Research Article

Journal of Physical Organic Chemistry

Received: 16 June 2008, Revised: 13 August 2008, Accepted: 13 August 2008,

(www.interscience.wiley.com) DOI 10.1002/poc.1445

Hierarchy of the non-covalent interactions in the alanine-based secondary structures. DFT study of the frequency shifts and electron-density features

M. V. Vener[a]*, A. N. Egorova[a], D. P. Fomin[a] and V. G. Tsirelson[a]

The alanine (Ala)-based cluster models of C5, C7, and C10 H-bonds are studied at the DFT/B3LYP level. CPMD/BLYP simulations of the infinite polyalanine α-helix (C13 H-bond) and the two-stranded β-sheets are performed. Combined use of frequency shifts and electron-density features enable us to detect and describe quantitatively the non-covalent interactions (H-bonds) defining the intrinsic properties of Ala-based secondary structures. The energies of the primary N—H···O H-bonds are decreasing in the following way: C13 > C5 ≥ C7 > C10. The energies of the secondary N—H···O, N—H···N, and H···H interactions are comparable to those of the primary H-bonds (~4.5 kcal/mol). Side chain–backbone C—H···O interaction is found to be the weakest non-covalent interaction in the considered species. Its energy is ~0.5 kcal/mol in the infinite polyalanine α-helix. Quantum-topological electron-density analysis is found to be a powerful tool for the detection of secondary non-covalent interactions (C=O···H—C and H···H) and bifurcated H-bonds, while the frequency shift study is useful for the identification and characterization of primary or secondary H-bonds of the N—H···O type. Copyright © 2008 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: H-bonds; secondary structures; topological analysis; frequency shifts

INTRODUCTION

Secondary structures are the first 3D level of organization in proteins and therefore they play an important structure-forming role.[1] The local geometries of the considered species are essentially controlled by the ability of the amide N—H groups to bind to a more or less remote carbonyl group, giving rise to the so-called C5, C7, C10, and C13 H-bonds.[2–5] Due to the technical limitations, the structure of peptides in the molecular crystalline form is usually studied by the low-resolution diffraction methods. As a result, the H atom positions cannot be obtained with the accuracy that is needed for understanding the nature of the non-covalent interactions at the atomic level. Spectroscopic studies (linear-infrared (IR), two-dimensional IR spectroscopy, vibrational circular dichroism) of oligo- and polypeptides in a solution are often blurred by the spectral overlap and solvent or temperature effects.[3] These factors complicate a quantitative description of non-covalent interactions defining the intrinsic local conformational preference of the peptide secondary structures.

During last decade, the laser spectroscopy, including infrared (IR)/UV double-resonance spectroscopy[6] is successfully applied for the investigation of non-covalent interactions, especially H-bonds, in the gas-phase capped peptides, that is oligopeptides in which the C- and N-termini are chemically protected. These modified peptides serve as the model molecules mimicking the central part of a peptide or protein chain and whose study focuses on the intrinsic properties of the chain itself.[5] However, existence of several quasi-isoenergetic conformers for the considered species complicates the assignment of the experimental bands in IR/UV hole-burning spectra.[7–9] To overcome these difficulties, the quantum-chemical computations are widely used. In particular, the geometry and energy of the non-covalent interactions were studied using the cluster[10–16] and infinite-model[17–21] approximations. The electron-density characteristics underlying the non-covalent interactions[22,23] and the harmonic frequencies[5,7–9,20,21,24–30] of the H-bonds were computed as well. In most of these papers either the particular secondary structure or the specific vibrational region of the capped peptides were studied. Moreover, the harmonic frequencies of the H-bonded groups and electron-density properties of the non-covalent interactions have never been considered simultaneously till now. This is why a precise understanding the non-covalent interactions that control the 3D character of the peptide secondary structures has not reach yet. Solution of this problem calls for a systematic approach, in which the effects of complexity, for example introduction of large residues, can be

---

* *Correspondence to: M. V. Vener, Department of Quantum Chemistry, Mendeleev University of Chemical Technology, Miusskaya Square 9, 125047 Moscow, Russia.*
E-mail: venermv@muctr.edu.ru

a *M. V. Vener, A. N. Egorova, D. P. Fomin, V. G. Tsirelson*
Department of Quantum Chemistry, Mendeleev University of Chemical Technology, Miusskaya Square 9, 125047 Moscow, Russia

studied step by step.[7] In the present paper, we attempt to develop such approach. It is based on the combined use of the DFT methods with/without periodic boundary conditions and the simultaneous consideration of the frequency shifts of the H-bonded groups and electron-density properties of the non-covalent interactions. The size of the species under consideration goes up to 11 residues and all types of the secondary structures, including the protein α-helixes are studied at the same computational level. The alanine (Ala)-based oligo- and polypeptides seems to be the best models capable of reflecting the intrinsic backbone preferences since the Ala residue bears the simplest generic side chain, that is a methyl group, and is therefore devoid of the extra (non-covalent) interactions.

The aim of this paper is twofold: (i) we aim to describe quantitatively the backbone–backbone and the simplest side chain–backbone interactions in Ala-based secondary structures; (ii) we intend to establish a hierarchy of these interactions. The theoretical study of conformational landscapes in the considered oligo- and polypeptides is beyond the scope of the present work.

THE SECONDARY-STRUCTURE MODELS AND COMPUTATIONAL METHODS

Two sets of the secondary-structure models were used in this study. The first one includes the cluster (gas-phase) models of the oligoalanines. In so doing, the conformers of CH_3—CO—Ala—NH—CH_3 molecule, corresponding to C5 and C7 H-bonds and conformers of CH_3—CO—Ala—Ala—NH—CH_3 molecule, corresponding to C10 H-bonds, were considered. The DFT calculations were carried out using the PC version[31] of the GAMESS(US) program package.[32] Geometries of the considered conformers were optimized at B3LYP/6-31+G**. As it was already demonstrated, this level of approximation gives a reasonable description of the structure,[5,8] relative stability,[9] and harmonic frequencies of oligopeptides.[5,26–28] The minimum-energy states of the structures have been confirmed by calculating the harmonic frequencies. The topological electron-density properties[33] were evaluated with the AIMPAC computer program suite.[34] The following electron-density features at the H···A (A = O, N, C, H) bond critical point in the electron density were considered: the values of electron density, ρ_b, and the potential energy density, V_b. The first value may be measured experimentally by means of the X-ray diffraction, while V_b is used to estimate the energy E_{int} of the particular non-covalent interaction:[35,36]

$$E_{int} = (1/2)V_b \tag{1}$$

Details of the electron-density topology analysis of the considered gas-phase species were published elsewhere.[37]

Infinite models of the polyalanine parallel and antiparallel two-stranded β-sheets and α-helix were taken from the literature.[18,38] The DFT calculations with periodic boundary conditions are carried out using the CPMD program package.[39] The CPMD calculations use the BLYP functional[40,41] with Troullier–Martins[42] pseudo-potentials for core electrons. The kinetic energy cut-off for the plane wave basis set is 100 Ry and k-space sampling is limited to the Γ point. The minimum-energy states of the structures have been confirmed by calculating the harmonic frequencies. The BLYP/plane-wave computations have been successfully used to study the structure and harmonic frequencies of periodical systems with H-bonds of different strength, for example refer References.[43–48] According to References,[47,48] the kinetic energy cut-off of 100 Ry yields reasonable results for the considered systems. Obtained geometries were used in computation of periodical electronic wave functions by the CRYSTAL98 program[49] at the B3LYP/6-31G** level of approximation. The topological analysis of the periodical electron density was performed by the TOPOND computer program.[50] Details of the electron-density topology analysis of the periodical models were published elsewhere.[51]

We focus on the N—H stretch region (amide A) which provides a direct information on the number and intensity of H-bonds in gas phase,[5] the amide I and II regions, corresponding to C=O stretching and N—H bending motions, respectively. Frequency shifts of the C^α—H stretching vibration are also considered; here C^α stands for the backbone carbon atom. The B3LYP/6-31+G** harmonic frequencies were scaled with the scaling factor of 0.9648,[52] while the unscaled values were used for the BLYP/plane-wave frequencies. Unfortunately, IR intensities are not available for harmonic vibrations computed by CPMD. This is why we cannot say which bands are observable in the IR and which in the Raman spectra of the infinite models.

RESULTS

Cluster models of oligoalanines

Computed characteristics of the three low-energy conformers of CH_3—CO—Ala—NH—CH_3 C5, C7$_{eq}$, and C7$_{ax}$ are given in Table 1. Relative stability of the considered species and the geometrical parameters of the H-bonded fragments computed at the B3LYP/6-31+G** level of approximation agree with the literature data.[13,14] Computed ρ_b values at the H···O bond critical point are found to be in reasonable agreement with the available data derived from the experimental electron density of the amino acid crystals[53] (Table 1). The most stable conformer C7$_{eq}$ (the inverse γ-turn) has no additional non-covalent interactions, while the local-minimum structure C7$_{ax}$ (the classic γ-turn) exhibits an additional (secondary) interaction of the C^β—H···O type (Fig. 1 in Reference[37]); here C^β stands for the side chain carbon atom. Its energy equals to 2.1 kcal/mol, that is smaller than that of the primary N—H···O interaction (~4.8 kcal/mol).

Calculated values of the frequency shifts for the N—H and C=O stretching and N—H bending vibrations agree with the available experimental data[54] fairly well (Table 1). In accord with the previous computations,[24] the relative IR intensity of the N—H stretch involved in H-bond is much larger that one of the free N—H group, while the opposite tendency is observed for the relative IR intensity of the C=O stretch (Table 1). According to our calculations, the absolute values of the frequency shift of the stretching vibrations of the C^β—H group involved into the C^β—H···O interaction is around 10 cm^{-1}. It is less than a width of 15 cm^{-1} usually assumed in simulations of IR spectra by positioning the Lorenzian bands.[55] We conclude that using the B3LYP/6-31+G** approximation for computations of the harmonic frequencies and quantum topological analysis of electron density in the cluster models of oligopeptides seems to be verified.

The C10 conformers of CH_3—CO—Ala—Ala—NH—CH_3 are shown in Fig. 2 of Reference.[37] Computed values of the geometrical parameters of the H-bonded fragments and topological properties at the H-bond critical point are given in

Table 1. Relative energies and structural parameters for H-bonds in B3LYP/6-31+G** (in regular letters) and MP2/aug-cc-pVDZ[a] (in italics) geometries of the $C7_{eq}$, C5, and $C7_{ax}$ conformers of CH_3–CO–Ala–NH–CH_3. Topological properties at the H···O bond critical point in these structures computed using B3LYP/6-31+G** (in regular letters) and the corresponding experimental values[b], obtained for the crystal structures of the six amino acids (in italics). Calculated values of the frequency shifts (in regular letters) and the corresponding experimental values[c] obtained for $C7_{eq}$ and C5 conformers of CH_3–CO–Ala–NH–CH_3 (in italics). The H-bonds are defined in Fig. 1 of this paper and Fig. 1 in Reference.[37]

			Structural parameters			Topological properties		Frequency shifts and relative intensities[d]		
Conformer	Relative energy, kcal/mol	H-bonded fragment	R(X···O), Å	R(H···O), Å	∠(XHO), deg	ρ_b, a.u.	$\lvert V_b \rvert$, a.u.	$\Delta\nu$(N–H), cm^{-1}	$\Delta\nu$(C=O), cm^{-1}	$\Delta\delta$(N–H), cm^{-1}
$C7_{eq}$	0.0[e]	N–H···O	2.968	2.092	143.0	0.020	0.0155	100 (~5)	34 (~0.5)	−49 (~1)
	0.0		*2.91*	*2.02*	*144.8*	*0.015 ÷ 0.035*		*115*	*25*	*−42*
C5	1.09	N–H···O	2.674	2.209	106.2	0.019	0.0157	39 (~3)	18 (~0.3)	−41 (~0.4)
								71	*17*	*−34*
$C7_{ax}$	2.51	N–H···O	2.858	1.934	149.8	0.028	0.0214	148 (~9)	24 (~0.3)	−42 (~1.5)
	2.28		*2.82*	*1.88*	*151.7*	*0.015 ÷ 0.035*		—	—	—
		C^{β}–H···O	3.138	2.553	112.6	0.010	0.0067			
			3.08	*3.08*	*113.5*	*~0.01*		—		—

[a] Reference.[13]

[b] Reference.[53]

[c] IR study of CH_3–CO–Ala–NH–CH_3 isolated in an argon matrix.[54]

[d] Scaled values of the computed harmonic frequencies of the 'free' group vibrations in $C7_{eq}$ are: ν(N–H) = 3506 cm^{-1}, ν(C=O) = 1689 cm^{-1}, δ(N–H) = 1483 cm^{-1}, here ν and δ denote stretching and in-plane bending vibrations, respectively. Relative intensities (I/I_0) are given in parentheses; I_0 stands for the computed IR intensity of the corresponding free group.

[e] Total energy is −495.898557 a.u.

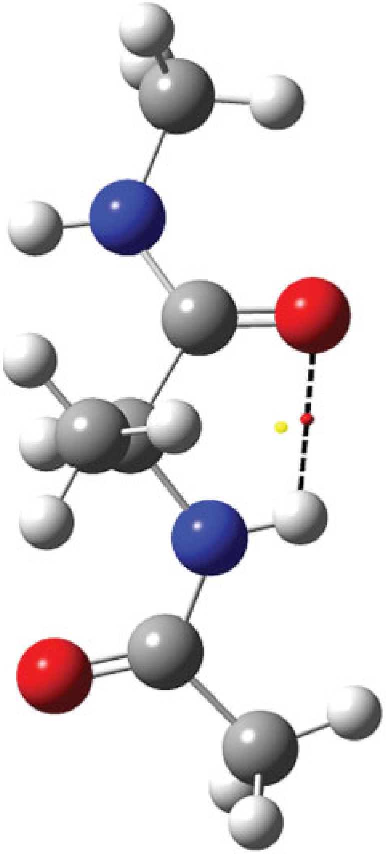

Figure 1. Molecular graphs and the critical point pattern in the C5 conformer of CH_3–CO–Ala–NH–CH_3. The small dots indicate the position of the critical points in the electron density: red dot stands for the H-bond critical point, while yellow dot stands for the ring critical points. The primary N–H. . .O H-bond (Table 1) is given by the broken line.

Table 2. In contrast to C7, the C10 structures exhibit bifurcated or 'three-center' H-bonds, *cf.* Figs. 1 and 2 in Reference.[37] Bifurcated H-bonds in the C10B and C10D conformers involve the H···H interaction. Its energy equals ~3.0 kcal/mol, in reasonable agreement with the value of 2.3 kcal/mol estimated for the N—H···H—C bond in valine.[56] The energy of the secondary C—H···O interactions in the C10 structures is ~2.0 kcal/mol.

Computed values of the frequency shifts and relative IR intensities of the harmonic vibrations of the groups forming H-bonds are given in Table 2. In accord with the literature data,[7] computed shifts of these vibrations are found to be very sensitive to the intrinsic local conformational preference of the considered secondary structure. The following observations can be made:

The Amide A motif. Computed $\Delta\nu$(N—H) values in the C10 structures (<100 cm^{-1}) are found to be smaller those ones, obtained for the C7 structures (100–150 cm^{-1}), *cf.* Tables 1 and 2, in agreement with the results of the IR investigations of short capped peptides chains in gas-phase.[7] Similar to the results obtained in Reference[57] for the N—H···O intramolecular H-bonded systems using the B3LYP/6-311++G** approximation, the intensity of the N—H stretching vibrations involved into H-bonds, increases strongly in comparison with that of the free N—H group.

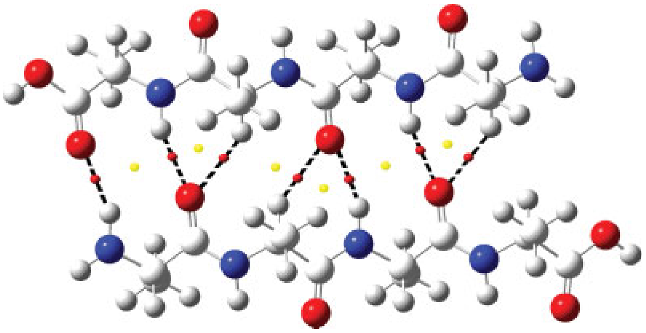

Figure 2. Molecular graphs and the critical point pattern in the infinite polyalanine antiparallel two-stranded β-sheet. Refer the captions of Fig. 1 for the color coding. The interstrand H-bonds (Table 3) are given by the broken lines. For the sake of simplicity, the terminal N and O atoms are saturated with H.

The Amide I motif is rather compact in C10A and C10D, $\Delta\nu(CO) < 40\ cm^{-1}$ (Table 2). It is larger (~60 cm^{-1}) in C10B due to the additional N_{i+1}—H_b···O_{i+3} interaction which generates γ-turn, in accord with the results of mid-IR spectroscopic studies of the capped peptides in the gas phase.[8] The intensity of the stretching vibrations of the C=O groups involved into H-bonds, practically does not change in comparison with that of the free C=O group.

The Amide II motif. In accord with the literature data,[26] the hydrogen-bonding formation down-shifts the amide I frequency and up-shifts the amide II band. Computed values of $\Delta\delta$(N—H) are around $-50\ cm^{-1}$ for the primary H-bonds, negative sign indicates a blue shift. Calculated shifts of the amide II band agree with the available experimental values of $-40\ cm^{-1}$ established for isolated peptides.[58]

C—H stretch region. Computed values of $\Delta\nu(C^{\alpha}$—H) vary from -33 to $-50\ cm^{-1}$. Negative numbers indicate a blue shift. These values are in agreement with the available experimental and calculated IR frequency shifts obtained for different systems with the C—H···Y fragment, where Y = O and Cl.[59] Due to this blue shift the C^{α}—H bands fall into the region of the CH_3 stretches. In accord with the literature data,[60] the intensity of the C^{α}—H stretching vibrations involved into the C^{α}—H···O interaction, is smaller that one of the free C^{α}—H group.

Polyalanines: infinite versus cluster models

It is known[2,16] that each carbonyl oxygen atom in the antiparallel β-sheet structure that juts into the region between the two strands has access to two sorts of proton donors (Fig. 2). This oxygen atom forms two types of the interstrand H-bonds: N—H···O and C^{α}—H···O. Geometrical parameters of these fragments are given in Table 3. Computed values of the (N)H···O distances 2.071 and 2.085 Å are found to be a bit longer the corresponding values 1.93–2.04 Å obtained for the antiparallel β-sheet models consisting of two strands of four glycine residues.[61] According to the values of ρ_b and V_b (Table 3), the N—H···O bonds are stronger than C^{α}—H···O. Their energies are 4.1 and 2.0 kcal/mol, respectively. Special attention should be paid to the intrastrand N—H···O bonds, that are considered in the cluster models of polypetides, for example refer References.[15,23,61] The five-member pseudo-cycles, corresponding to

Table 2. Relative energies, geometrical parameters of the X–H$\cdots$Y fragments, where X = C, N, Y = O, N, H, and topological properties at the H-bond critical point, computed for the C10 conformers of CH_3–CO–Ala–Ala–NH–CH_3

			Structural parameters			Topological values		Frequency shifts and relative intensities[a]			
Con-for- mer	Relative energy, kcal/mol	H-bonded fragment or intraction groups	$R(X\cdots Y)$, Å	$R(H\cdots Y)$, Å	$\angle$(XHY), deg	ρ_b, a.u.	$\lvert V_b\rvert$, a.u.	$\Delta\nu$(N–H), cm^{-1}	$\Delta\nu$(C=O), cm^{-1}	$\Delta\delta$(N–H), cm^{-1}	$\Delta\nu$(C$^{\alpha}$–H), cm^{-1}
C10A	0.0[b]	N_i–$H_b\cdots O_{i+3}$	3.143	2.172	159.7	0.015	0.0113	52 (~10)	24 (~1)	−45 (~2)	—
		N_i–$H_b\cdots N_{i+1}$	2.768	2.353	103.4	0.016	0.0122		—		—
		N_{i+1}–$H\cdots N_{i+2}$	2.793	2.332	106.5	0.016	0.0118	—	—	−26 (~2)	—
C10B	1.11	N_i–$H_b\cdots O_{i+3}$	3.378	2.433	155.1	0.009	0.0064	22 (~3)	56 (~1)	−64 (~2)	—
		N_{i+1}–$H_b\cdots O_{i+3}$	3.054	2.031	139.2	0.024	0.0187	83 (~4)		−13 (~3)	—
		$C^{\beta}{}_{i+1}$–$H\cdots O_{i+2}$	3.176	2.553	92.6	0.010	0.0064	—	14 (~1)	—	—
		N_i–$H_b\cdots H_b$–N_{i+1}	—	1.930	—	0.013	0.0098	—	—	—	—
C10D	2.88	N_i–$H_b\cdots O_{i+3}$	3.005	2.026	160.8	0.020	0.156	99 (~8)	38 (~1)	−37 (~1)	—
		$C^{\alpha}{}_{i+2}$–$H_b\cdots O_{i+3}$	3.434	2.597	133.0	0.009	0.0057	—		—	−33 (~0.5)
		$C^{\alpha}{}_{i+2}$–$H_b\cdots H$–$C^{\alpha}{}_{i+1}$	—	1.901	—	0.015	0.0095	—	—	—	−48 (~0.5)[c]

The H-bonds are defined in Fig. 2 of Reference.[37] H_b indicates that H-atom participates in a bifurcated H-bond. Calculated values of the frequency shifts and relative intensities of harmonic vibrations of the groups forming H-bonds are given in the last four columns.

[a] Scaled values of the computed harmonic frequencies of the 'free' group vibrations in C10A are: ν(N–H) = 3498 cm^{-1}, ν(C=O) = 1688 cm^{-1}, δ(N–H) = 1468 cm^{-1}, ν(C$^{\alpha}$–H) = 3002 cm^{-1}. Relative intensities I/I_0 are given in parentheses. I_0 stands for the computed IR intensity of the corresponding free group.

[b] Total energy is −743.243722 a.u.

[c] The data for the $C^{\alpha}{}_{i+1}$–H group.

Table 3. Geometrical parameters of the X–H···O fragments, X = N or C$^{\alpha}$ and topological properties at critical point of the H-bonds, computed for the infinite polyalanine antiparallel β-sheet and the α-helix

		Structural parameters			Topological properties	
Structure	H-bonded fragment	R(X···O), Å	R(H···O), Å	∠(XHO), deg	ρ_b, a.u.	$\|V_b\|$, a.u.
Antiparallel β-sheet[a]	N–H···O	3.058	2.071	161.3	0.018	0.0137
	N–H···O	3.071	2.085	161.0	0.018	0.0132
	C$^{\alpha}$–H···O	3.361	2.414	144.5	0.011	0.0073
	C$^{\alpha}$–H···O	3.378	2.425	145.6	0.010	0.0071
α-helix	N$_i$–H···O$_{i+4}$	2.945–2.956	1.948–1.960	163.4–163.7	0.024	0.0178
	C$^{\beta}_i$–H···O$_{i+3}$	3.893–3.928	3.037–3.075	135.1–135.5	0.003	0.0015

[a] The H-bonds are defined in Fig. 2.

the intrastrand H-bonds, are not detected in the present study for the infinite antiparallel β-sheet. Possible explanation of this finding is based on the different structural parameters of the intrastrand N—H···O fragments in the infinite β-sheet and the C5 conformer of CH_3—CO—Ala—NH—CH_3. The H···O distance is 2.37 Å in the β-sheet structure, while it equals to 2.209 Å in the C5 conformer (Table 1).

Computed values of the selected harmonic frequencies of the β-sheet structure are given in Table 4. Frequencies of the N—H, C$^{\alpha}$—H and C═O groups are doubled due to the fact that every second group does not involved into interstrand H-bond formation. For the sake of simplicity, the latter groups will be called as 'free.'

The Amide A motif. Computed values of stretching frequencies of the N—H groups, involved into the interstrand H-bonds, vary from 3374 to 3379 cm^{-1} (Table 4). The hydrogen-bonding network down-shifts the ν(N—H) frequencies on ~85 cm^{-1}, in accord with the experimental value of ~75 cm^{-1} established for the cyclic gramicidin S (a 15-residue peptide[62]) forming internal bonds similar to a β-sheet structure.

The Amide I motif. The frequencies of the 'free' C═O groups vary from 1629 to 1646 cm^{-1} (Table 4). The frequencies of the C═O groups involved into intrastrand H-bonds shifted on ~40 cm^{-1} to red. It should be noted that the simulated IR spectrum for the antiparallel two-stranded β-sheet segment with octaamide strands shows two bands in the considered region with the gap ~50 cm^{-1}.[63] Very recently, the Amide I region for the antiparallel β-sheets was studied in the cluster approximation.[30] In four-stranded β-sheets based upon four glycine residues the most intense band locates at ~1650 cm^{-1},[30] that is close to the 1594–1613 cm^{-1} region, computed in the present study.

The Amide II motif locates in the 1460–1510 cm^{-1} region. In accord with previous simulations,[63] it splits into two bands separated by ~50 cm^{-1}.

Table 4. Selected harmonic frequencies (cm^{-1}) of the infinite polyalanine antiparallel β-sheet and α-helix

	α-helix		
Antiparallel β-sheet	BLYP/plane wave	Experiment[a]	Assignment
3463–3464	—	—	ν(N–H), free
3374–3379	3350–3352; 3328–3341	3321	ν(N–H), H-bonded
3044–3047	—	—	ν(C$^{\alpha}$–H), H-bonded
3018–3031	3033–3034; 3028–3029	2987	$\nu(CH_3)_{as}$
2994–2995	2981–2984	2970[b]	ν(C$^{\alpha}$–H), free
2960	2967–2968	2929	$\nu(CH_3)_s$
1629–1646	—	—	ν(C═O), free
1594–1613	1619–1621; 1585–1599	1658	ν(C═O), H-bonded
1504–1512	1501–1521	1518	δ(N–H), H-bonded
1456–1459	—		δ(N–H), free

[a] Experimental data from polarized IR and Raman spectra[66] on oriented films of α-helical polyalanine. Frequencies of a symmetry species, that are IR and Raman active, are given in the table.
[b] Computed value.[66]

C—H stretch region. Due to the H-bond formation, the C^{α}—H stretch frequencies shift to blue by ~50 cm^{-1} (Table 4). This band overlaps with the stretching vibrations of the methyl groups. This is why its experimental detection seems to be problematic.

We were unable to localize the infinite polyalanine parallel β-sheet structure that corresponds to the global or local minimum. Variations of some CPMD internal parameters (kinetic energy cut-off, convergence criteria for optimization runs, step length in the numerical calculation of harmonic frequencies, etc.) change the number and values of the imaginary frequencies. However, the infinite parallel β-sheet always has, at least, one imaginary frequency ~35 cm^{-1}.

All C═O and N—H groups in the infinite polyalanine α-helix involved into the hydrogen-bonding network. Computed values of the (N)H···O distance equal to 1.948–1.960 Å (Table 3). These values agree nicely with the minimum H···O distances of ~1.95 Å computed in[64] for α-helix polyalanines contaning up to 18 amino acid residues. In accord with the previous computational study,[15] each C_i═O group contacts two H—N donors: one from a 13-member loop (H—N_{i+4}) and the another one from the 10-member loop (H—N_{i+3}). CPMD calculations give for the latter contact R(O···H) = 3.036 Å. This distance is a bit shorter that one in the C_i═O···H—C^{β}_{i+3} contact of ~3.06 Å, refer Table 3. However, the topological electron-density analysis does not detect the non-covalent interaction between the C_i═O and H—N_{i+3} groups. This result shows the limited applicability of the geometrical criteria in detection of the non-covalent interactions. According to our calculations, refer Table 3, each carbonyl in the infinite polyalanine α-helix forms the two types of H-bonds: C_i═O···H—N_{i+4} and C_i═O···H—C^{β}_{i+3}. Their energy is ~5.5 and ~0.5 kcal/mol, respectively. The C_i═O···H—C^{β}_{i+3} side chain–backbone interaction is absent in polyglycines. We can speculate that due to it Ala stabilizes the α-helix conformation, while glycine, in contrast, destabilizes it.[65]

Computed values of the harmonic frequencies are in reasonable agreement with the experimental data from polarized IR and Raman spectra[66] on oriented films of α-helical polyalanine, refer Table 4. The largest deviations exist for the CH_3 stretches. However, the calculated value of the splitting between the two CH_3 bands, ~60 cm^{-1}, agrees with the experimental value of 58 cm^{-1}.

The Amide A motif. Computed values of the N—H stretching frequencies locate around 3340 cm^{-1}. We note that IR–UV double-resonance spectra of gramicidin A–D are characterized by broad absorption band around 3300 cm^{-1} that was assigned to the N—H stretch of the helix-like structures.[62]

The Amide I motif. The frequencies of the C═O stretches split into two bands, around 1620 and 1590 cm^{-1} (Table 4). Very recently, the Amide I region of α-helical capped polyalanines acetyl$(Ala)_N NH_2$ was studied in the ONIOM and pure DFT approximation.[29] As the peptide becomes longer, the Amide I band becomes both more intense and more red shifted. For $N = 17$ the most red-shifted absorption locates at 1657 cm^{-1},[29] that is close to the values obtained in the present study.

The Amide II motif locates in the 1501–1521 cm^{-1} region in agreement with the 1510–1540 cm^{-1} region, calculated for the infinite polyalanine α-helix using DFT in the generalized gradient approximation.[21]

DISCUSSIONS

In the present work, we try to take advantage of the combined use of the topological electron-density analysis and harmonic frequencies computations in order to establish a hierarchy of the non-covalent interactions in the Ala-based secondary structures. First of all, we discriminate between the primary and secondary interactions. This is quite natural classification, since the primary N—H···O interaction defines the type of the secondary structure, giving rise to the so-called C5, C7, C10, and C13 H-bonds. Its strength, estimated using different computed properties of the N—H···O fragment, is presented in Table 5. The C13 H-bond is the strongest primary interaction in the considered species. The energies of the primary N—H···O H-bonds are decreasing in the following way:

$$C13 > C5 \sim C7 > C10 \qquad (2)$$

All parameters, presented in Table 5, give a similar scale of strength. On the other hand, the use of the particular property may lead to ambiguous conclusions. Indeed, consideration of the H···O distance is not sufficient due to the fact that the N—H···O fragment is nonlinear to different extent in different types of the secondary structures. The computed values of the electron density are quite close to each other. Frequency shifts are very sensitive to the particular type of the H-bond, however to get an unambiguous result, one has to consider several spectral regions. Summing up, a quantitative description of the primary

Table 5. Scale of strengths (expressed in terms of frequency shifts, the values of electron density at critical point of the considered H-bonds, and the interaction energies E_{int}[a]) for the primary N–H···O H-bonds in the alanine-based secondary structures

Type of the interaction	R(H···O), Å	$\Delta\nu$(N–H), cm^{-1}	$\Delta\nu$(C═O), cm^{-1}	$\Delta\delta$(N–H), cm^{-1}	ρ_b, a.u.	E_{int}, kcal/mol
C5	2.209	39	18	−41	0.019	4.9
C7[b]	2.092	100	34	−29	0.020	4.8
C10[b]	2.172	52	24	−45	0.015	3.5
β-sheet structure	2.081	~85	~35	−50	0.018	4.1
C13[c]	1.950	110–130	~25 and ~60	−55	0.024	5.5

[a] Interaction energy E_{int} is estimated from the computed V_b value for the considered interaction, refer Eqn (1).
[b] The primary N–H···O H-bond in the global-minimum conformer of the considered capped oligoalanine.
[c] The frequency shifts of the vibrations in α-helix were computed using the harmonic frequencies of the corresponding 'free' groups in the antiparallel β-sheet structure.

Table 6. Energies (kcal/mol) of the secondary non-covalent interactions in the alanine-based secondary structures

Non-covalent interaction	Energy (structure)
N–H···O	5.8 (C10B)
N–H···N	3.7 (C10A)
H···H	3.1 (C10B); 3.0 (C10D)
C^{α}–H···O	3.3 (β-sheet); 1.8 (C10D)
C^{β}–H···O	2.1 ($C7_{ax}$); 2.0 (C10B); 0.5 (C13)

interactions requires at least two different computed properties of the N—H···O fragment.

Secondary interactions cover backbone–backbone N—H···O, N—H···N, H···H, C^{α}—H···O H-bonds, and the side chain–backbone C^{β}—H···O interaction. The energies of the secondary N—H···O, N—H···N, and H···H interactions are comparable to those of the primary H-bonds, *cf.* Tables 5 and 6. These secondary interactions play an essential role in stabilizing the C10 structures. Energy of the C^{α}—H···O H-bonds varies from ~3.3 kcal/mol in the antiparallel β-sheet structure to ~1.8 kcal/mol for the C10D conformer of CH_3—CO—Ala—Ala—NH—CH_3. It should be noted that the average stabilization energy of a C^{α}—H···O interaction in transmembrane proteins was estimated to be 1.4 kcal/mol.[67] The C^{α}—H···O H-bonds are characterized by relatively large frequency shifts, Tables 2 and 4. Their experimental detection seems to be possible in the case of the C^{α}—D···O bonds, because the deuteration shifts the C^{α}—D stretching frequency far away from those of the methyl groups. We note that the dependence of the C^{α}—D stretching frequency on conformational preferences of the Ala oligopeptides has been studied very recently.[27]

The C^{β}—H···O side chain–backbone interactions appear to be the weakest in the Ala-based secondary structures (Table 6). Their energies vary from ~2 kcal/mol in C7 and C10 structures to 0.5 kcal/mol in the infinite polyalanine α-helix. According to the results of the present study, the C^{β}—H···O interactions are practically undetectable by the IR spectroscopy due to the small values of frequency shifts of the C^{β}—H stretching vibration.

Comparison of the results obtained in the present study with those from References [23,29,30,37,61] highlights differences in the properties of the polypeptides, computed using the cluster and infinite-model approximation. (i) In the cluster model of the antiparallel β-sheet, C5 *intrastrand* H-bonds are either disturbed or enhanced upon the formation of *interstrand* H-bonds.[61] In the infinite model C5 *intrastrand* H-bonds are disturbed strongly upon the formation of *interstrand* H-bonds. This is why the five-member pseudo-cycles, corresponding to the intrastrand H-bonds, are not detected for the infinite antiparallel β-sheet. (ii) Weak interstrand interactions of the C^{β}—H···H—$^{\beta}C$ and C^{α}—H···H—$^{\alpha}C$ type were detected for the cluster models of the two-stranded β-sheets.[23,37] However, they are not detected in the infinite-model approximation. This is due the geometrical changes in the H···H distances, caused by optimization of the infinite polyalaninie antiparallel β-sheet in the BLYP/plane-wave approximation performed in the present work. Resulted values of these distances appear to be much longer those, used in Reference[37], for example 3.326 Å *versus* 2.787 Å for the C^{β}—H···H—$^{\beta}C$ fragment and 2.476 Å *versus* 2.289 Å for the C^{α}—H···H—$^{\alpha}C$ fragment. (iii) In the cluster models position of the Amide I region depends on the number of residues.[29,30] It is usually shifted to blue as compared to its position in the corresponding infinite model.

CONCLUSIONS

Combined use of the topological electron-density analysis and harmonic frequencies computations enable us to establish the hierarchy of the non-covalent interactions in the Ala-based oligo- and polypeptides. The energies of the primary N—H···O H-bonds are decreasing in the following way: C13 > C5 ≥ C7 > C10. The energies of the secondary N—H···O, N—H···N, and H···H interactions are comparable to those of the primary H-bonds (~4.5 kcal/mol). Secondary side chain–backbone C^{β}—H···O interaction is found to be the weakest non-covalent interaction in the considered species. Its energy varies from ~0.5 kcal/mol in the infinite polyalanine α-helix to ~2 kcal/mol in the $C7_{ax}$ conformer (classic γ-turn).

We conclude that the quantum-topological electron-density analysis and frequency shifts study may be considered as complimentary. The first one is a powerful tool for the detection of secondary C—H···O and H···H interactions and bifurcated H-bonds, while the second is useful for the identification of primary or secondary N—H···O H-bonds.

Results of the present study highlight the differences in the properties of the polypeptides, computed in the cluster and infinite-model approximation. (i) The five-member pseudo-cycles, corresponding to the intrastrand H-bonds, are not detected for the infinite antiparallel β-sheet, because C5 *intrastrand* H-bonds are disturbed strongly upon formation of *interstand* H-bonds. In the cluster approximation C5 *intrastrand* H-bonds are either disturbed or enhanced upon the formation of *interstand* H-bonds (ii) In the cluster model of the antiparallel β-sheet and α-helix the position of the Amide I region depends on the number of residues. It is usually shifted to blue as compared to its position in the corresponding infinite model.

Acknowledgements

This study was supported by the Russian Federal Agency for Education (Program 'Development of the Highest-School Scientific Potential: 2006–2008,' Project 2.1.1.5051) and Russian Foundation for Basic Research (grants 07-03-00702 and 08-03-00515-a). MVV thanks Prof. J. Hutter and Dr C. Tuma for help in numerical calculations and Drs K. A. Lyssenko, J. Ireta, and I. G. Shenderovich for useful discussions.

REFERENCES

[1] J. Venkatraman, S. C. Shankaramma, P. Balaram, *Chem. Rev.* **2001**, *101*, 3131–3152, and references therein.
[2] J. Bella, H. M. Berman, *J. Mol. Biol.* **1996**, *264*, 734–742.
[3] E. Vass, M. Hollósi, F. Besson, R. Buchet, *Chem. Rev.* **2003**, *103*, 1917–1954, and references therein.
[4] A. K. Thakur, R. Kishore, *Biopolymers* **2006**, *81*, 440–449.
[5] V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel, M. Mons, *J. Phys. Chem. A* **2007**, *111*, 7347–7354.
[6] T. S. Zwier, *J. Phys. Chem. A* **2001**, *105*, 8827–8869, and references therein.

[7] W. Chin, F. Piuzzi, I. Dimicoli, M. Mons, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033–1048.

[8] I. Compagnon, J. Oomens, G. Meijer, G. von Helden, *J. Am. Chem. Soc.* **2006**, *128*, 3592–3597.

[9] V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel, M. Mons, *Angew. Chem. Int. Ed.* **2007**, *47*, 2463–2466.

[10] A. Perczel, M. A. McAllister, P. Csaszar, I. G. Csizmadia, *J. Am. Chem. Soc.* **1993**, *115*, 4849–4858.

[11] R. Kaschner, D. Hohl, *J. Phys. Chem. A* **1998**, *102*, 5111–5116.

[12] I. Topol, S. K. Burt, T. Deretey, T.-H. Tang, A. Perczel, A. Rashin, I. G. Csizmadia, *J. Am. Chem. Soc.* **2001**, *123*, 6054–6060.

[13] R. Vargas, J. Garza, B. P. Hay, D. A. Dixon, *J. Phys. Chem. A* **2002**, *106*, 3213–3218.

[14] R. Improta, V. Barone, *J. Comput. Chem.* **2004**, *25*, 1333–1341.

[15] Z. Varga, A. Kovács, *Int. J. Quant. Chem.* **2005**, *105*, 302–312.

[16] S. Scheiner, *J. Phys. Chem. B* **2006**, *110*, 18670–18679.

[17] R. Improta, V. Barone, K. N. Kudin, G. E. Scuseria, *J. Am. Chem. Soc.* **2001**, *123*, 3311–3322.

[18] J. Ireta, J. Neugebauer, M. Scheffler, A. Rojo, M. Galván, *J. Phys. Chem. B* **2003**, *107*, 1432–1437.

[19] J. Rossmeisl, I. Kristensen, M. Gregensen, K. W. Jakobsen, J. K. Nørskov, *J. Am. Chem. Soc.* **2003**, *125*, 16383–16386.

[20] S. Franzen, *J. Phys. Chem. A* **2003**, *107*, 9898–9902.

[21] L. Ismer, J. Ireta, S. Boeck, J. Neugebauer, *Phys. Rev. E* **2005**, *71*, 031911 (5pages).

[22] T.-H. Tang, E. Deretey, S. J. K. Jensen, I. G. Csizmadia, *Eur. Phys. J. D* **2006**, *37*, 217–222.

[23] R. Parthasarathi, S. S. Raman, V. Subramanian, T. Ramasami, *J. Phys. Chem. A* **2007**, *111*, 7141–7148.

[24] H. J. Bohr, K. J. Jalkanen, M. Elstner, K. Frimand, S. Suhai, *Chem. Phys.* **1999**, *246*, 13–36.

[25] M. S. de Vries, P. Hobza, *Annu. Rev. Phys. Chem.* **2007**, *58*, 585–612, and references therein.

[26] N. S. Myshakina, S. A. Asher, *J. Phys. Chem. B* **2007**, *111*, 4271–4279.

[27] N. G. Mirkin, S. Krimm, *J. Phys. Chem. A* **2007**, *111*, 5300–5303.

[28] C. A. Rice, I. Dauster, M. Suhm, *J. Chem. Phys.* **2007**, *126*, 134213 (8 pages).

[29] R. Wieczorek, J. J. Dannenberg, *J. Phys. Chem. B* **2008**, *112*, 1320–1328.

[30] R. Viswanathan, J. J. Dannenberg, *J. Phys. Chem. B* **2008**, *112*, 5199–5208.

[31] A.A. Granovsky, *PC GAMESS, Version 7.0, http://classic.chem.msu.su/gran/gamess/index.html*

[32] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[33] R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Oxford University Press, New York, **1990**. 458 p.

[34] W. F. Bieger-Konig, R. F. W. Bader, T.-H. Tang, *J. Comput. Chem.* **1982**, *3*, 317–328.

[35] E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170–173.

[36] L. Sobczyk, S. J. Grabowski, T. M. Krygowski, *Chem. Rev.* **2005**, *105*, 3513–3560.

[37] M. V. Vener, A. N. Egorova, D. P. Fomin, V. G. Tsirelson, *Chem. Phys. Lett.* **2007**, *440*, 279–285.

[38] J. Rossmeisl, J. K. Nørskov, K. W. Jacobsen, *J. Am. Chem. Soc.* **2004**, *126*, 13140–13143.

[39] J. Hutter, M. Parrinello, A. Alavi, D. Marx, M. Tuckerman, W. Andreoni, A. Curioni, E. Fois, U. Röthlisberger, P. Giannozzi, T. Deutsch, D. Sebastiani, A. Laio, J. VandeVondele, A. Seitsonen, S. Billeter, *CPMD, 3.11.1, IBM Research Laboratory and MPI für Festkörperforschung, Stuttgart*, **1995–2007**.

[40] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[41] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[42] N. Troullier, J. L. Martins, *Phys. Rev. B* **1991**, *43*, 1993–2006.

[43] S. V. Churakov, B. Wunder, *Phys. Chem. Minerals* **2004**, *31*, 131–141.

[44] M. V. Vener, J. Sauer, *Phys. Chem. Chem. Phys.* **2005**, *7*, 258–263.

[45] P. Rodzievicz, S. M. Melikova, K. S. Rutkowski, F. Buda, *ChemPhysChem* **2005**, *6*, 1719–1724.

[46] A. Jezierska, J. J. Panek, A. Koll, J. Mavri, *J. Chem. Phys.* **2007**, *126*, 205101 (9 pages).

[47] S. Tosoni, C. Tuma, J. Sauer, B. Civalleri, P. Ugliengo, *J. Chem. Phys.* **2007**, *127*, 154102 (11pages).

[48] A. Jezierska, J. J. Panek, A. Koll, *ChemPhysChem* **2008**, *9*, 839–846.

[49] V. R. Saunders, R. Dovesi, C. Roetti, M. Causa, N. M. Harrison, R. Orlando, C. M. Zikovich-Wilson, *CRYSTAL 98 User's Manual*, Universita di Torino, Torino, **1998**. 198 p.

[50] C. Gatti, *TOPOND98 User's Manual*, CNR-CSRSRC, Milano, **1999**. 24 p.

[51] M. V. Vener, A. V. Manaev, A. N. Egorova, V. G. Tsirelson, *J. Phys. Chem. A* **2007**, *111*, 1155–1162.

[52] J. P. Merrick, D. Moran, L. Radom, *J. Phys. Chem. A* **2007**, *111*, 11683–11700.

[53] R. Flaig, T. Koritsanszky, D. Dittrich, S. Wagner, P. Luger, *J. Am. Chem. Soc.* **2002**, *124*, 3407–3417.

[54] Y. Grinie, M. Avignon, C. Gerribou-Lagrange, *J. Mol. Stuct.* **1975**, *24*, 293–307.

[55] P. Bour, T. A. Keiderling, *J. Mol. Struct. (Theochem)* **2004**, *675*, 95–105.

[56] W. Yu, Z. Lin, Z. Huang, *ChemPhysChem* **2006**, *7*, 828–830.

[57] S. Wojtulewski, S. J. Grabowski, *Chem. Phys.* **2005**, *309*, 183–188.

[58] J. M. Bakker, C. Plützer, I. Hünig, T. Häber, I. Compagnon, G. von Helden, G. Meijer, K. Kleinermanns, *ChemPhysChem* **2005**, *6*, 120–128.

[59] P. Hobza, Z. Havlas, *Theoretica Chimica Acta* **2002**, *108*, 325–334, and references therein.

[60] K. Hermansson, *J. Phys. Chem. A* **2002**, *106*, 4695–4702.

[61] R. Viswanathan, A. Asensio, J. J. Dannenberg, *J. Phys. Chem. A* **2004**, *108*, 9205–9212.

[62] A. Abo-Riziq, B. O. Crews, M. p., Callahan, L. Grace, M. S. de Vries, *Angew. Chem. Int. Ed.* **2006**, *45*, 5166–5169.

[63] J. Kubelka, T. A. Keiderling, *J. Am. Chem. Soc.* **2001**, *123*, 12048–12058.

[64] R. Wieczorek, J. J. Dannenberg, *J. Am. Chem. Soc.* **2004**, *126*, 14198–14205.

[65] A. L. Lehninger, *The Molecular Basis of Cell Structure and Function*, Worth Publ. Inc., New York, **1972**. 1119 p.

[66] S.-H. Lee, S. Krimm, *Biopolymers* **1998**, *46*, 283–317.

[67] H. Park, J. Yung, C. Seok, *J. Phys. Chem. B* **2008**, *112*, 1041–1048.