Fig. 8(a).

The energy barrier as a function of a scan variable [see Fig. 8(a)] for the transition between conformations 1 and 2 is shown in Fig. 9. The energy dependence has been calculated with and without relaxation of all the atoms in the system. In the case of alanine hexapeptide with the sheet secondary structure the barrier height for the transition $1 \to 2$ is significantly higher than for the transition $2 \to 1$, being equal to 0.095 and 0.023 eV, respectively. The normal vibration mode frequency, corresponding to the twisting of the polypeptide chain is equal to 6.24 cm^{-1} and was calculated with the B3LYP/STO-3G method, where STO-3G is a basis set, whose acronym stands for Stater-type-orbitals simulated by three Gaussians added together. Using Eq. (10) one derives the transition times at room temperature: $\tau^{1\to 2}_{6\times Gly} \sim 211$ ps, $\tau^{2\to 1}_{6\times Gly} \sim 13$ ps.

Let us now consider alanine hexapeptide with the helix secondary structure. The potential energy surface for this polypeptide is shown in Fig. 8(b). The positions of minima on this surface are shifted significantly compared to the cases discussed above. This change takes place because of the in-

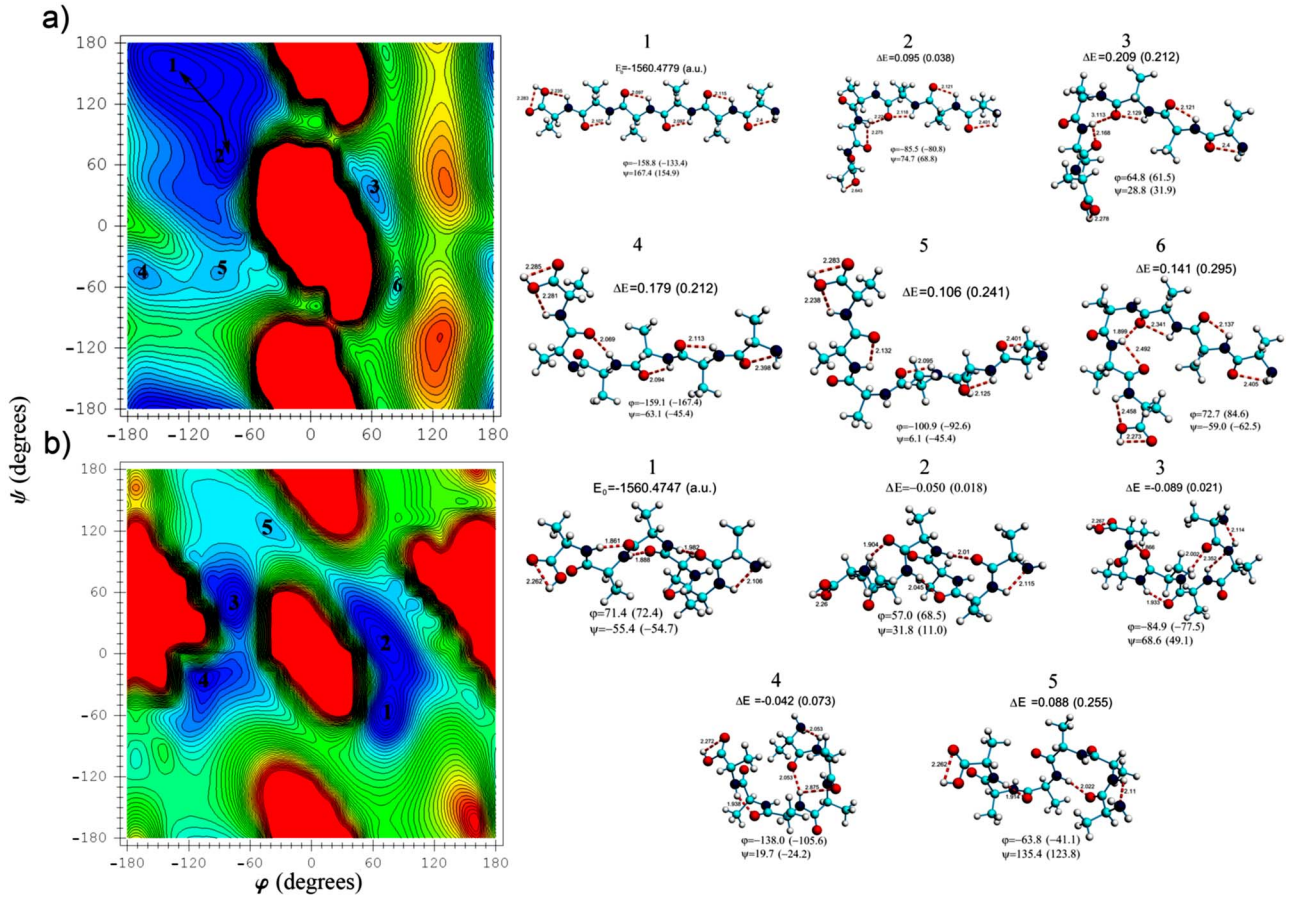

FIG. 8. (Color online) Potential energy surface for the alanine hexapeptide with the sheet secondary structure [part (a)] and with the helix secondary structure [part (b)] calculated by the B3LYP/6−31G(2d,p) method. The energy scale is given in Fig. 4. Numbers mark the energy minima on the potential energy surface. Images of optimized conformations of the alanine hexapeptide are shown near the corresponding energy landscape. Values of angles φ and ψ, as well as the relative energies of the conformations, are given analogously to that in Fig. 5.

fluence of the secondary structure of the polypeptide on the potential energy surface. The geometries of the most stable conformations are shown on the right-hand side of Fig. 8(b).

For the alanine hexapeptide with the helix secondary structure there is a maximum at $\varphi \sim 180°$ and $\psi \sim 40°$ in addition to the central maxima on the potential energy surface. This maximum appears because of the repulsive interaction of the outermost amino acids side radicals.

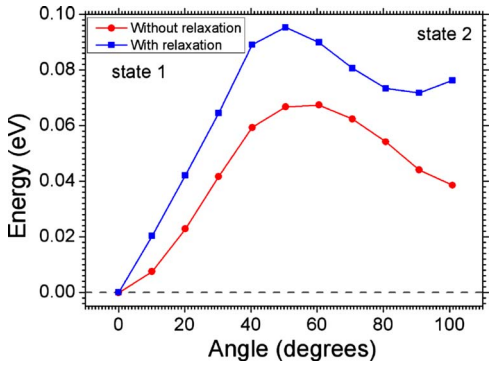

FIG. 9. (Color online) Transitions barriers between conformations $1 \leftrightarrow 2$ of the alanine hexapeptide with the sheet secondary structure. Circles and squares correspond to the barriers calculated without and with relaxation of all degrees of freedom in the system.

It is worth noting that for some conformations of alanine hexapeptide the angles φ and ψ change significantly when the relaxation of all degrees of freedom in the system is accounted [see, for example, conformations 1, 5 in Fig. 8(a) and conformations 2, 4 in Fig. 8(b)]. This means that the potential energy surface of the alanine hexapeptide in the vicinity of the mentioned minima is very sensitive to the relaxation of all degrees of freedom. However, the calculation of the potential energy surface with accounting for the relaxation of all degrees of freedom is unfeasible task. Indeed, one needs about 2000 h of computer time (Pentium Xeon 2.4 GHz) for the calculation of the potential energy surface for the alanine hexapeptide. To perform an analog calculation with accounting for the relaxation, about 5 y of computer time would be needed. Nevertheless, the potential energy surface calculated without accounting for the relaxation carries a lot of useful information. Thus, one can predetermine stable conformations of polypeptide, which then can be used as starting configurations for further energy minimization.

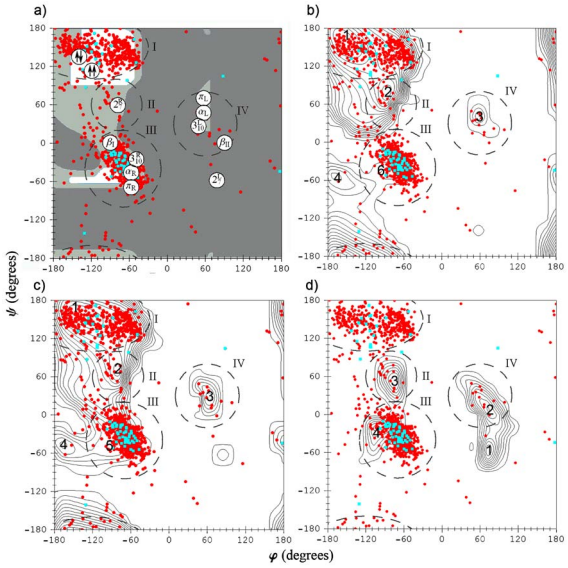

FIG. 10. (Color online) Comparison of angles φ and ψ of alanine residues in protein structures selected from the Brookhaven Protein Data Bank [9,44] with the steric diagram for polyalanine [45] [part (a)]. Comparison of angles φ and ψ of alanine residues in protein structures selected from the Brookhaven Protein Data Bank [9,44] with the minima on the calculated potential energy surfaces for alanine tripeptide (b); alanine hexapeptide in sheet conformation (c); alanine hexapeptide in helix conformation (d). Transparent rhomboids correspond to alanines surrounded with alanines, while filled circles correspond to alanines surrounded by other amino acids. Dashed ellipses mark the regions of higher concentration of the observed angles.

F. Comparison of calculation results with experimental data

Nowadays, the structure of many proteins has been determined experimentally Ref. [9]. Knowing the protein structure one can find the angles φ and ψ for each amino acid in the (Fig. 10) protein.

In Fig. 10(a) we show a map of the allowed and forbidden conformations for alanine residues in polyalanine chains taken from Ref. [45] (steric Ramachandran diagram). This map was obtained from pure geometrical considerations, in which the structure of the polypeptide was assumed to be fixed and defined by the interatomic van der Waals interaction radii. Depending on the distances between the atoms one could distinguish three regions: completely allowed, conventionally allowed, and forbidden. The conformation is called completely allowed if all the distances between atoms of different amino acids are larger than some critical value $r_{ij} \geqslant r_{max}$. Conventionally allowed regions on the potential energy surface correspond to the conformations of the polypeptide, in which the distances between some atoms of different amino acids lie within the interval $r_{min} \leqslant r_{ij} < r_{max}$. All other conformations are referred to as forbidden. The values of r_{min} and r_{max} are defined by the types of interacting atoms and can be found in the textbooks (see, e.g., [45]). In Fig. 10(a) we mark the completely allowed regions with white, the conventionally allowed regions with light gray, and the forbidden regions with a dark gray color. In this figure we mark the points, which correspond to the geometries of alanine, whose

TABLE II. Angles φ and ψ corresponding to the most prominent polyalanine secondary structures.

Structure type	φ (deg.)	ψ (deg.)
Right-handed (left-handed) 2_7 helix	−78 (78)	59 (−59)
Right-handed (left-handed) 3_{10} helix	−49 (49)	−26 (26)
Right-handed (left-handed) α-helix (4_{13})	−57 (57)	−47 (47)
Right-handed (left-handed) π-helix (5_{16})	−57 (57)	−70 (70)
Parallel β sheet ($\uparrow\uparrow$)	−119	113
Antiparallel β sheet ($\uparrow\downarrow$)	−139	135
β-turn of type I	−90	0
β-turn of type II	90	0

periodical iteration leads to the formation of chains with specific secondary structure. In Table II we compile the values of angles φ and ψ, which correspond to the most prominent polyalanine secondary structures. For the illustrative purposes we mark these points by white circles with the corresponding type of the secondary structures typed in. Thus, 2_7^R, 2_7^L are the right-handed and the left-handed 2_7 helix; 3_{10}^R, 3_{10}^L are the right-handed and the left-handed 3_{10} helix; α_R, α_L are the right-handed and the left-handed α-helix (4_{13}); π_R, π_L are the right-handed and the left-handed π-helix (5_{16}); $\uparrow\uparrow$, $\uparrow\downarrow$ are the parallel and antiparallel β sheets. β_I, β_{II} correspond to the β-turns of types I and II, respectively.

Note that not all of the structures listed above are present equally in proteins. In Fig. 10(a) we show the distribution of the angles φ and ψ of alanine residues in protein structures selected from the Brookhaven Protein Data Bank [9,44]. It is possible to distinguish four main regions, in which most of experimental points are located. In Fig. 10 these regions are schematically shown with dashed ellipses. Note, that these ellipses are used for illustrative purposes only, and serve for a better understanding of the experimental data. The regions in which most of the observed angles φ and ψ are located correspond to different secondary structures of the polyalanine. Thus, region I corresponds to the parallel and antiparallel β-sheets. Region II corresponds to the right-handed 2_7^R helix. Region III corresponds to the right-handed α_R-helix, right-handed π_R-helix, right-handed 3_{10}^R helix, and β-turn of type I. Region IV corresponds to the left-handed α_L-helix, right-handed π_L-helix, left-handed 3_{10}^L helix, and β-turn of type II. In some cases there are several types of secondary structure within one domain. In the present work we have not studied the secondary structure of proteins systematically enough to establish the univocal correspondence of the observed experimental points to different types of the secondary structure.

Let us now compare the distribution of angles φ and ψ experimentally observed for proteins with the potential energy landscape calculated for alanine polypeptides and establish correspondence of the secondary structure of the calculated conformations with the predictions of the simple Ramachandran model.

Region I corresponds to the minimum 1 on the both potential energy surfaces of the alanine tripeptide [Fig. 10(b)] and the alanine hexpeptide with the secondary structure of

the sheet conformation [Fig. 10(c)]. These conformations correspond exactly to the alanine chains in the β-sheet conformation [see Figs. 5 and 8(a)]. Note that there is no minimum in that region of the potential energy surface for alanine hexapeptide with the secondary structure of helix [see Fig. 10(d)].

Region II corresponds to the minimum 2 on the both potential energy surfaces Figs. 10(b) and 10(c), as well as to the minimum 3 on the potential energy surface Fig. 10(d). On the steric diagram for polyalanines this region corresponds to the right-handed 2_7^R helix. The structure of conformations 2 on the surfaces of Figs. 10(b) and 10(c) differs from the structure of this particular helix type. Only the central alanines, for which the angles φ and ψ in Figs. 10(b) and 10(c) are defined, have the structure of the 2_7^R helix. Thus, one can refer to the conformations 2 as to the mixed states, where the central part of the polypeptide chain has the conformation of helix and the outermost parts have the conformation of sheets. Conformation 3 on the surface Fig. 10(d) is also a mixed state. Here one can distinguish one turn of the 3_{10}^R helix and two turns of the 2_7^R helix [see Fig. 8(b)].

Region III corresponds to the structure of right-handed α_R-helix, right-handed 3_{10}^R helix, right-handed π_R-helix, and β-turn. It corresponds to minima 6, 5, and 4 on the potential energy surfaces, Figs. 10(b)–10(d), respectively. Conformation 6 cannot be assigned to any specific type of secondary structure because the chain is too short. Note that conformation 6 is even not a stable one on the potential energy surface of the alanine tripeptide. The most probable types of secondary structures in that region of the potential energy surface are right-handed α_R-helix and β-turn. However, for the formation of a single turn of the α_R-helix (or for the formation of the β-turn) at least four amino acids are needed. Conformation 5 on the potential energy surface of the alanine hexapeptide can be characterized as a partially formed β-turn because the alanine, for which the dihedral angles φ and ψ in Fig. 10(c) are defined has the geometry of β-turn, but its neighbor forms a β-sheet [see Fig. 8(a)]. Conformation 4 on the potential energy surface Fig. 8(b) changes significantly after accounting for the relaxation of all degrees of freedom in the system, and gets outside the region III. In this conformation one can locate fragments of right-handed 2_7^R and 3_{10}^R helixes. The point corresponding to the minimum 4 (after accounting for the relaxation) lies outside regions II and III because angles φ and ψ in Fig. 10(d) are defined for the amino acid between two helix fragments.

Region IV is represented by the structure of the left-handed α_L-helix, the left-handed 3_{10}^L helix, left-handed π_L-helix and β-turn of type II. The fragments with those types of secondary structures are very rare met in native proteins. To form these structures it is necessary to have at least four amino acids, therefore minima 3 on the potential energy surface for alanine tripeptides cannot be compared to any type of the mentioned secondary structures. Region IV corresponds to the conformations 3 and 2 on the surfaces of Figs. 10(c) and 10(d), respectively. Conformation 3 on the surface of Fig. 10(c) corresponds to a partially formed β-turn, because the alanine, for which the dihedral angles φ and ψ in Fig. 10(c) are plotted has the configuration of the β-turn but the neighboring amino acid in the polypeptide chain forms the β-sheet [see Fig. 8(a)]. Conformation 2 on the potential energy surface of Fig. 10(d) lies outside the region IV, but accounting for the relaxation of all degrees of freedom shifts the minimum on the potential energy surface to the allowed region of left-handed α_L and 3_{10}^L helix [see Fig. 8(b)]. The geometry of conformation 2 is similar to the geometry of the left-handed 3_{10}^L helix [see Fig. 8(b)]. The main differences in the structure are caused by the insufficient length of the polypeptide chain to form a regular helix structure.

IV. CONCLUSION

In the present paper the multidimensional potential energy surfaces for amino acid chains consisting of three and six alanines have been investigated and the conformational properties of these systems with respect to the twisting of the polypeptide chain have been described. The calculations have been carried out within *ab initio* theoretical framework based on the density functional theory and accounting for all the electrons in the system. We have determined stable conformations and calculated the energy barriers for transitions between them. Using a thermodynamic approach, we have estimated times of the characteristic transitions between the conformations. It was demonstrated that the transition times lie within the picosecond region. Our estimates are compared with the available molecular-dynamics simulations results, and the correspondence between the results of the two different methods is reported. A strong barrier asymmetry between neighboring stable conformations on the potential energy surface was found.

We compared values of angles φ and ψ for alanine residues experimentally observed in real proteins with the coordinates of minima on the potential energy surfaces. This comparison showed that all profound minima on the potential energy surfaces correspond to the regions in which experimentally measured values of φ and ψ are located. We have also analyzed how the secondary structure of polypeptide chains influences the formation of the potential energy landscapes. For the chains of six amino acids with the secondary structures of sheets and helix the influence of the secondary structure on the stable conformations of the molecule has been demonstrated.

The results of this work can be utilized for modeling more complex molecular systems. For example, the suggested model for the estimation of the characteristic transition times can be used for longer polypeptide chains, also consisting of different amino acids and for estimates of the time of proteins folding. It is also possible to use the results of the present work for testing the applicability and accuracy of different model approaches for the polypeptide description requiring much less computer time than *ab initio* calculations.

ACKNOWLEDGMENTS

This work is partially supported by the European Commission within the Network of Excellence Project EXCELL, by INTAS under the Grant No. 03-51-6170 and by the Rus-

sian Foundation for Basic Research under Grant No. 06-02-17227-a. We are grateful to Dr. A. Korol and Dr. O. Obolensky for their help in preparation of this manuscript. The possibility to perform complex computer simulations at the Frankfurt Center for Scientific Computing is also gratefully acknowledged.

[1] M. Karas and F. Hillenkamp, Anal. Chem. **60**, 2299 (1988).
[2] F. Hillenkamp and M. Karas, Int. J. Mass. Spectrom. **200**, 71 (2000).
[3] M. Karas, U. Bahr, I. Fournier, M. Gluckmann, and A. Pfenninger, J. Mass Spectrom. **226**, 239 (2003).
[4] M. Wind and W. Lehmann, J. Anal. At. Spectrom. **19**, 20 (2004).
[5] J. Fenn, M. Mann, C. Meng, S. Wong, and C. Whitehouse, Science **246**, 64 (1989).
[6] S. B. ndsted Nielsen, J. Andersen, P. Hvelplund, B. Liu, and S. Tomita, J. Phys. B **37**, R25 (2004).
[7] A. Finkelstein and O. Ptizin, *Physics of Proteins* (Moscow University Press "Universitet," Moscow, 2002).
[8] A. Mülberg, *Protein Folding* (St. Petersburg University Press, St. Petersburg, Russia, 2004).
[9] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, and P. Bourne, Nucleic Acids Res. **28**, 235 (2000).
[10] A. Rubin, *Biophysics: Theoretical Biophysics* (Moscow University Press, "Nauka," 2004).
[11] T. Head-Gordon, M. Head-Gordon, M. J. Frisch, Ch. L. Brooks, and J. A. Poplet, J. Am. Chem. Soc. **113**, 5989 (1991).
[12] I. Gould, W. Cornell, and I. Hillier, J. Am. Chem. Soc. **116**, 9250 (1994).
[13] Z. Wang and Y. Duan, J. Comput. Chem. **25**, 1699 (2004).
[14] A. Percel, O. Farkas, I. Jákli, I. Topol, and I. Csizmadia, J. Comput. Chem. **24**, 1026 (2003).
[15] I. Húdaky, P. Húdaky, and A. Percel, J. Comput. Chem. **25**, 1522 (2004).
[16] R. Improta and V. Barone, J. Comput. Chem. **25**, 1333 (2004).
[17] R. Vargas, J. Garza, B. Hay, and D. Dixon, J. Phys. Chem. A **106**, 3213 (2002).
[18] R. Kashner and D. Hohl, J. Phys. Chem. A **102**, 5111 (1998).
[19] D. Wei, H. Guo, and D. R. Salahub, Phys. Rev. E **64**, 011907 (2001).
[20] S. Woutersen, Y. Mu, G. Stock, and P. Hamm, Chem. Phys. **266** (2001).
[21] S. Woutersen, R. Pfister, Y. Mu, D. Kosov, and G. Stock, J. Chem. Phys. **117**, 6833 (2002).
[22] Y. Mu and G. Stock, J. Phys. Chem. B **106**, 5294 (2002).
[23] Y. Mu, D. Kosov, and G. Stock, J. Phys. Chem. B **107**, 5064 (2003).
[24] P. Nguyen and G. Stock, J. Chem. Phys. **119**, 11350 (2003).
[25] W. R. P. Scott, P. H. Hunenberger, I. G. Trioni *et al.*, J. Phys. Chem. A **103**, 3596 (1999).
[26] A. D. MacKerell, D. Bashford, M. Bellott *et al.*, J. Phys. Chem. B **102**, 3586 (1999).
[27] W. D. Cornell, P. Cieplak, C. I. Bayly *et al.*, J. Am. Chem. Soc. **117**, 5179 (1995).
[28] H. Torii and M. Tasumi, J. Raman Spectrosc. **29**, 81 (1998).
[29] R. Schweitzer-Stenner, F. Eker, Q. Huang, and K. Griebenow, J. Am. Chem. Soc. **123** (2001).
[30] Y. Levy and O. Becker, J. Chem. Phys. **114**, 993 (2001).
[31] Z. Shi, C. Olson, G. Rose, R. Baldwin, and N. Kallenbach, Proc. Natl. Acad. Sci. U.S.A. **99**, 9190 (2002).
[32] A. Garcia, Polymer **45**, 669 (2004).
[33] A. Yakubovitch, I. Solov'yov, A. Solov'yov, and W. Greiner, e-print physics/0406093.
[34] A. Yakubovitch, I. Solov'yov, A. Solov'yov, and W. Greiner, e-print physics/0406094.
[35] *Atomic Clusters and Nanoparticles*, edited by C. Guet, P. Hobza, F. Spiegelman, and F. David, Proceedings of Les Houches Summer School of Theoretical Physics, LXXIII, 2000 (EDP Sciences and Springer-Verlag, Berlin, 2001).
[36] *Latest Advances in Atomic Cluster Collisions Fission, Fusion, Electron, Ion and Photon Impact*, edited by A. Solov'yov and J.-P. Connerade (World Scientific Press, Singapore, 2004).
[37] L. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer-Verlag, New York, 1986).
[38] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[39] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[40] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[41] J. B. Foresman and Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods* (Gaussian Inc, Pittsburgh, PA, 1996).
[42] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[43] A. Bax, Protein Sci. **12**, 1 (2003).
[44] S. Sheik, P. Sundararajan, A. Hussain, and K. Sekar, Bioinformatics **18**, 1548 (2002).
[45] D. Voet and J. Voet, *Biochemistry* (John Wiley and Sons, New York, 2004).