JCTC Journal of Chemical Theory and Computation

Electro-Optical Parameters for Computation of Nonresonance Raman Scattering Intensities of Peptides

Vineet Gupta,[†] Konstantin S. Smirnov,*[,‡] Daniel Bougeard,[‡] and Poonam Tandon[†]

Department of Physics, Lucknow University, 226007 Lucknow, India, and LASIR, Université des Sciences et Technologies de Lille, CNRS, Bât. C5, 59655 Villeneuve d'Ascq, France

Received November 26, 2008

Abstract: A set of electro-optical parameters (EOPs) of cylindrical zero-order bond polarizability model (BPM) for chemical bonds found in peptides was obtained from the results of quantum-chemical computations. The calculation of the polarizability tensors and the Raman scattering activities of four test molecules (N-methylacetamide, N,N-dimethylacetamide, dialanine, and diglycine) has shown that the BPM calculated quantities are in good agreement with the reference data obtained in experiments or derived by quantum-chemical calculations. The mean molecular polarizabilities are reproduced with the maximum relative error of 1.6%. A good agreement was obtained for the Raman activity of stretching vibrations, whereas a limited performance of the EOPs was found for the vibrational modes with a significant contribution of the bending internal coordinates involving H atoms. The origins of the discrepancies are analyzed, and the ways of improvement of the model performance are discussed.

1. Introduction

Response of electronic structure of molecules to an external electric field and the dependence of the response on a variation of molecular geometry play a fundamental role in reactivity and optical activity of molecules. For an electric field of low intensity the response is determined by the dipole polarizability tensor, whereas the change of the tensor with the geometry of the molecule is defined by the derivatives of the tensor with respect to displacements of nuclei. The latter quantities are directly related to the intensities in the Raman scattering spectra of the molecule. Raman spectroscopy has recommended itself as a fast, reliable, and nondestructive means of studying the structure sensitive dynamics of systems in the time scale from femtoseconds to picoseconds.

The calculation of the polarizability and the quantitative modeling of the Raman scattering intensities of peptides and proteins can be of significant help for understanding the relationship between structure, function, and dynamics of these (macro-)molecules.[1,2] Furthermore, such modeling can be of significant help for developing both structural and potential models of these systems. The simulation of spectral signatures of the vibrational dynamics is often limited to the computation of the density of vibrational states, that is a poor indicator of the validity of the models since for systems with many atoms the densities of vibrational states computed with different potential models strongly resemble each other because of the absence of any selection rules.[3] On the other hand, the infrared and Raman spectra in which the modes are discriminated according to selection rules often provide more valuable information, but examples of the IR and Raman spectra calculations are scarce, particularly because of the lack of intensity parameters.

The polarizability tensor as well as its derivatives with respect to atomic displacements can be computed in different ways. The modern quantum-chemical methods permit the evaluation of these quantities with a high precision, but such calculations are still fairly limited to systems of small size, especially if the derivatives of the polarizability tensor have to be computed. For real systems containing hundreds/thousands of atoms the only feasible way is to use parametric models that permit a fast calculation of the quantities of interest. Several such models exist[4−8] among which the bond

* Corresponding author phone: +33 320336139; fax: +33 320436755; e-mail: Konstantin.Smirnov@univlille1.fr.

[†] Lucknow University.

[‡] Université des Sciences et Technologies de Lille.

10.1021/ct800510y CCC: $40.75
Published on Web 04/08/2009

polarizability model[4] (BPM) permits a fast and reliable approach for computing both the polarizability and the Raman scattering intensities.

The BPM uses a set of parameters (electro-optical parameters, EOPs) to construct, via the topology of chemical bonds and the geometry of the system, the polarizability tensor and to describe the change of the polarizability with geometry variations. As in any parametric model, the quality of the parameters determines the reliability and predictive power of the model. Experimentally, EOPs can be extracted from data of Raman scattering experiments, but the resulting values of the parameters are often ambiguous because the number of observables is generally smaller than the number of EOPs to be determined. The ill-conditioned mathematical problem as well as the fact that measurements of the absolute Raman intensities are necessary to obtain absolute values of parameters has as a consequence the fact that sets of EOPs were determined only for few systems, e.g. alkanes,[9] while most of the works employing the BPM use only relative values of the parameters. It is therefore highly desirable to have a way of deriving reliable sets of EOPs whose values can be interpreted with a physical sense but not as solution of a mathematical problem.

Recently, an approach for obtaining EOPs of BPM from results of *ab initio* quantum-chemical calculations of molecular models was proposed.[10] The advantage of the procedure is that it overcomes the two problems mentioned above because the calculations are in principle capable of providing the necessary observables. The sets of EOPs derived in this way for systems with different types of chemical bonds (alkanes[10] and aluminosilicates[11]) were shown to be in good agreement with available experimental and theoretical data, although the EOPs for aluminosilicates with partially ionic bonds were found to have a limited transferability to periodic systems.[11]

The present paper reports the development of a set of the BPM electro-optical parameters for chemical bonds relevant to molecules of biological interest, in particular to peptides. The paper is organized as follows. The next section (section 2) discusses the BPM and the procedure used to derive the EOPs. This discussion is followed by a presentation of the molecular models of the training and validation sets and by details of quantum-chemical calculations whose results are used as reference data in the development. Results of the development and the analysis of discrepancies and of limitations of the model are given in section 3, which also presents the application of the obtained set of EOPs to the calculation of the polarizability and of the Raman spectra of molecules in the validation set. The conclusions of the work are summarized in section 4.

2. Model and Computational Procedures

2.1. Bond Polarizability Model. This section provides a brief description of BPM; the interested reader can find a thorough presentation of BPM in refs 9, 12, and 13. The BPM represents the polarizability tensor **A** of a molecule as the sum of polarizability tensors of the bonds in the molecule

$$\mathbf{A} = \sum_{i \in \text{bonds}} \mathbf{a}_i \tag{1}$$

with $\mathbf{a}_i$ being the polarizability tensor of bond i in the fixed Cartesian frame. The tensors a_i can be rewritten as

$$\boldsymbol{\alpha}_i = \mathbf{R}_i \alpha_i \mathbf{R}_i^{-1} \tag{2}$$

where $\boldsymbol{\alpha}_i$ is the bond polarizability tensor in its principal axes, and $\mathbf{R}_i$ is the rotation matrix between the bond principal and the Cartesian frames. The three nonzero components of the α_i tensor are the longitudinal (L) and two transversal (T and T') bond polarizabilities. The polarizability tensor at the equilibrium $\mathbf{A}^0$ is therefore equal to

$$\mathbf{A}^0 = \sum_{i \in \text{bonds}} \mathbf{R}_i \boldsymbol{\alpha}_i^0 \mathbf{R}_i^{-1} \tag{3}$$

and the Cartesian components of the $\mathbf{A}^0$ tensor can be written as

$$A_{pq}^0 = \sum_d \sum_b \alpha_d^0(b) \sum_{i \in b} p_d(i) q_d(i) \tag{4}$$

where $p, q = x, y, z$, while the indexes d and b run over the principal components ($d = L, T, T'$) of the bond polarizability tensor $\boldsymbol{\alpha}_i$ and bond types, respectively.

The intensity I_k of the peak corresponding to the vibrational mode k in the Raman spectrum of molecule reads[14]

$$I_k \propto \frac{\omega_s^4}{\omega_k(1 - \exp(-\beta\hbar\omega_k))} \left\{ \frac{45(\bar{\alpha}')^2 + 7(\gamma')^2}{45} \right\} \tag{5}$$

where ω_k and ω_s are the angular frequencies of the mode k and of the scattered radiation, respectively, and β and $\hbar$ have their usual meaning. The quantities $\bar{\alpha}'$ and γ' in eq 5 stand for the trace and anisotropy of the polarizability tensor derivative $\partial\mathbf{A}/\partial Q_k$ with respect to the normal coordinate Q of mode k; these are calculated in the usual way[14]

$$\bar{\alpha}' = \frac{1}{3}(A'_{xx} + A'_{yy} + A'_{zz}) \tag{6}$$

$$(\gamma')^2 = \frac{1}{2}[(A'_{xx} - A'_{yy})^2 + (A'_{yy} - A'_{zz})^2 + (A'_{zz} - A'_{xx})^2 + 6(A'^2_{xy} + A'^2_{xz} + A'^2_{yz})] \tag{7}$$

with $A'_{pq} \equiv \partial A_{pq}/\partial Q_k$. Note that eq 5 was obtained assuming a linearly polarized incident radiation.

The Raman scattering intensity I_k of the mode k is therefore determined by the tensor $\mathbf{A}'_k$ that can be written as

$$\mathbf{A}'_k \equiv \frac{\partial \mathbf{A}}{\partial Q_k} = \sum_{i \in \text{bonds}} \left\{ \frac{\partial \mathbf{R}_i}{\partial Q_k} \boldsymbol{\alpha}_i^0 \mathbf{R}_i^{-1} + \mathbf{R}_i \boldsymbol{\alpha}_i^0 \frac{\partial \mathbf{R}_i^{-1}}{\partial Q_k} + \mathbf{R}_i \left[\sum_j \boldsymbol{\alpha}'_{ij} \left(\frac{\partial s_j}{\partial Q_k} \right) \right] \mathbf{R}_i^{-1} \right\} \tag{8}$$

where the index j runs over the internal coordinates s_j. The first two terms in eq 8 account for the variation of the molecular polarizability with changes of bond orientations in the normal mode Q_k, while the last term explains the change of the bond polarizability with the variation of internal coordinates. The $\boldsymbol{\alpha}_i^0$ and $\boldsymbol{\alpha}'_{ij} \equiv (\boldsymbol{\alpha}_i/\partial s_j)_0$ quantities in eq 8 are known as equilibrium and valence electro-optical parameters of the bond polarizability model, respectively.

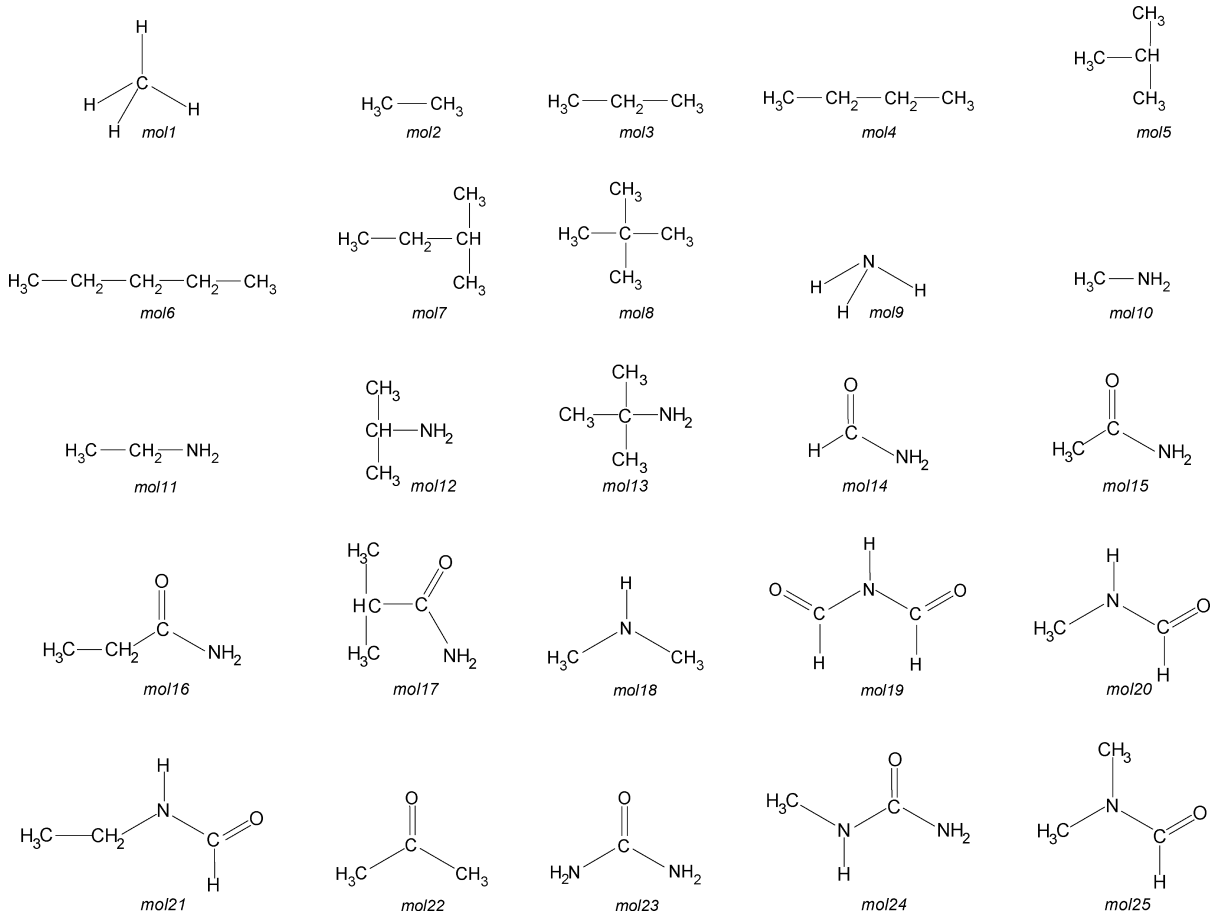

Figure 1. Molecules of the training set used to derive electro-optical parameters.

In a similar way as eq 4 the components of the $\mathbf{A}'_k$ tensor can be represented as a linear combination of the EOPs with coefficients depending on the molecular geometry (matrices $\mathbf{R}_i$) and on the variations of the internal coordinates in the normal modes (derivatives $\partial s_j/\partial Q_k$). All these quantities can be obtained from the molecular geometry and from the Cartesian displacements of atoms in the normal mode k. Hence, for a molecule containing N atoms the eqs 3 and 8 form a system of $6 \times (3N - 5)$ linear equations with EOPs as unknowns.

In the following we shall limit ourselves to a cylindrical zero-order BPM which implies that (i) all but $\boldsymbol{\alpha}'_{ii}$ parameters are equal to zero (zero-order model) and that (ii) in the tensor $\boldsymbol{\alpha}_i^0$ and in the tensor $\boldsymbol{\alpha}'_{ii}$ the two transversal components are equal to each other (cylindrical model). The former assumption reduces eq 8 to

$$\mathbf{A}'_k = \sum_{i \in \text{bonds}} \left\{ \frac{\partial \mathbf{R}_i}{\partial Q_k} \boldsymbol{\alpha}_i^0 \mathbf{R}_i^{-1} + \mathbf{R}_i \boldsymbol{\alpha}_i^0 \frac{\partial \mathbf{R}_i^{-1}}{\partial Q_k} + \mathbf{R}_i \left[\boldsymbol{\alpha}'_{ii} \left(\frac{\partial s_i}{\partial Q_k} \right) \right] \mathbf{R}_i^{-1} \right\} \tag{9}$$

while the latter limits the number of EOPs to four per bond type. A model that retains terms with $j \neq i$ in the last sum of eq 8 is known as the first-order BPM.

To derive EOPs with the approach described above one needs a database of the Cartesian components of the $\mathbf{A}^0$ and $\mathbf{A}'_k$ tensors for a set of molecules with known geometry. As such information is not generally available from experimental data, the database was generated with the use of quantum-chemical calculations as described in the next section. The $6 \times (3N - 5)$ linear equations for each molecule are then combined to form an extended set of equations that can be solved by a linear least-squares (LLSQ) method yielding statistically meaningful values of sought parameters. The proposed approach differs from that used previously in refs 10 and 11 in the way of creating the database. Instead of calculating the polarizability tensor derivatives with respect to bond lengths of individual bonds by the finite-differences procedure,[10,11] the present approach uses the derivatives of the tensor in normal modes obtained in an analytical way. As the derivatives determine the Raman scattering activity of the vibrational modes, it is therefore expected that the resulting EOPs will provide reliable Raman scattering intensities, while keeping a high precision in the calculation of the equilibrium polarizability.

2.2. Training Set. The training set of molecules used to obtain the A^0_{pq} and $\partial A_{pq}/\partial Q_k$ quantities consisted of eight linear and branched alkanes and of 17 small organic molecules with C−H, C−C, N−H, C−N, and C═O bonds (Figure 1). Previously, it was found that values of EOPs can be influenced by both the chemical environment of the bond and the position and orientation of the bond in the molecule.[15,16] Nevertheless, to keep the number of parameters in reasonable limits, it was supposed that all the bonds of the same type can be described by the same set of EOPs. The only exception was done for the C−N bonds which were

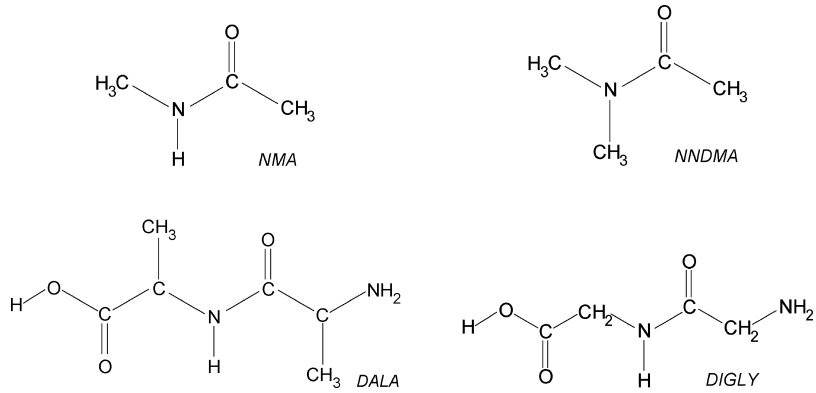

Figure 2. Molecules used in the validation set: N-methylacetamide (NMA), N,N-dimethylacetamide (NNDMA), dialanine (DALA), and diglycine (DIGLY).

separated into two types corresponding to C−N single and C−N partially double (peptide) bonds with bond lengths of *ca.* 1.45 Å and 1.36 Å, respectively. Therefore, six different types of bonds with a total number of 24 EOPs parameters were considered in the model.

For each molecule of the training set, the molecular geometry was optimized without symmetry constraints, and then the analysis of the vibrational normal modes was performed. The Cartesian displacements of atoms in the modes and the matrix of the Cartesian derivatives of the polarizability tensor produced in the calculations were then used to compute the $\partial A_{pq}/\partial Q_k$ quantities. The quantum-chemical calculations were carried out at the B3LYP level with the use of the pVTZ Sadlej basis set that was specially designed for the calculation of molecular electric properties, especially polarizability.[17,18] Halls and Schlegel have also shown that this basis set provides very good Raman activities for vibrational modes of small molecules.[19] The quantum-chemical calculations were performed with the G03 program.[20] The linear least-squares fitting of EOPs was done using the singular value decomposition method.[21]

The optimized geometry of the molecules of the training set is available as Supporting Information.

2.3. Validation Set. The derived set of EOPs was tested in the calculation of the polarizability tensors and Raman activities of N-methylacetamide (NMA), N,N-dimethylacetamide (NNDMA), dialanine (DALA), and diglycine (DIGLY) molecules (Figure 2) which can be considered as elementary building blocks of many peptide structures. Taking into account the fact that the experimental vibrational spectra of molecules in a condensed phase can substantially be affected by intermolecular interactions as well as reveal anharmonic effects (overtones and combination bands, Fermi resonance), the reference Raman activities R of the vibrational modes of the molecules were obtained in G03 calculations at the B3LYP/Sadlej basis set level after complete optimization of molecular geometries. These quantities were compared with those computed using the cylindrical zero-order BPM with obtained EOPs. In agreement with the definition used in the G03 program for the Raman activity the R_k quantities were calculated as

$$R_k = 45(\bar{\alpha}'_k)^2 + 7(\gamma'_k)^2 \qquad (10)$$

with $\bar{\alpha}'_k$ and γ'_k defined in eqs 6 and 7. One easily sees the relation of the $R_k$quantity with the intensity I_k in the Raman spectrum (eq 5). As the set of EOPs developed in the present work does not include C−O and O−H bonds, the calculations for the DALA and DIGLY molecules were done with the EOPs for C−O bond equal to those of C−N one, while the EOPs for the O−H bond were taken from ref 11.

3. Results and Discussion

Figure 3a,b presents the correlation graphs for the Cartesian components of the polarizability tensor A^0_{pq} and of the derivatives of the polarizability tensor A'_{pq} in the normal modes obtained in the quantum-chemical calculations and with the BPM for the molecules in the training set. The polarizability tensor components are, however, not usually measured quantities, and the response of a molecule to an applied electric field is commonly characterized by the mean molecular polarizability $\bar{\alpha} = (A_{xx} + A_{yy} + A_{zz})/3$. The BPM reproduces the mean polarizabilities of the 25 molecules of the training set with an rms error of 0.054 Å^3 and with the maximum relative error of 2.1% for the mean polarizability of NH_3 molecule. The good correlation seen in Figure 3a as well as the small discrepancy between the BPM computed and reference values of $\bar{\alpha}$ indicate that the equilibrium polarizability can be well described with the set of EOPs derived using the proposed approach. The good overall agreement seen in Figure 3b between the DFT and BPM computed A'_{pq} derivatives also implies that the set of EOPs obtained in the fitting procedure should yield correct Raman activities.

Figure 4a shows a correlation graph for the R_k values for the normal modes of the molecules in the training set. Although the accord between the DFT and BPM computed quantities can be considered as rather satisfactory, the inset in the figure reveals significant errors for some vibrational modes with a low activity. To identify the source of errors Figure 4b,c displays the correlation graphs for the $\bar{\alpha}'_k$ and γ'_k quantities, respectively, which determine the activity. One sees that a much better correlation exists for the trace of the polarizability tensor derivative than for the tensor's anisotropy.

The different extent of correlation in Figure 4b,c can in part be explained by different rms errors for the $\bar{\alpha}'_k$ and γ'_k

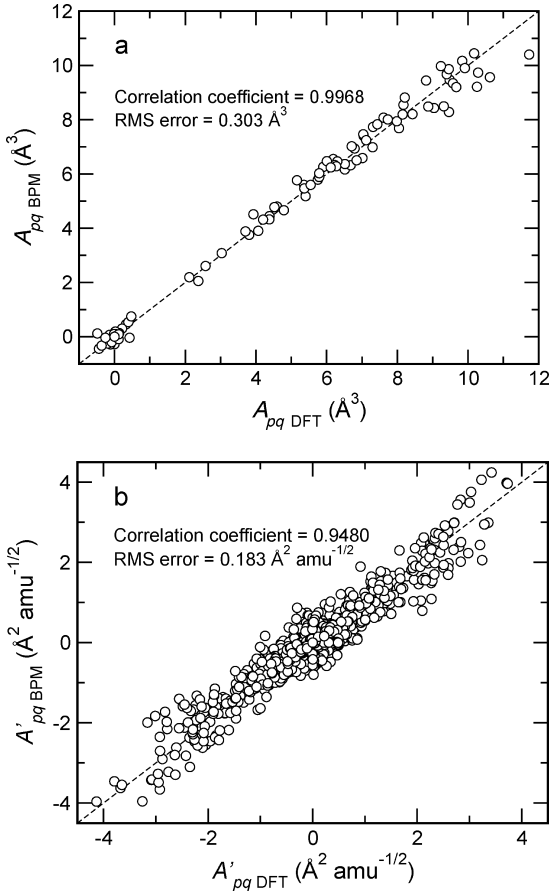

Figure 3. Correlation between the Cartesian components A_{pq} of the polarizability tensor (a) and between the components of the polarizability tensor derivatives $\partial A_{pq}/\partial Q_k$ (b) of molecules in the training set calculated by DFT and with BPM. Diagonal dashed line is the identity line.

derivatives. Assuming the Gaussian propagation rule, the variances σ^2 of these quantities can be estimated as

$$\sigma^2_{\bar{\alpha}'} \approx \frac{1}{3}\sigma^2_{A'} \tag{11}$$

$$\sigma^2_{\gamma'} \approx \left[1 + \frac{6(A'^{\,2}_{xy} + A'^{\,2}_{xz} + A'^{\,2}_{yz})}{(\gamma')^2}\right]\sigma^2_{A'} \tag{12}$$

where $\sigma^2_{A'}$ is the variance of the components of the $\mathbf{A}'$ tensor. One immediately sees that the rms error of the $\bar{\alpha}'$ derivative is smaller than that of A'_{pq} ones; the value of the error estimated to 0.106 Å^2 from eq 11 is in good agreement with that found for the fitted values and reported in Figure 4b. On the other hand, the rms error of the anisotropy derivative is always larger than the error of A'_{pq} quantities, and it lies in the range $\sigma_{A'}$ to $\sqrt{3}\sigma_{A'}$; for some vibrational modes the error can be comparable to the γ' value itself. This feature can result in a large error for the activity of modes with a high depolarization ratio whose intensity is largely determined by the γ'_k quantity. The above consideration however does not explain the scatter of points in Figure 4c that exceeds the estimated error for the γ'_k quantity, especialy for the values close to the graph's origin. This fact points to a deficiency of the used BPM model for the computation of the polarizability anisotropy derivative. It is worthy to note

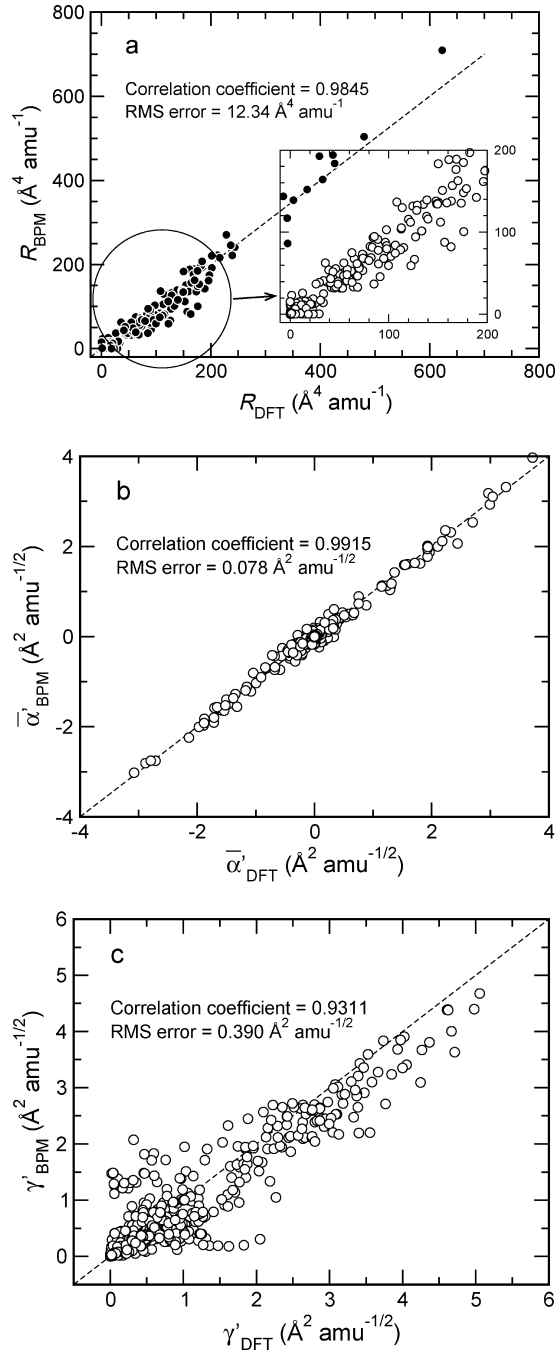

Figure 4. Correlation between the Raman activity (a), the derivative of mean polarizability $\bar{\alpha}'$ (b), and the anisotropy of the derivative of polarizability tensor γ' (c) calculated by DFT and with BPM for normal modes of molecules of the training set. Diagonal dashed line is the identity line.

that such a problem is inherent not only in the BPM but also in other models; this issue was already addressed in refs 22 and 23. An extension of the model to get around the deficiency is discussed below.

Table 1 reports the electro-optical parameters obtained in the LLSQ fit and compares them with those derived from the experimental Raman scattering intensities[24] or obtained in previous theoretical calculations.[10,25,26] A good agreement for parameters of different sets can only be found for the EOPs of C−H bond, while for other bonds the values of parameters vary significantly, particularly the longitudinal

Table 1. Values of Electro-Optical Parameters[f]

bond	equilibrium parameters (in Å^3) α_L^0	α_T^0	valence parameters (in Å^2) α'_L	α'_T	ref
C−H	*0.818 (0.092)*	*0.569 (0.087)*	*2.665 (0.090)*	*0.409 (0.061)*	
	0.779 (0.101)	0.489 (0.054)	2.743 (0.049)	0.353 (0.033)	10[e]
	0.82	0.50	2.14	0.44	24
	0.83	0.54	3.90	0.80	25
	0.83−0.88	0.53	2.43	0.45	26
N−H	*0.826 (0.141)*	*0.661 (0.126)*	*2.502 (0.196)*	*0.456 (0.133)*	
	1.50	0.50	1.0	0.24	24
C−C	*1.047 (0.302)*	*0.255 (0.273)*	*2.742 (0.493)*	*0.111 (0.349)*	
	1.677 (0.195)	0.127 (0.113)	2.863 (0.053)	0.243 (0.036)	10[e]
	0.7	0.50	3.0	0.5	24
	1.290−1.495	0.240−0.265	1.900	0.500−1.000	26
C−N$_s$[a]	*1.274 (0.389)*	*0.322 (0.328)*	*1.371 (0.810)*	*0.037 (0.561)*	
	0.49	0.44[d]	2.5	0.5	24
C−N$_p$[b]	*1.611 (0.444)*	*0.000*[c] *(0.393)*	*1.906 (0.866)*	*0.534 (0.597)*	
	0.28	0.49	2.0	0.50	24
C═O	*2.341 (0.402)*	*1.202 (0.352)*	*4.445 (0.824)*	*0.469 (0.576)*	
	2.56	1.04[d]	8.3	0.75[d]	24
	1.61	0.80	3.76	0.94	25, 26

[a] CN simple bond. [b] CN peptide bond. [c] The LLSQ fit resulted in value of α_T^0 parameter equal to −0.011 Å^3; the parameter was set to zero in the table. [d] Mean value of two transversal components. [e] See also Table 1 in ref 10 for a compilation of experimental values. [f] Statistical uncertainties are given in parentheses. Values given in italics refer to the results of the present work.

ones. There can be several reasons for such a discrepancy. First, it may reflect the ambiguity of EOPs derived on the basis of experimental Raman intensities. Analysis of Table 1 reveals that some parameters found in the literature have rather unexpected values thus indicating that the set may be one of several possible mathematical solutions of the problem. For example, one can expect values of N−H bond parameters in the range of those for C−H and O−H bonds, that is not the case for the EOPs obtained in ref 24. In contrast, the EOPs derived in the present work follow the expected trends. The present approach can therefore be of substantial help in choosing the solution that is meaningful from the physical viewpoint. Second, some differences in the parameters can be due to the use of a first-order BPM in refs 24−26 The use of the first-order BPM is, however, expected to influence the values of the principal parameters to a small extent.

The analysis of the standard errors of the EOPs in Table 1 shows that some parameters have a large uncertainty. The reason for this is 2-fold. First, not all chemical bonds are equally represented in the database; this can cause a larger uncertainty of EOPs for less frequently sampled bonds. Thus, the 25 molecules of the training set contain 148 C−H bonds as compared to 13 C−N$_p$ bonds; note that the C−H bond EOPs have the smallest errors. Second, large uncertainty can point to a variation of the parameter while going from one molecule to another and in normal modes. Thus, Table 1 shows that valence EOPs of C−H and C−C bonds obtained in ref 10 generally have smaller uncertainties. These parameters were derived from the variations of molecular polarizability upon change of the lengths of *individual bonds* in molecules of training set, whereas the values of present EOPs were derived from the variation of the **A** tensor in *normal modes*. Large uncertainty of parameter can therefore be indicative of an important influence of other internal coordinates on the component of the bond polarizability tensor. The set of parameters reported in Table 1 can thus be considered as a generic set that may need further refinement for a particular class of molecules (at the expense of generality).

The set of EOPs was applied to the calculation of the polarizability and of the Raman scattering activity of the molecules of the validation set (Figure 2). Values of the components A_{pq}^0 of the polarizability tensor of the molecules are reproduced with an rms error of 0.869 Å^3, and a mean absolute error in the $\bar{\alpha}$ quantity of the four molecules is equal to 0.116 Å^3 with the maximum relative error of 1.6%. Note that the maximum discrepancies in the mean polarizability were obtained for the dialanine and diglycine molecules for which the values of EOPs for the C−O and O−H bonds of the carboxyl group were taken in a somewhat arbitrary way. The Raman activities R_k are reproduced with an rms error of 14.081 Å^4 amu^{-1} that is close to the rms for this quantity for the vibrational modes of the molecules in the training set.

Figures 5−8 compare the calculated Raman spectra of the molecules in the validation set. The spectra were obtained by the convolution of spectrum of Raman activities R_k with a Lorentzian function of 4 cm^{-1} width. One sees that intensities in the high-frequency region of reference spectra are well reproduced with the BPM. Given that the intensities are entirely due to C−H and N−H stretching vibrations, the good agreement between the BPM calculated and reference DFT spectra points to the fact that the valence EOPs for these bonds have reliable values. The same analysis concerns the mode at 1710−1730 cm^{-1} due to the stretching of the C═O bond (99% of the coordinate in potential energy distribution). The Raman activity of this mode is reproduced by the BPM model with a precision better than 10% thus indicating correct values of the valence parameters.

The most important differences between the spectra in the midfrequency range occur between 1200 and 1500 cm^{-1}; the BPM underestimates the peaks intensities in this region as compared to the DFT results. To our mind this fact points to the limits of the zero-order BPM. The vibrational modes in this energy range involve variation of many angle-bending

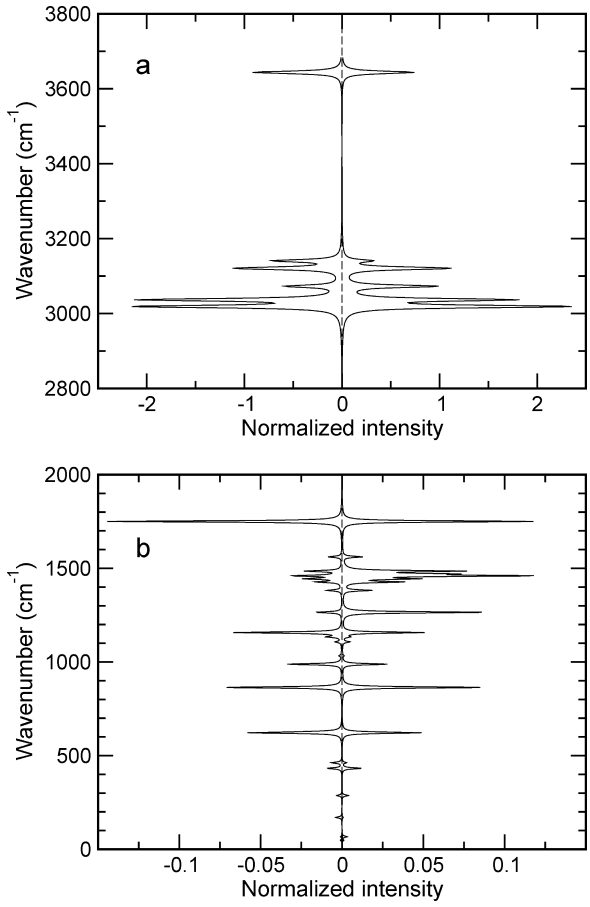

Figure 5. High-frequency (a) and low-frequency (b) parts of the Raman activity spectrum of N-methylacetamide molecule calculated with DFT (positive values of *x*-axis) and with BPM (negative values of *x*-axis). The intensity in the spectra were normalized on the total intensity and multiplied by 100; the intensities calculated with the BPM were made negative for comparison.

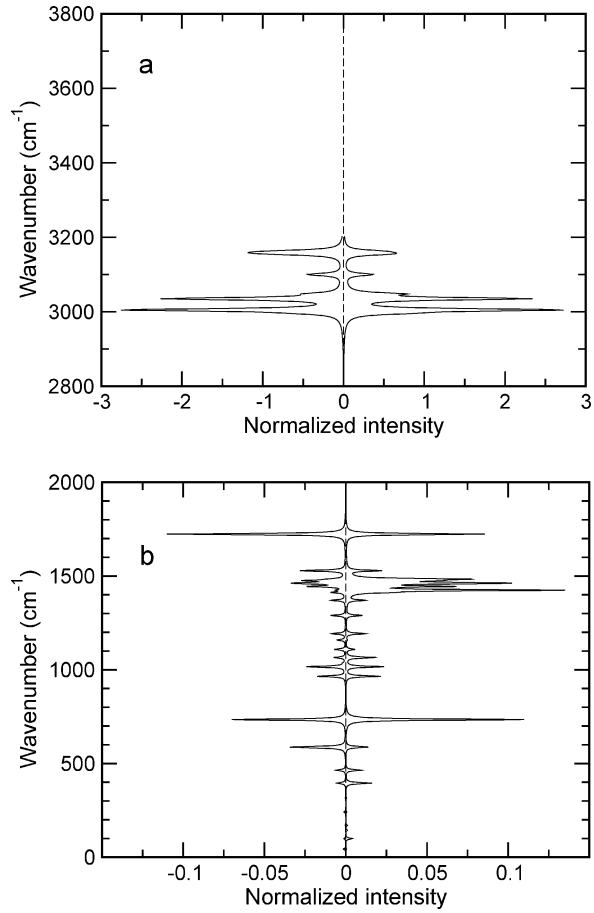

Figure 6. High-frequency (a) and low-frequency (b) parts of the Raman activity spectrum of N,N-dimethylacetamide molecule calculated with DFT (positive values of *x*-axis) and with BPM (negative values of *x*-axis). The intensity in the spectra were normalized on the total intensity and multiplied by 100; the intensities calculated with the BPM were made negative for comparison.

internal coordinates, particularly those including H-atoms. The zero-order BPM attempts to describe the related change in the molecular polarizability solely by contributions due to changes in bonds orientations, while it lacks the change induced in the bond polarizabilities upon variation of the bending coordinates. The missing polarizability variation could be simply described by a change of the equilibrium EOPs. However, such a modification will alter components of the equilibrium polarizability tensor, and, therefore, the improvement in the Raman scattering intensities will occur at a sacrifice of quality of the A^0_{pq} quantities.

The limited performance of BPM in describing the Raman intensity of bending modes has already been mentioned by Van Hemert and Blom.[23] In agreement with the exprimental work of Montero and Bermejo,[27] the authors of ref 23 pointed to the need of explicit consideration of parameters depending on the bending coordinates. As it was mentioned above, indirect support in the favor of such an extension is provided by large statistical uncertainties of some EOPs derived in the present work (Table 1). The extension would necessitate to keep in eq 8 α'_{ij} terms for some angle-bending internal coordinates s_j, that will, however, notably increase the number of EOPs.

Figure 9 compares the experimental Raman spectra of the N-methylacetamide and N,N-dimethylacetamide molecules in the gas phase[28] with those computed by eq 5 using the characteristics of experimental setup. To mimic the rotational broadening of Raman peaks in the experimental spectra, the theoretical ones were convoluted with a Lorentzian function with HWHM = 20 cm^{-1}. Besides the aforementioned discrepancy in the region of 1200−1500 cm^{-1}, the figure displays a very good agreement between the experimental and calculated spectra once more indicating that the EOPs can be used for the quantitative modeling of both the Raman scattering intensity and polarizability of peptide structures.

As the purpose of the EOPs model is the computation of large systems, we have also calculated the Raman spectrum of polyglycine I (PGI) by normal-mode analysis. The vibrational dynamics of this system has been the subject of many computational studies,[29−35] but the analysis of the dynamics in those works was limited to the calculation of the frequencies of vibrational modes and their dispersion curves, classification of modes according to the symmetry, and to the analysis of the heat capacity. Neither of the studies reported the computation of the intensities in the infrared and/or Raman spectra of the system, and the refinements of

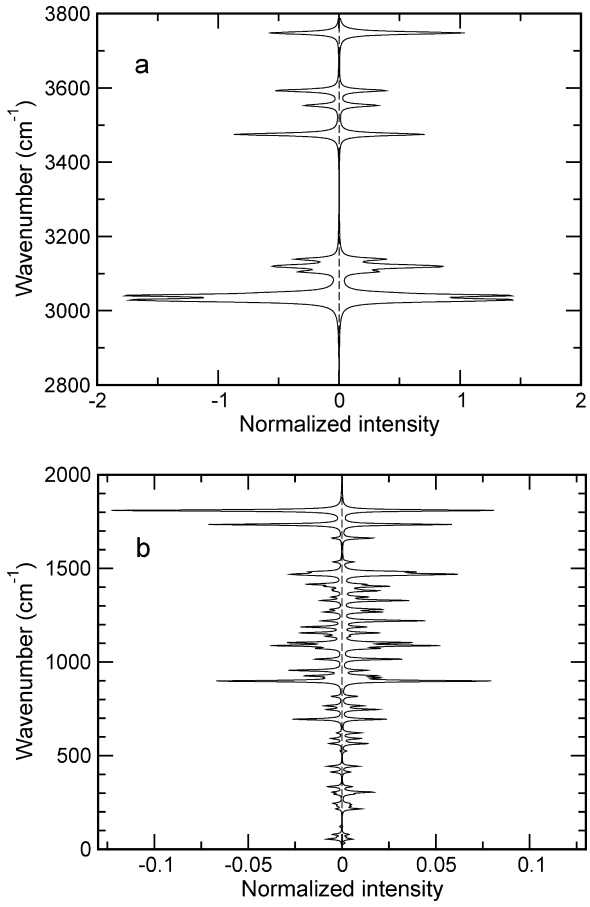

Figure 7. High-frequency (a) and low-frequency (b) parts of the Raman activity spectrum of dialanine molecule calculated with DFT (positive values of *x*-axis) and with BPM (negative values of *x*-axis). The intensity in the spectra were normalized on the total intensity and multiplied by 100; the intensities calculated with the BPM were made negative for comparison.

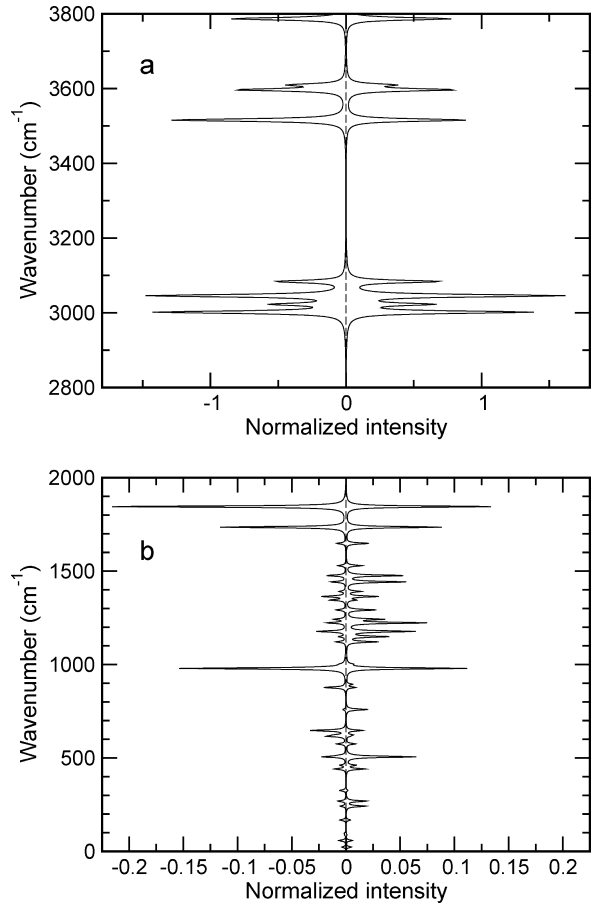

Figure 8. High-frequency (a) and low-frequency (b) parts of the Raman activity spectrum of diglycine molecule calculated with DFT (positive values of *x*-axis) and with BPM (negative values of *x*-axis). The intensity in the spectra were normalized on the total intensity and multiplied by 100; the intensities calculated with the BPM were made negative for comparison.

the force field parameters were therefore carried out solely on the basis of the comparison of vibrational frequencies. In order to evaluate the influence of the force field on the calculated intensities two force fields of a different nature were chosen: a valence force field by Dwivedi and Krimm[33] (100 internal coordinates, 70 force constants) and a Urey−Bradley force field (148 internal coordinates, 31 force constants) used by Porwal et al.[35] The system treated with the former force field was a three-dimensional periodic model of PGI structure, whereas the calculations using the latter force field considered a single infinite polyglycine chain. In both models the coordinates of atoms corresponded to the low molecular weight PGI structure determined by Kajava.[36]

The inspection of the calculated frequencies indicates a good reproduction of the data cited in the previous works and a good agreement with experimentally determined frequencies. However, the potential energy distribution significantly varies between the force fields although the frequencies and general form of the modes are very similar. This fact finds confirmation in Figure 10, which presents the computed Raman spectra in the region 2000−200 cm^{-1} and compares them with the experimental spectrum adapted from ref 37. All the spectra reveal the most intense Raman peaks at ca. 1000 cm^{-1}. In the spectral range from 2000 to 1000 cm^{-1} the relative intensities of the Amide I and Amide II peaks are well reproduced by the calculations with the former being more intense, whereas the intensity of the Amide III peak is underestimated. The weakness of the intensity model discussed above results in an underestimation of the Raman intensity of the peaks due to the angle deformation modes involving CH bonds. Below 1000 cm^{-1} an intense peak is present at 700−730 cm^{-1} in the calculated spectra, while it is absent in the experimental one; the relative intensities of other peaks vary but are in a fair agreement.

Analysis of the potential energy distribution in the 700 cm^{-1} mode shows that it involves the C−Cα stretching and C═O in-plane bending vibrations coupled with angle-bending deformations Cα NC + CCα N of the chain. 90% of the Raman activity comes from the anisotropy of the C═O bond polarizability tensor, and thus the Raman activity of the mode is almost entirely due to the in-plane bending vibration of the bond. Table 1 shows that the EOPs determining the anisotropy of C═O bond polarizability have relatively low statistical uncertainties and that values of the anisotropy obtained in the present work and experimentally determined well agree with each other. Consequently, the presence of the peak with such a large intensity in the

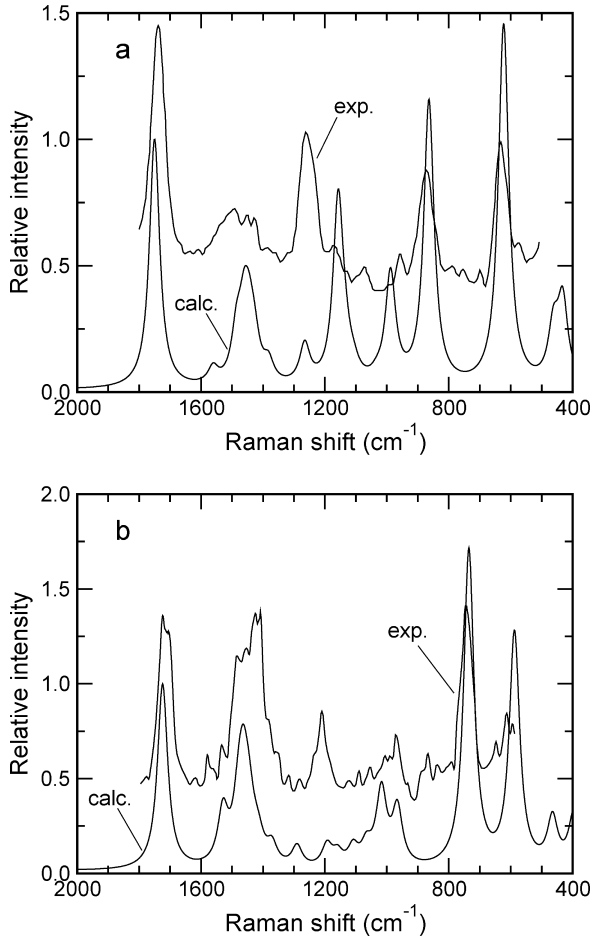

Figure 9. Experimental[28] and calculated Raman scattering spectra of N-methylacetamide (a) and N,N-dimethylacetamide (b) molecules.

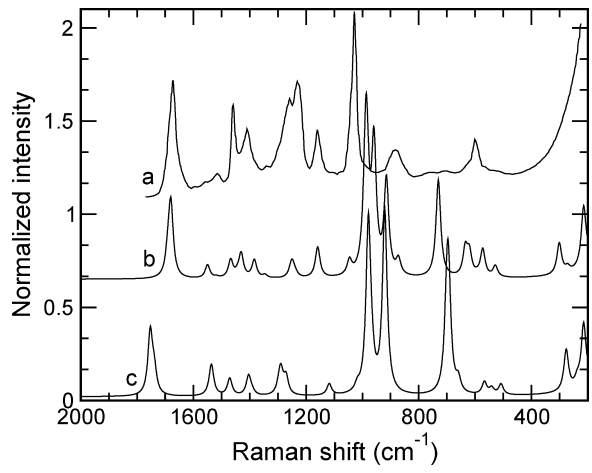

Figure 10. Experimental (a) and calculated Raman spectra of polyglycine I: (b) spectrum computed with a valence force field, ref 33; (c) spectrum obtained with a Urey−Bradley force field, ref 35. The experimental spectrum was adapted from Figure 1 of ref 37. The spectra were shifted along the *y* axis for clarity.

spectrum can unlikely be due to incorrect values of the electro-optical parameters. To our mind, this fact rather points to a problem in the assignment of the experimental bands and to the necessity of performing a new refinement of polyglycine force fields based on an alternative assignment. The corresponding work should go along the following lines:

• precise definition of the structure of experimental samples (e.g., polyglycine I or II) and recording of the vibrational spectra with the use of isotopic substitution permitting an increase in the number of observables,

• use of quantum-chemical calculations to identify the most relevant force constants having a physical sense; the step should also include a precise definition of the set of internal coordinates used in the force field development,

• refinement of the force constants using the experimental data and the parametric models for the spectral intensities. This stage should be followed by an ultime test of the force field transferability.

Such a complete study is well out of the scope of the present work dealing with the EOPs for polypeptides; an example of such a development for aluminosilicate structures can be found in refs 38 and 39. The present results clearly indicate that the computation of the Raman spectra can be of significant help for judging the force field quality.

Besides the limitations inherent in the bond polarizability model, there is another factor affecting the quality of the calculated Raman activities. As it follows from eq 8, the derivative of the polarizability tensor in a normal mode depends on the $(\partial s_j/\partial Q_k)$ quantities that are, in their turn, defined via the derivatives of the Cartesian coordinates with respect to Q_k (and *vice versa*). Obviously, these quantities depend on the potential model used in the vibrational analysis, and thus the choice of the potential model can have a consequence on the computed Raman activities, as it can be seen in Figure 10. To further assess such an influence we have performed a calculation of the Raman activities of the modes of NMA molecule using the vectors of Cartesian displacements obtained in the vibrational analysis of the molecule at the MP2/6-311G++(3df,3pd) level. The comparison of the activities with those computed with the B3LYP/Sadlej basis set level shows that for some modes the R_k values can vary up to 15−20%, and therefore the choice of the force field plays an important role for the Raman intensities computed with the BPM model. The 15% confidence interval is likely the precision that can be achieved with parametric models for Raman intensity calculations.

4. Conclusions

A procedure for obtaining electro-optical parameters of bond polarizability model from results of quantum-chemical computations was applied to derive the EOPs for chemical bonds relevant to biologically important molecules, particularly to peptides. Calculation of the mean polarizability of four test molecules (N-methylacetamide, N,N-dimethylacetamide, dialanine, and diglycine) has shown that the set of parameters reproduces the reference data with a mean absolute error of 0.116 Å^3 and the maximum relative error of 1.6%. The calculation of the nonresonant Raman scattering activities of the normal modes of the molecules has revealed a good correspondence of the BPM computed activities with the reference ones in the region of C−H and N−H stretching vibrations (2800−3800 cm^{-1}), while a systematic underestimation of the Raman activity was found in the energy range

from 1200 to 1500 cm^{-1} that is characteristic to vibrational modes involving variations of angle-bending internal coordinates. The agreement in this spectral region can be improved by changing the equilibrium EOPs for the bonds involved in the coordinates, however, at the expense of the quality in the polarizability tensor components and in the Raman activities of some other modes. An extension of the model to a first-order BPM is expected to improve the quality of the Raman intensity calculations; work in this direction is in progress. Despite the existing deficiencies the resulting set of electro-optical parameters is capable of describing both the polarizability tensor and its derivatives with respect to variations of geometry of large molecules. This conclusion is confirmed by the comparison of the calculated Raman scattering spectra of the N-methylacetamide and N,N-dimethylacetamide molecules with the experimental ones. The results of the calculation of the Raman spectrum of the polyglycine I structure suggest the necessity of revising the force fields for the system and illustrate the usefulness of the Raman spectra computations in the development of force fields of spectroscopic quality. The EOPs can be used in studies of the conformational dependence of the polarizability of peptides and in the Raman spectra simulations by molecular mechanics or molecular dynamics methods.

The comparison of EOPs derived in the present work with those found in the literature shows sometimes significant differences that can reflect the well-known ambiguity of the parameters obtained from experimental data because of ill-conditioning of the corresponding mathematical problem. Hence, the proposed procedure allows one to choose among mathematically equivalent sets of EOPs that which has a physical sense. Furthermore, the approach permits an easy extension to new types of chemical bonds. Such an extension necessitates additional quantum-chemical computations of molecular models having chemical bonds of the new types that are followed by constrained or unconstrained LLSQ fit.

Acknowledgment. The financial support from the Indo-French Centre for the Promotion of the Advanced Research (IFCPAR) under grant no. 3305-1 is gratefully acknowledged.

Supporting Information Available: Structures of molecules in the training and validation sets optimized at B3LYP/pVTZ Sadlej basis set level. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Schweitzer-Stenner, R. *J. Raman Spectrosc.* **2001**, *32*, 711–732.

(2) Schweitzer-Stenner, R. *Vib. Spectrosc.* **2006**, *42*, 98–117.

(3) To, T. T. Molecular Dynamics Study of the Vibrational Spectra of Silica Glasses. Ph.D. Thesis, University of Sciences and Technologies of Lille, Villeneuve d'Ascq, France, 2007. Available online: http://tel.archives-ouvertes.fr/tel-00282362/fr/.

(4) Wolkenstein, M. W. *C. R. Acad. Sci. URSS* **1941**, *30*, 791–794.

(5) Applequist, J.; Carl, J. R.; Fung, K. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.

(6) Maple, J. R.; Ewing, C. S. *J. Chem. Phys.* **2001**, *115*, 4981–4988.

(7) Wang, J.; Vie, X. Q.; Hou, T.; Xu, X. *J. Phys. Chem. A* **2007**, *111*, 4443–4448.

(8) Truchon, J. F.; Nicholls, A.; Iftimie, R. I.; Roux, B.; Bayly, C. I. *J. Chem. Theory Comput.* **2008**, *4*, 1480–1493.

(9) Gussoni, M. In *Advances in Infrared and Raman Spectroscopy*; Clark, R. J. H., Hester, R. E., Eds.; Heyden & Son: London, 1980; Vol. 6, Chapter 2, pp 61−126.

(10) Smirnov, K. S.; Bougeard, D. *J. Raman Spectrosc.* **2006**, *37*, 100–107.

(11) Smirnov, K. S.; Bougeard, D.; Tandon, P. *J. Phys. Chem. A* **2006**, *110*, 4516–4523.

(12) Long, D. A. *Proc. Roy. Soc. (London)* **1953**, *A217*, 203–221.

(13) Volkenstein, M. V.; Gribov, L. A.; Eliashevich, M. A.; Stepanov, B. I. *Molecular Vibrations*; Nauka: Moscow, 1972.

(14) Long, D. A. *The Raman effect*; John Wiley & Sons: Chichester, 2002.

(15) Gough, K. M.; Dwyer, J. R. *J. Phys. Chem. A* **1998**, *102*, 2723–2731.

(16) Gough, K. M.; Dwyer, J. R.; Dawes, R. *Can. J. Chem.* **2000**, *78*, 1035–1043.

(17) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.

(18) Sadlej, A. J. *Theor. Chim. Acta* **1992**, *79*, 123–140.

(19) Halls, M. D.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 8819–8824.

(20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision D.01*; Gaussian, Inc.: Wallingford, CT, 2004.

(21) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing*; Cambridge University Press: New York, 1992; pp 670−675.

(22) Applequist, J.; Quicksall, C. O. *J. Chem. Phys.* **1977**, *66*, 3455–3459.

(23) Van Hemert, M. C.; Blom, C. E. *Mol. Phys.* **1981**, *43*, 229–250.

(24) Furer, V. L.; Alekseev, V. V. *Zh. Prikl. Spektrosk. (USSR)* **1986**, *45*, 951–955.

(25) Yokoyama, I.; Miwa, Y.; Machida, K. *J. Phys. Chem.* **1991**, *95*, 9740–9746.

(26) Yokoyama, I.; Miwa, Y.; Machida, K.; Umemura, J.; Hayashi, S. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 400–413.

(27) Montero, S.; Bermejo, D. *Mol. Phys.* **1976**, *32*, 1229–1232.

(28) Triggs, N. E.; Valentini, J. J. *J. Phys. Chem.* **1992**, *96*, 6922–6931.

(29) Fukushima, K.; Ideguchi, Y.; Myazawa, T. *Bull. Chem. Soc. Jpn.* **1963**, *36*, 1301–1307.

(30) Gupta, V. D.; Trevino, S.; Boutin, H. *J. Chem. Phys.* **1968**, *48*, 3008–3015.

(31) Abbe, Y.; Krimm, S. *Biopolymers* **1972**, *11*, 1817–1839.

(32) Moore, W. H.; Krimm, S. *Biopolymers* **1976**, *15*, 2439–2464.

(33) Dwivedi, A. M.; Krimm, S. *Macromolecules* **1982**, *15*, 177–185.

(34) Fillaux, F.; Fontaine, J. P.; Baron, M. H.; Leygue, N.; Kearley, G. J.; Tomkinson, J. *Biophys. Chem.* **1994**, *53*, 155–168.

(35) Porwal, V.; Misra, R. M.; Tandon, P.; Gupta, V. D. *Indian J. Biochem. Biophys.* **2004**, *41*, 34–39.

(36) Kajava, A. V. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *D55*, 436–442.

(37) Small, E. W.; Fanconi, B.; Peticolas, W. L. *J. Chem. Phys.* **1970**, *52*, 4369–4379.

(38) Ermoshin, V. A.; Smirnov, K. S.; Bougeard, D. *Chem. Phys.* **1996**, *209*, 41–51.

(39) Creton, B.; Bougeard, D.; Smirnov, K. S.; Guilment, J.; Poncelet, O. *J. Phys. Chem. C* **2008**, *112*, 10013–10020.

CT800510Y