Phonon spectra and thermodynamic properties of the infinite polyalanine α helix: A density-functional-theory-based harmonic vibrational analysis

Lars Ismer,[1] Joel Ireta,[1] Sixten Boeck,[1,2] and Jörg Neugebauer[1,2]
[1]*Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany*
[2]*Theoretische Physik, Universität Paderborn, Warburger Straβe 100, 33098 Paderborn, Germany*
(Received 4 May 2004; revised manuscript received 10 January 2005; published 23 March 2005)

We have performed a density-functional theory harmonic vibrational analysis of the infinite polyalanine α helix. The calculated phonon dispersion spectrum shows excellent agreement to available experimental data, except for the high frequency hydrogen stretching modes which show characteristic shifts due to anharmonic effects. A major advantage compared to previously performed empirical force field studies is that long range effects such as electrostatic interaction and polarization are intrinsically taken into account for characterizing hydrogen bond formation in the helix. Our results indicate that these effects are crucial to accurately describe the low frequency acoustical branches and lead to a significantly better agreement with experiment for the specific heat in the low temperature range.

DOI: 10.1103/PhysRevE.71.031911 PACS number(s): 87.15.Aa, 87.15.La, 87.15.He

An accurate description of the vibrational properties of the secondary structure of proteins is crucial for a detailed understanding of its structure, thermodynamic stability and functionality. For example, the formation of a specific secondary structural element (such as, e.g., an α helix, a 3_{10} helix, a β sheet, or a γ turn) is directly reflected in characteristic changes (fingerprints) in the vibrational spectrum which can be experimentally detected by spectroscopic methods [e.g., infrared (IR) and Raman spectroscopy] [1]. Further, particularly the lower frequency vibrational branches strongly determine the thermodynamic stability (free energy) of the secondary structure and also crucially affect atomic rearrangement (protein folding) and conformational fluctuation processes of biological importance [2]. A key interaction for the stabilization of the secondary structure is intrachain hydrogen bonding [3]. It is now well established that strength and formation of hydrogen bonds are reflected both in the vibrational modes of the peptide-backbone (amide modes) [1] as well as in the low frequency rigid-peptide (acoustical) modes [4]. Thus, for getting an accurate theoretical description of the vibrational properties an adequate treatment of the intrachain hydrogen bonding in the secondary structure is essential.

A problem in describing hydrogen bonds is that due to their polar nature they may strongly interact with each other. Thus, embedding an isolated hydrogen bond in an array of hydrogen bonds, such as, e.,g., in a regular α helix, strongly affects its dipole moment and bond strength [5,6]. This may lead to significant cooperative phenomena—the strength and thus the impact on the vibrational spectrum of a hydrogen bond depends on the environment, i.e., whether it is located in a small finite piece of the secondary structure or in long regular (periodic) structures. Thus, realistic models for the secondary structure must take hydrogen bond cooperativity into account.

A model system to study the α–helical secondary structure is alanine. Alanine is known to have a high propensity to form α helices [7] and due to its simple side chain and the possibility to synthesize long regular helices it has been strongly investigated in the past [4,8,9]. The theoretical studies were based on an infinite regular structural model for the α helix and employed empirical force fields which were optimized to fit the experimental spectra. While this approach allowed a fundamental understanding of the vibrational dynamics of polypeptides, theoretical studies revealed significant differences between experimental data and force field results for the specific heat [4,8], indicating substantial errors in the low frequency vibrational branches. Thus, calculations which fully account for long-range hydrogen bond cooperativity effects are highly desirable.

Previous studies showed the high reliability and accuracy of density functional theory in the generalized gradient approximation (DFT-GGA) to describe hydrogen bond strength and geometry (see [10] and references therein). Recently DFT-GGA has been used to study selected vibrations such as the Amide branches of infinite β sheets [11] or vibrations of short finite helices [12,13]. However a study of the complete vibrational spectrum of the secondary structure of proteins that considers its full electronic structure and that would allow a fully ab initio thermodynamic treatment is missing.

We have therefore employed DFT-GGA in this work to calculate the full vibrational spectrum of an infinite polyalanine chain in α–helical conformation and compared the derived frequencies and thermodynamic data to experiment. As will be shown this approach provides an accurate description of the phonon spectra (both with respect to absolute position and dispersion) and a significantly improved description of the thermodynamic properties (specific heat).

The infinite regular polyalanine α helix has been described as a crystal with, in terms of cylindrical coordinates, one-dimensional lattice periodicity: each lattice point n contains as basis an L-alanine peptide unit, which consists of $p=10$ atoms at positions $u_{i,\alpha}^{\mathrm{cyl}}$, where i goes over the p atoms and α goes over the three cylindrical directions r, ϕ, and z. The lattice vector, which transforms one representative peptide unit into its nth nearest neighbor is given by

 ©2005 The American Physical Society

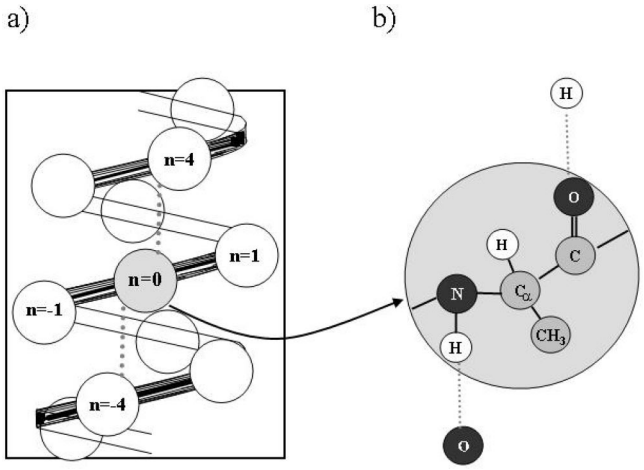

FIG. 1. (a) Schematic geometry of an infinite α helix as modeled in our supercell approach. The supercell contains 3 full turns of the helix, i.e., 11 peptide units. (b) Schematic geometry of a single alanine peptide unit. The dotted lines denote intrachain N-H···O hydrogen bonds.

$$\mathbf{R}_n^{\mathrm{cyl}} = (0, n\Theta, nh). \tag{1}$$

Here Θ is the helix twist and h is the helix pitch, i.e., the increment per peptide unit along the z axis. The helix axis is chosen to be parallel to the z axis.

The DFT calculations have been performed using a plane wave pseudopotential approach as implemented in the SFHIngX code [14]. The helix has been modeled using an orthorhombic supercell which contains 11 peptide units and three turns (Fig. 1). All calculations have been performed using the Perdew, Burke, and Ernzerhoff (PBE) GGA-functional [15] and a plane wave energy cutoff of 70 Ry. To obtain full self-consistency and completely stable convergence for the electronic system we have employed the discretized expression for the exchange-correlation functional [16]. The influence of spurious Hellman-Feynman forces hereby arising from the discretization has been reduced by using a twice as fine mesh than that imposed by the sampling theorem (details will be given elsewhere [17]).

Using the equilibrium geometry we have computed the (symmetry reduced) force constant matrix in terms of Cartesian coordinates:

$$K^{\mathrm{Cart}}_{(n,i,\alpha)(0,j,\alpha')} = \frac{\partial F_{(n,i,\alpha)}}{\partial u^{\mathrm{Cart}}_{(0,j,\alpha')}}. \tag{2}$$

Here, α and α' go over the three Cartesian directions x, y, and z. $F_{(n,i,\alpha)}$ is the αth component of the Hellman-Feynman force acting on the ith atom in the nth peptide and $u^{\mathrm{Cart}}_{0,j,\alpha'}$ is the displacement of the jth atom along α' in a representative peptide (with $n=0$). The derivatives in Eq. (2) have been calculated by a two-point symmetric finite difference approach. Thus, a total of 60 inequivalent structures has been considered. A value of 0.01 Bohr for the displacement has been found to give converged results of the finite difference quotients.

After converting the force constant matrix $\mathbf{K}^{\mathrm{Cart}}$ to the respective matrix $\mathbf{K}^{\mathrm{cyl}}$ in cylindrical coordinates a phase dependent quadratic $3p$ dimensional force constant matrix is obtained by convolution:

$$\mathbf{K}^{\mathrm{cyl}}(\varphi) = \sum_{n=-L}^{n=+L} \mathbf{K}^{\mathrm{cyl}}_{0,n} \exp(in\varphi). \tag{3}$$

The fact that the helix exhibits a translational symmetry in cylindrical coordinates allows us to use the same methodology as for phonons in crystals [18,19]. Specifically, the phase vector φ is analog to the k vector in periodic crystals and denotes the vibrational phase difference between adjacent peptide units. The upper and lower bound L for the summation in Eq. (3) are a cutoff and limit intrachain vibrational coupling up to the Lth nearest neighbor. Since our supercell contains 11 peptides, the maximum nearest neighbor interaction which can be resolved is $L=5$.

To obtain the vibrational frequencies and eigenstates of the helix the matrices $\mathbf{K}^{\mathrm{cyl}}(\varphi)$ have been transformed to mass-weighted Cartesian coordinates yielding the dynamical matrices:

$$\mathbf{D}(\varphi) = \mathbf{M}^{-1/2}\mathbf{B}^{\dagger}\mathbf{K}^{\mathrm{cyl}}(\varphi)\mathbf{B}\mathbf{M}^{-1/2}. \tag{4}$$

Here the matrix $\mathbf{M}^{-1/2}$ is a diagonal matrix, containing the square roots of the atomic masses of the peptide unit, and the $\mathbf{B}$ matrix [19,20] is constructed from the first-order derivatives of the cylindrical with respect to Cartesian coordinates. The equations of motion can then be expressed by the secular equation:

$$\det[\mathbf{D}(\varphi) - \omega_i^2(\varphi)\mathbf{1}] = 0. \tag{5}$$

Solving this equation for a given phase angle φ yields the according vibrational frequencies $\omega_i(\varphi)$. We note that Eqs. (3)–(5) are exact only on a special set of phase angles compatible with the periodic boundary conditions introduced by the supercell approach:

$$\varphi_n = \frac{2\pi n}{2L+1}, \quad -L \leqslant n \leqslant L. \tag{6}$$

To obtain the vibrational frequencies in between these points we did not use Eq. (5) (because it gives the wrong long wavelength limit for the acoustical branches) but a cubic spline interpolation scheme. As result we have obtained 30 continuous vibrational branches, the so-called phonon dispersion relation. Note, that due to symmetry considerations $\omega(\varphi)=\omega(-\varphi)$ and $\omega(\varphi)=\omega(\varphi+2\pi)$. Thus it is sufficient to display the phonon dispersion relation for $0 \leqslant \varphi \leqslant \pi$, i.e., in the first half of the vibrational Brillouin zone.

All 30 vibrational branches $\omega_i(\varphi)$ obtained are real, verifying that the helix is in a local minimum of the total energy surface. Careful tests regarding the numeric accuracy of our results give an error bar of 5 cm^{-1}. The eigenvectors have been analyzed—a qualitative classification of the individual branches is given in Table I. A more detailed discussion of the eigenvectors will be given elsewhere [17].

The calculated frequencies $\omega_i(\varphi)$ are shown in Fig. 2 together with experimental data points from polarized IR and

TABLE I. Qualitative classification of the vibrational branches based on an analysis of the corresponding eigenvectors. Notation: *s*=stretching, *b*=bending, *t*=torsional, *r*=rotational, *tr*=translational. Amide nomenclature according to Ref. [21].

Id	Type	Name
1-2	acoustical branches	
3-9	backbone vs side group *r* or *tr*	
10	backbone *t*, N-H *b*	Amide 5
11-13, 17	backbone *b* or *t*	
14-16	backbone or C_α–C_β *s*, C_α–*H b*	
18, 19	backbone *s*, N-H *b*	Amide 3
20-23	C_β-H *b*	
24	N-C *s*, N-H *b*	Amide 2
25	C-O *s*	Amide 1
26-29	C_β-H *s*, C_α-H *s*	Amide B
30	N-H *s*	Amide A

Raman spectra [9] on oriented films of α–helical polyalanine. Note that since only vibrational modes at high-symmetry points are optically active, experimental data are restricted to a few points in the Brillouin zone: In case of the α helix the *A* point at $\varphi=0^\circ$ and the E_1 point at $\varphi=99.57^\circ$ are IR and Raman active. The E_2 point at $\varphi=160.86^\circ$ is only Raman active. To allow a direct comparison with experiment the calculated frequencies have been scaled by a uniform factor of 1.02. Comparing with experiment we have to differentiate between the 25 branches lower than 2000 cm^{-1} and the 5 branches above 2000 cm^{-1}. For the first group of branches we find an overall excellent agreement. The only noticeable disagreement is for branch 10 at 600 cm^{-1} which is upshifted by $\approx$50 cm^{-1} compared to experiment. For all other branches below 2000 cm^{-1} the deviations are smaller than 20 cm^{-1}. For the second group of branches, which are all related to hydrogen stretching modes, a characteristic upward (blueshift) of $\approx$100 cm^{-1} compared to experiment is found. This deviation is expected to be mainly due to anharmonic contributions which are known to give shifts in the range of 50–250 cm^{-1} for hydrogen stretching vibrations [22].

We note that the excellent agreement between theory and experiment is similar to what has been achieved in previous studies based on force field models [8,9]. However, in contrast to these methods which are optimized to fit the vibrational spectra at available experimental data points, the DFT calculations performed here are free of any experimental input parameters. Therefore, it is interesting to check the accuracy of the force field models in the region not accessible in experiment, i.e., away from the high-symmetry (optical active) points in the Brillouin zone. The part of the vibrational spectrum where deviations are most likely to occur are branches exhibiting large dispersion and/or a complex shape of the $\omega_i(\varphi)$ dependence.

A comparison of our DFT results to the force field models showed hereby significant differences for two classes of such branches. The first class is given by the vibrations most directly involved in the hydrogen bonds [Fig. 1(b)], i.e., the Amide A, 1, and 2 branches (Table I). The dispersion of these branches is characterized by long ranged vibrational couplings between peptides, which are third, fourth, or even fifth nearest neighbors as can be inferred from the off-diagonal elements in the dynamical matrix. Particularly the dispersive splitting of the Amide A branch where we observe a splitting of 32 cm^{-1} is completely absent in force field calculations [8,9].

The second type of highly dispersive branches are the acoustical modes. For the two lowest lying acoustic branches we find significant shifts with respect to the force field calculations [8,9]: Compared to the latter our data are blue shifted by up to 25 cm^{-1} [Fig. 1(b)], i.e., DFT predicts a "harder" helix compared to force field results. Most likely, this difference arises from the limited accuracy in describing the intrachain hydrogen bonding interactions and from neglecting nonlinear, cooperative effects within the force fields used in Refs. [8,9]. In fact, recent DFT-GGA studies showed, that hydrogen bond cooperativity strengthens individual hydrogen bonds by more than a factor of two [5]. We therefore conclude, that this effect hardens interpetide force constants and consequently blueshifts the acoustical branches. To

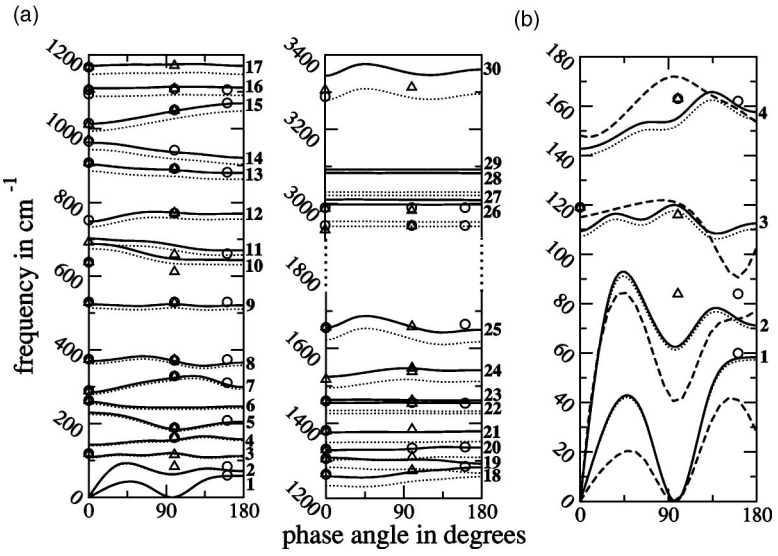

FIG. 2. (a) Calculated phonon dispersion relation (solid line for scaled frequencies; dotted line for unscaled frequencies) as function of the phase angle for polyalanine in α–helical conformation. Experimental data from polarized IR (triangles) and Raman (circle) measurements [9]. (b) Comparison for the 4 lowest vibrational branches to force field results (dashed line, from Fig. 3 in [9]).

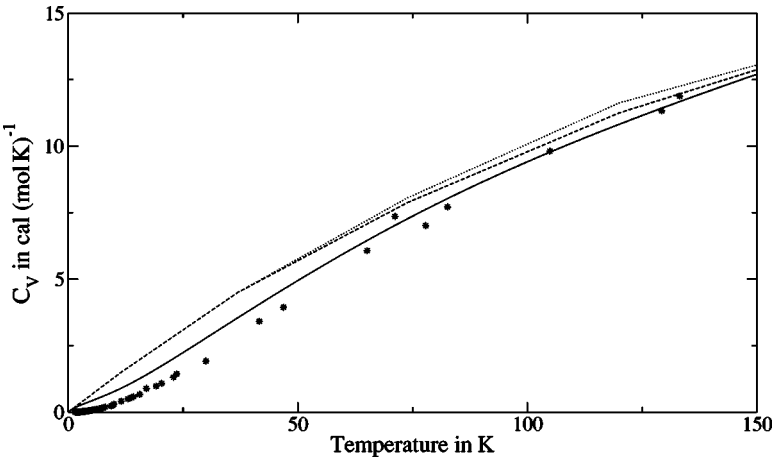

FIG. 3. Comparison of the specific heat of the polyalanine α helix: experimental results (circles [23]), force-field calculations of Fanconi *et al.* (dotted [4]) and of Datye *et al.* (dashed [8]), and our DFT-PBE results (solid line).

verify this explanation we have performed an identical vibrational study for an infinite polyalanine chain in the fully extended conformation, where hydrogen bonds and thus hydrogen bond cooperativity are completely absent. The acoustical branches of this conformation are found to be redshifted by 20–40 cm^{-1}, verifying that the absence of hydrogen bonds softens the back bone against mechanical deformations. A more detailed discussion of the above analysis together with decomposition into short and long-range interactions will be given elsewhere [17].

Since the low frequency branches have a fundamental meaning for the thermodynamic and mechanical properties of the helix, as they dominate the vibrational entropy and the low temperature heat capacity, it is important to check whether the deviations between force fields and DFT give rise to changes in these quantities. We have therefore calculated the heat capacity:

$$C_V = k_B \int_0^\infty g(\omega) \frac{\theta^2(\omega,T)\exp(-\theta(\omega,T))}{[1-\exp(-\theta(\omega,T))]^2} d\omega. \qquad (7)$$

Here, $g(\omega)$ denotes the density of states, as derived from the phonon dispersion relation, $\theta(\omega,T)=\hbar\omega_i(\varphi)/k_BT$, T is the temperature, k_B the Boltzmann constant, and $\hbar$ the Planck constant. Figure 3 shows the calculated heat capacity as function of temperature. Further included are experimental data [23] reported for crystalline α–helical polyalanine and results from force field calculations [4,8]. Note that we have calculated the specific heat at constant volume, whereas the experimental values are reported for constant pressure. However, the difference between these quantities is small for low temperatures [8] and has thus been neglected. For the following discussion we will restrict on temperatures below 150 K, as for larger temperatures the experimental values show a drastic increase in the specific heat, most likely due to residual water molecules in the experimental sample [23]. As can be seen in Fig. 3 the agreement with experiment is significantly improved for our DFT results in comparison to the force field results, i.e., the observed shift in the acoustical branches leads to significant changes in the heat capacity.

It is interesting to note that deviations are not only in value but exhibit also a qualitatively different behavior: The force field models predict in the low temperature region up to $\approx$50 K a linear behavior ($C_V \sim T$). Following the Debye law, this dependence is characteristic for a one-dimensional periodic crystal at low temperatures, whereas for two- or three-dimensional crystals a T^2/T^3 behavior is predicted. The deviation of experiment from the linear behavior has been thus interpreted by interhelix couplings [8]: as the experimental values have been obtained for crystalline polyalanine these lateral couplings were expected to become important at low temperatures. However, our results clearly demonstrate that the deviation from linearity is not due to interhelix couplings but an intrinsic property of the isolated α helix and results from the complex dispersion of the low frequency branches.

A discussion of the remaining discrepancy between experimental values for the heat capacity and our results remain speculative. Possible sources present in experiment but not included in a harmonic DFT-GGA study are on the one hand van der Waals attractions, which could lead to a hardening of the low frequency force constants and on the other hand positive (hardening) anharmonicity, which is expected to occur for large amplitudes in the acoustical branches. Further interhelix interactions in the experimental crystal (see above) may occur. In conclusion, by employing density-functional theory a significant improvement has been achieved in predicting the acoustical branches—which describe the mechanical or elastic properties—and crucially influence folding and structure formation of proteins. Since the resolution of the phonon-dispersion relation of our approach exceeds experimental data (where only the high symmetry points A, E_1, and E_2 are accessible) the calculated data may be used as benchmark to construct a new generation of improved force fields, which correctly treat long ranged peptide-peptide interactions and take hydrogen bond cooperativity accurately into account.

[1] A. Elliot and E. Ambrose, Nature (London) **165**, 921 (1950).
[2] C. L. Brooks III, M. Karplus, and B. Pettitt, *Advances in Chemical Physics* (Wiley, New York, 1988), Vol. LXXI.
[3] L. Pauling and R. Corey, Proc. R. Soc. London, Ser. B **141**, 10 (1953).
[4] B. Fanconi, W. E. Small, and W. L. Peticolas, Biopolymers **10**, 1277 (1971).
[5] J. Ireta *et al.*, J. Phys. Chem. B **107**, 1432 (2003).
[6] R. Wieczorek and J. Dannenberg, J. Am. Chem. Soc. **125**, 14065 (2003).
[7] C. N. Page and J. M. Scholtz,Biophys. J. **75**, 422 (1998).
[8] V. K. Datye and S. Krimm, J. Chem. Phys. **84**, 12 (1986).
[9] S. Lee and S. Krimm, Biopolymers **46**, 283 (1998).
[10] J. Ireta, J. Neugebauer, and M. Scheffler, J. Phys. Chem. A **108**, 5692 (2004).
[11] S. Franzen, J. Phys. Chem. A **107**, 9898 (2003).
[12] A. Moran and S. Mukamel, Proc. Natl. Acad. Sci. U.S.A. **101**, 506 (2004).
[13] C. J. Wang *et al.*, J. Phys. Chem. A **108**, 8160 (2004).
[14] S. Boeck *et al.* (in preparation); http://www.sfhingx.de.
[15] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[16] J. A. White and D. M. Bird, Phys. Rev. B **50**, R4954 (1994).
[17] L. Ismer, J. Ireta, and J. Neugebauer (unpublished).
[18] P. Higgs, Proc. R. Soc. London, Ser. A **133**, 472 (1953).
[19] S. Lee and S. Krimm, Chem. Phys. **230**, 277 (1998).
[20] E. Wilson, J. Chem. Phys. **7**, 1047 (1939).
[21] S. Krimm and J. Bandekar, Adv. Protein Chem. **38**, 181 (1986).
[22] B. G. Johnson, P. Gill, and J. Pople, J. Chem. Phys. **98**, 5612 (1993).
[23] M. Daurel, P. Delhaes, and E. Dupart, Biopolymers **14**, 801 (1975).