THE JOURNAL OF PHYSICAL CHEMISTRY A

ARTICLE

pubs.acs.org/JPCA

Theoretical Modeling of Peptide α-Helical Circular Dichroism in Aqueous Solution

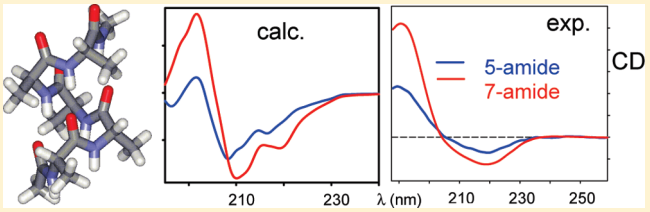
Jakub Kaminský,*,† Jan Kubelka,*,‡ and Petr Bouř*,†

†Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo nám. 2, 166 10 Prague, Czech Republic

‡Department of Chemistry, University of Wyoming, 1000 East University Avenue, Laramie, Wyoming 82071, United States

Supporting Information

ABSTRACT: Reliable modeling of protein and peptide circular dichroism (CD) spectra in the far UV presents a challenge for current theoretical approaches. In this study, the time-dependent density functional theory (TDDFT), configuration interaction with single excitation (CIS), and transition dipole coupling (TDC) were used to assess the most important factors contributing to the CD spectra of the α-helical secondary structure. The dependence on the peptide chain length and also the role of the flexibility and solvent environment were investigated with a model oligopeptide Ac-$(Ala)_N$-NH-Me, (N = 1, ..., 18). Both the TDDFT and TDC-like methods suggest that the CD curve typical for the α-helix arises gradually, but its basic characteristic is discernible already for peptides with 4−5 amino acid residues. The calculated dependence was in a qualitative agreement with experimental spectra of short α-helices stabilized by the histidine-metal binding. The TDDFT computations of the CD were found to be unusually sensitive to the basis set and solvent model. Explicit hydration and temperature fluctuations of the peptide geometry, simulated with the aid of molecular dynamics (MD), significantly influenced the CD and absorption spectral shapes. An extensive averaging over MD configurations is thus required to obtain a converged spectral profile in cluster simulations. On the other hand, both the TDDFT and TDC models indicate only a minor influence of the alanine side chains. The CIS and TDC calculations also point toward a relatively small effect of the helix−helix interaction on the CD spectral profiles. For a model system of two helices, the CIS method predicted larger changes in the spectra than TDC. This suggests other than interactions between peptide chains, such as mutual polarization, can have a minor, but measurable, effect on the CD spectrum.



INTRODUCTION

Protein structure and folding are intensely studied using a wide variety of experimental spectroscopic techniques. While nuclear magnetic resonance (NMR)[1] and X-ray crystallography[2] provide structural information up to positions of individual atoms, the lower-resolution techniques, such as the electronic circular dichroism (CD),[3,4] are frequently used for efficient monitoring of global conformational changes. CD spectra have been traditionally interpreted on empirical or semiempirical grounds.[5,6] For example, α-helical proteins have a characteristic positive band around 190 nm and two negative lobes at 208 and 222 nm. The 3_{10}-helical CD spectra are generally believed to be very similar.[7,8] The β-sheet structure exhibits a positive CD band around 198 nm and a negative feature at 218 nm.[5] The spectrum of so-called random coil structure also shows a characteristic signal with an opposite sign pattern, very similar to that of a left-handed polyproline-like helix,[9,10] only much weaker.

In practice, however, measured spectra can deviate from the canonical curves. In principle, these variations can be utilized to gain additional information about protein structural properties. For example, it has been suggested that a long α-helix can be distinguished from several shorter α-helices according to their CD signal.[11] This obviously requires understanding the dependence of the spectra on the helix length, which has been the subject of several studies.[12] In proteins, a number of other effects may also become significant, such as solvent exposure of some structural elements and burial of others, and the interaction between two (or more) α-helices in close proximity.[13] Reliable simulations of the CD spectra thus could provide valuable clues about the sensitivity of the CD spectral signatures to such effects.

The theoretical foundations of the CD phenomenon are well established,[14−17] and first predictions of the CD for various types of peptide secondary structure, based on simplified models, date back to the 1950s.[18−20] The right-handed structure of α-helices was predicted from the CD method before it was confirmed by X-ray.[21] Nowadays, a full quantum-chemical treatment and predictions of the spectra are possible with the aid of many program packages, most commonly using the time-dependent density functional theory (TDDFT).[22−24]

However, the precision of the ab initio/TDDFT computational methods for larger systems is limited, and a theoretical explanation of all experimental spectral features is still challenging. Gas phase models often do not provide realistic spectra.[25] Better results have been obtained by combined MM/QM methods,

Received: November 1, 2010
Revised: January 18, 2011
Published: February 15, 2011

taking into account peptide flexibility and the solvation.[26,27] Another obstacle lies in the limited and unpredictable precision of TDDFT. Sometimes, precise wave function computations are performed on smaller subsystems and used to predict spectral properties of a larger molecule.[28−34] In such semiempirical treatment, the chromophores (amide groups) are most often considered separately, and simplified models are employed to approximate their interactions. In the transition dipole coupling (TDC) scheme, the dipole−dipole interaction dominates,[35−38] whereas more complex forces are considered in the so-called matrix methods, allowing also for the interchromophore charge transfer.[39] These and other simplified simulation approaches have been recently reviewed.[40]

In this work, we employ TDDFT as the primary method for modeling the spectral dependence on the peptide chain length, as this approach allows for long-range delocalized polarization effects, and it is not dependent on molecule-dependent empirical parameters. Indeed, as shown below, some results significantly differ from those obtained with the individual chromophore-based models (e.g., TDC), which we have also tested. The simulated spectral shapes and length dependence are compared to an experimental model based on short helices stabilized by a metal ion. Molecular dynamics (MD) computations are performed, which indicate that geometry fluctuations of the peptide and surrounding water cause a large dispersion of the CD signal and inhomogeneous broadening of spectral lines. Therefore, a significant number of geometries must be averaged for a converged spectrum. Similar sensitivity to the environment and the need to average many MD snapshots were also recently observed for simulations of NMR spectra.[41] On the other hand, the tertiary peptide structure, modeled here as an interaction of two model peptide chains, leads only to relatively minor, although potentially detectable, contributions to the spectral patterns.

METHODS

Quantum Computations. Initial geometries of Ac-$(Ala)_N$-NH-Me ($N = 1-18$) and Ac-$(Gly)_4$-NH-Me were generated with the MCM program[42] using standard α-helical geometry parameters ($\varphi = -57°$, $\psi = -47°$, $\omega = 180°$).[43] For the length-dependence test, the CD spectra were calculated after reoptimization of the structures at the HF/6-31G level with all torsions fixed, using the Turbomole (version 5.10)[44] and Gaussian 03[45] programs at the TDDFT[46] or CIS[47] approximation level. The B3LYP functional was used by default as it provides reasonable amide electronic transition energies.[26] The 6-31G** and 6-311++G** Pople-style basis sets were used as specified; minor computations were done with the SVP basis (default in Turbomole software) as specified below. The COSMO[48,49] continuum solvent correction (known as CPCM in Gaussian programs) was added in this case, with default parameters for water.

Molecular Dynamics Simulations. The hydration and flexibility factors were explored with the aid of the Tinker molecular dynamics package.[50] The Ac-$(Ala)_4$-NH-Me peptide in the α-helical conformation (using the standard φ, ψ angles, as above) was constructed and soaked in a periodic cubic water box 18.64 Å wide.[50] The standard Amber99[51] force field was used for the simulations. Although later variants of Amber were proposed to provide better peptide secondary structure,[52] the influence of the force field on CD was not investigated, as it goes far beyond the scope of the present study. However, to mimic the experimental structure of protein helices, not stable for the shorter fragments, we performed two types of dynamics: (1) with fixed peptide geometry, to model the hydration only, and (2) with weak harmonic restrictions (with the force constants of 0.02 kcal/mol/deg) imposed on φ and ψ, to qualitatively investigate the effect of the peptide geometry fluctuations on the spectra.

After minimization and equilibration, geometries were stored each 500 steps during an NpT MD propagation, using integration time step of 1 fs, pressure 1 atm and temperature 300 K. The total simulation time was 50 ps as longer simulations did not change the hydration pattern. The CD spectra were calculated for the MD snapshots at the TDDFT level, as described above, using either the peptide geometry only or including the hydrogen-bonded water molecules (those within 3.6 Å from the amide nitrogen or oxygen).

Semiempirical Approximations. In addition to the TDDFT approach, the transition dipole coupling (TDC) model[53] and the Hirst matrix method[39] were used to calculate the CD spectra. The matrix method results are not shown, as they are essentially identical to the TDC ones. To simulate the effect of solvation, a set of 90 clusters of *N*-methylacetamide (NMA) and water molecules from the first hydration sphere, obtained previously,[26] were used for the TDC calculations. The electronic transition energies and electric dipole moments of each cluster were calculated at the BPW91[54]/Sadlej-pVTZ[55]/CPCM level, transferred to the polypeptide helix, and the spectra were averaged. The BPW91 functional was chosen instead of B3LYP because it gave the transition wavelengths of NMA slightly closer to the experiment. In trial computations, the polarization triple-ζ Sadlej-pVTZ[55] basis set provided similar results as a larger pc-3[56] basis set, but in a shorter time. Finally, the spectral curves were obtained from the simulated absorption and CD intensities, by a convolution with a Gaussian function, using the full width at half height of 10 nm by default for the TDC model. For the quantum methods, we used 8 nm in Figures 5 and 8 and 15 nm for all the other cases. All presented spectra are normalized to one amide group.

Experimental Section. To obtain experimental models of very short α-helices, we utilized the stabilizing effect of a metal ion bound to a histidine pair in i^{th} and $(i + 4)^{th}$ positions.[57] The Ac-HAAAH-NH_2 and Ac-HAAAHAA-NH_2 (H = histidine, A = alanine) oligopeptides were synthesized using the standard solid phase FMOC strategy, and their CD spectra were recorded with and without the presence of Co^{2+} ions (20 mM), on a JASCO 810 spectropolarimeter. The peptides are presumably mostly disordered under these conditions; however, differential change is supposed to provide approximate spectra of short α-helices.[57] Spectra of the penta- and heptapeptide α-helix were thus calculated as the differences between the CD signals obtained with and without the metal.

RESULTS AND DISCUSSION

Basis Set and Solvent Dependence. For all approaches, but in particular for TDDFT, the CD simulation is very sensitive to computational parameters. As an example, in Figure 1 the Ac-Ala-NH-Me diamide UV absorption and CD spectra are plotted as calculated at three levels of theory. All the B3LYP/6-31G**, B3LYP/6-311++G** and B3LYP/6-311++G**/CPCM methods provide similar absorption patterns in the n-π^* (∼230 nm) and $\pi-\pi^*$ (<200 nm) regions, although the 6-31G** intensity is significantly smaller (note the 4× multiplication factor in the Figure 1). The CPCM solvent correction causes a shift of the

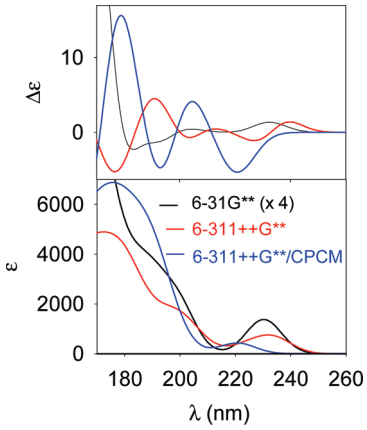


Figure 1. Ac-Ala-NH-Me, CD and absorption spectra calculated with several different basis sets, and with the CPCM solvent correction, using the B3LYP functional.

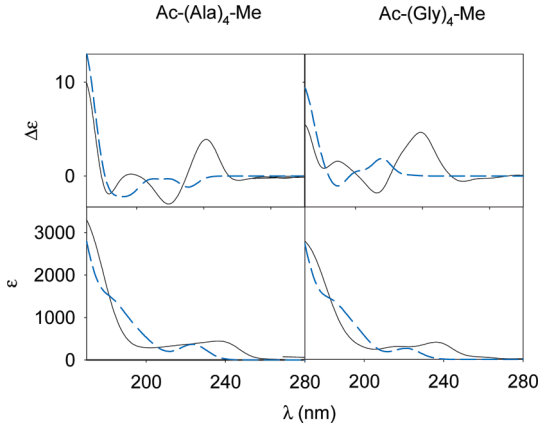


Figure 2. B3LYP/6-31G** CD and absorption spectra of alanine and glycine tetrapeptides, calculated in vacuum (solid black lines) and with the CPCM water environment (dashed blue line).

absorption threshold to higher energy (lower wavelength), similarly as observed previously for the NMA.[26] More significant differences are observed in CD signals throughout the entire spectral region. Closer analysis reveals that the angles between the electric and magnetic transition dipole moments remain very close to 90° (Table S1 in Supporting Information (SI)), which makes the resultant rotational strength much more sensitive to the parameter changes then the absorption. Also the long computational time needed for higher computational levels and longer peptides makes an accurate simulation of the spectral shape problematic.

Similar effects can be observed for longer peptides: in Figure 2, left, the vacuum Ac-$(Ala)_4$-NH-Me B3LYP/6-31G** spectra are compared to the CPCM results at the same level of theory. Here, unlike for the dipeptide (Figure 1), the aqueous environment causes an overall decrease of the CD signal. The UV absorption intensity is not significantly influenced by CPCM, but the absorption threshold is again shifted to lower wavelengths. This can be explained by the formation of the internal hydrogen bonds in the tetrapeptide, so that the amide group transitions are less sensitive to the outer water environment.

On the other hand, removal of the methyl side-chains of alanine causes only small changes in the spectra (cf. the Ac-$(Gly)_4$-NH-Me simulation on the right-hand side of Figure 2), indicating

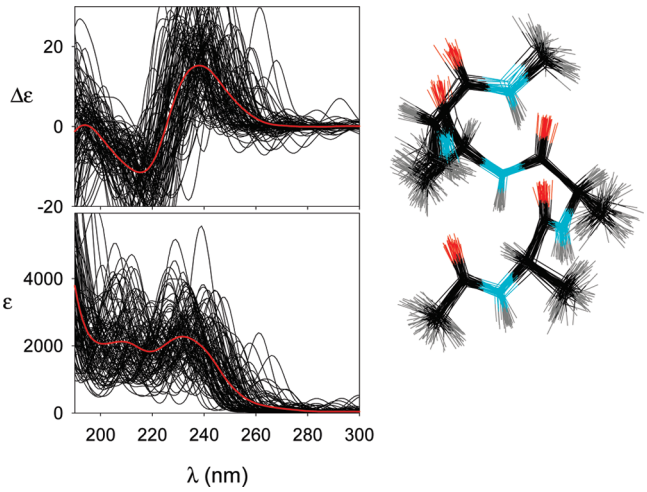


Figure 3. Simulated (B3LYP/6-31G**) CD ($\Delta\varepsilon$) and absorption (ε) spectra of Ac-$(Ala)_4$-NH-Me for 100 geometries obtained from MD trajectory. The average is plotted by the red lines; the geometry variation is indicated on the right-hand side by the overlap of a few randomly selected structures.

the relative isolation of the amide group electron system from the side-chains. This is consistent with several previous TDC and matrix model studies,[31] which suggested that although charge transfer across the peptide chain can influence spectral intensities, an electronically isolated chromophore model explains the most important spectral features.

Influence of Water Hydrogen Bonding and Peptide Dynamics on CD. Even the limited peptide movement under the constraints used to enforce the helical conformation causes significant variations of UV absorption and CD intensities. As can be seen in Figure 3 for the spectra of 100 selected MD geometries, the absorption varies within ~50−150% of the average intensity, but the CD intensity variations are significantly greater, and even lead to sign changes at particular wavelengths. Even larger dispersion of the absorption and CD is apparent when explicitly hydrated geometries are considered (Figure S1).

Obviously, the observed spectra would correspond to the average over the individual geometries (red lines in Figures 3 and S1) The MD simulations therefore clearly indicate that the peptide and solvent fluctuations are the main source of the inhomogeneous band broadening in the experiment. They also imply that for realistic simulations with explicit water modeling a large number of peptide-water structures or clusters needs to be averaged. These results are in agreement with the previous observations for N-methylacetamide,[26] where the hydration and geometry variations caused significant changes in the electronic transition energies and intensities. The hydrogen-bonded water molecules enable a partial charge transfer during the electronic excitations, which significantly influences the amide electronic spectra. Extended solvent polarization models were recently proposed[58,59] to properly account for the charge transfer effects. Unfortunately, increasing the level of the approximation for realistic α-helical models is currently not feasible. For example, obtaining a sufficient number of tetra-alanine spectra for averaging took about two months of the CPU time (64 bit AMD, 2.2 GHz). Similar needs for taking into account a large number of structures or clusters for adequate averaging have been noted in simulations of NMR[41,60] and ROA[61] spectra.

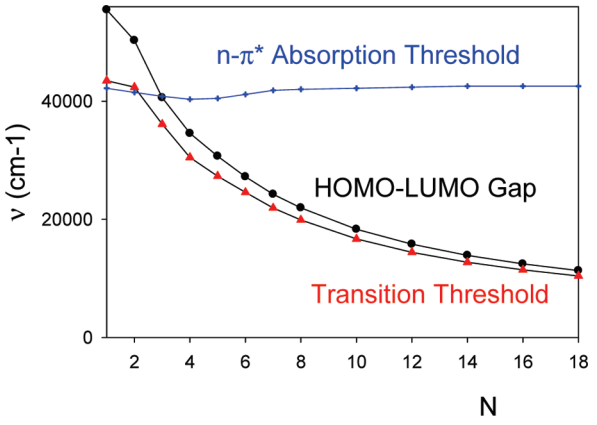


Figure 4. Calculated (B3LYP/6-31G**) HOMO−LUMO gap, singlet-transition threshold and approximate threshold for the observable (n-π^*) absorption in Ac-$(Ala)_N$-NH-Me.

To test the robustness of the simulations we modified the MD/averaging procedure, so that 1) the MD structures were partially optimized at the B3LYP/SVP level by the normal mode method,[62] and 2) a continuum solvent environment (COSMO[63]) was added to the TDDFT calculations. These variations, however, had only minor influence, and indicate that the resultant spectral shapes and inhomogeneous broadening of the spectral lines are relatively independent of these simulation parameters.

Alternatively, the explicit hydration was built into the TDC model, where the snapshots of the NMA molecule and hydrogen-bonded waters from a MD simulation were used as the source of the dipole moments. The simulated spectra (Figure S2 in SI) exhibit a similar variance to those obtained by TDDFT, confirming the dominance of the solvent influence over dynamical fluctuations. The TDC model even provides more realistic spectral profiles, in particular the negative n-π^* signal around 225 nm, as seen experimentally. For NMA, the B3LYP functional gives lower wavelength for the $\pi-\pi^*$ and n-π^* transitions than the BPW91, but the absorption profiles are similar (Figure S3).

Length Dependence of the Permanent Dipole and Orbital Energies. The contribution of the permanent α-helical dipole to the stability of α-helices in proteins is well-known.[64] For example, four-helix bundle proteins are to a large extent stabilized by the electrostatic interactions of the helical macrodipoles.[65] Presumably, the polarization affects the amide chromophore and hence the spectral properties. The helical macrodipole may also become an important factor in influencing the spectra of two or several interacting α-helices, which is investigated below. The calculated dependence of the dipole moment on the peptide length for our oligoalanine model (Figure S4) confirms that the mutual polarization of the amide groups, i.e. the dipole per one amide, is larger in longer α-helices. The dipole magnitude is in agreement with previous results.[66] As expected, the larger 6-311++G** basis set and the CPCM environment enhance the polarizability, and provide larger dipoles than the vacuum 6-31G** calculation. The dipole starts to increase for $N = 3$ when a complete α-helical turn can be formed. A slight change in the trend is also observed for $N = 6$ when a second turn is completed, so that at least one amide group is hydrogen bonded on both ends.

Similarly, the orbital energies (Figure S5), in particular the HOMO−LUMO gap, converge relatively slowly. From the graph in Figure S5, the limit for infinitely long helices cannot be reliably extrapolated; however, the dependence indicates that longer helices are better conductors. This is also consistent with the known experimental dependence of the helical conductivity on the electric field, and the semiconductor properties.[67]

According to empirical observations, the computed HOMO−LUMO gap is approximately in the middle of singlet and triplet transition thresholds, providing exact functional and complete basis sets are used.[68] For an approximate one, the relation of orbital and excitation energies may not be so tight.[69] However, as documented in Figure 4, for the Becke's B3LYP functional used in our case, the transition threshold (energy of the lowest-energy transition) follows the gap quite faithfully. Interestingly, the apparent absorption threshold, i.e. the approximate edge of the n-π^* band, is changing much less, because the lowest-energy transitions have negligible absorption intensities. These satellite low-energy electronic transitions exist only in vacuum and disappear in computations where the solvent CPCM environment is included, as also observed previously.[26]

Dependence of the Spectra on Peptide Length. A general form of the CD spectra dependence of the α-helix length has been discussed for a long time, as it is important for interpretation of spectroscopic experiments.[70] Although it is generally believed that native proteins cannot form helices shorter than about 14 amino acid residues,[71] shorter experimental models are available.[12] Some theoretical works argue that the CD spectrum of a short α-helix is significantly different from that of a longer one.[72]

The length dependence of the CD spectra of Ac-$(Ala)_N$-NH-Me α-helices ($N = 1-18$) is illustrated in Figure 5, as simulated by the TDC, B3LYP/6-31G**, B3LYP/6-31G**/CPCM and B3LYP/6-311++G**/CPCM models. For TDC, a randomly selected NMA/water cluster was used as a source of the transition dipoles and energies. We also note that a CIS/6-31G** calculation (not shown) provided band shapes similar to B3LYP/6-31G**, but with unrealistic transition wavelengths.

As can be seen from Figure 5 the B3LYP calculations are quite sensitive to the detailed parametrization. Similarly to the diamide (Figure 1) the CPCM solvent correction generally leads to larger CD intensities, while the absorption is approximately the same as in the gas phase. Also, without the solvent, the absorption and CD spectral shapes converge more slowly with the number of amides than for CPCM. Within CPCM, the larger basis set (6-311++G**) yields very similar absorption spectrum as the smaller one (6-31G**), but the CD intensity is quite different. For 6-311++G** the n-π^* minimum (at ~224 nm) is much deeper, and is present for all peptide lengths. Unlike the 6-31G* calculation, the positive CD $\pi-\pi^*$ band (~190 nm) is well developed with 6-311++G**, in good agreement with the TDC model as well as a typical experiment. At about nine amide units, the second "n-π^*" minimum appears near 208 nm in the B3LYP/6-311++G**/CPCM calculation; this feature is even more pronounced in the TDC and B3LYP/6-31G**/CPCM spectra. Note that it is generally accepted that ideal α-helices provide the two (208 and 222 nm) CD minima, of approximately equal intensity.

In spite of the differences between various models, several common trends in the length dependence can be generalized. All the models predict the diamide ($N = 1$) spectra to be radically different from longer oligomers. The diamide n-π^* absorption maximum (around 225 nm) is shifted by 4−8 nm to shorter wavelengths from the maxima of the longer peptides. On the

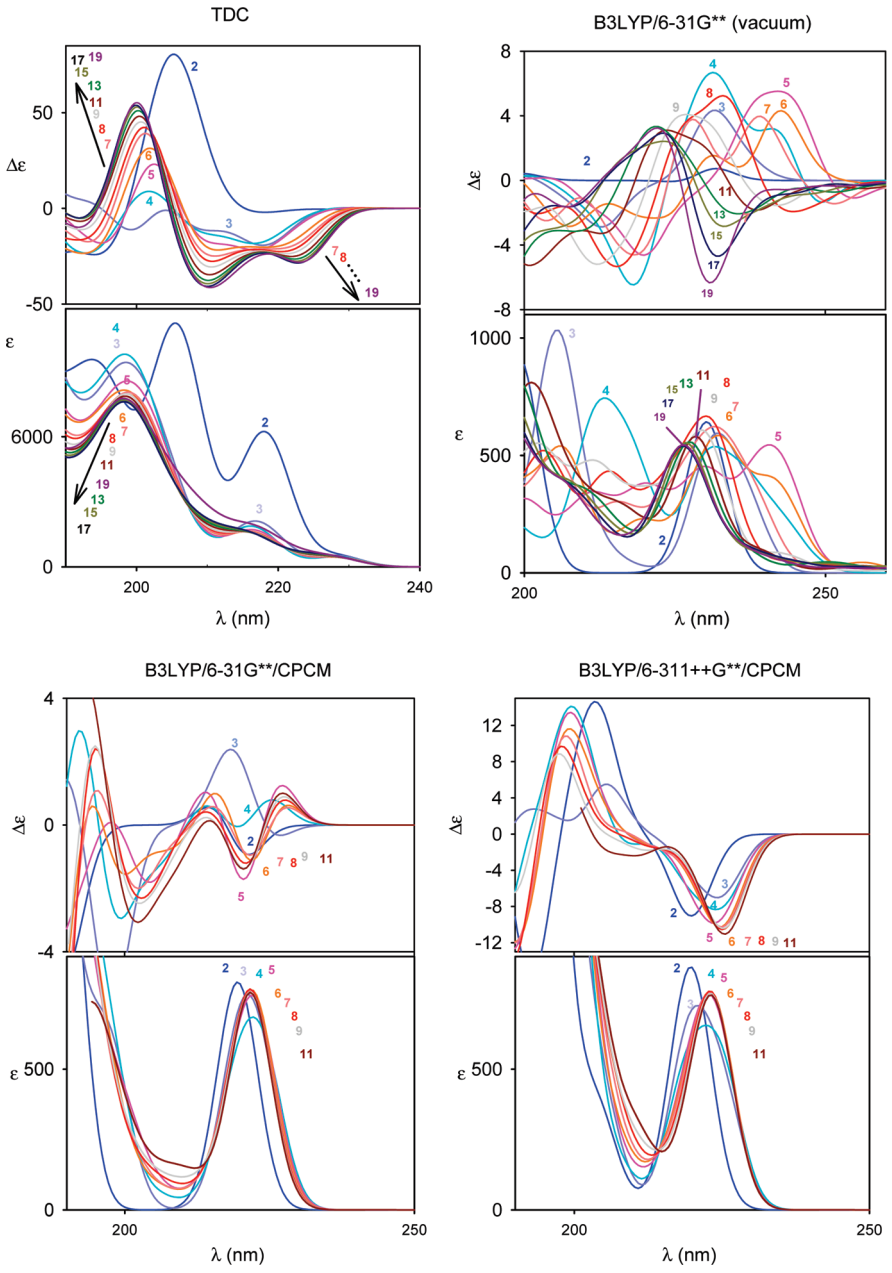


Figure 5. Dependence of the CD and absorption spectra of Ac-$(Ala)_N$-NH-Me on the peptide length (number of amides = $N + 1$ is indicated), as calculated at the four approximation levels.

other hand, the absorption profiles remain quite similar, and absorption spectra of helices with more than four amide groups ($N > 5$) are virtually indistinguishable.

Focusing on the B3LYP/6-311++G**/CPCM results, the highest level of theory used, it is apparent that the CD profiles stabilize more slowly than the absorption. Nevertheless, the n-π^* negative band intensities for the tetra- and penta-amide, for example, differ only by about 10%. We do not extensively discuss in detail the low-wavelength $\pi-\pi^*$ ("far UV") region, because of the expected computational error and many overlapping transitions.[26,31] Also, an energy cutoff that is necessary for practical TDDFT computations cannot be set too high, in order to keep the computational time reasonable for larger molecules.

The simplest TDC model (Figure 5) qualitatively resembles the B3LYP/6-311++G**/CPCM calculation, which is consistent with the traditional assumption that the peptide CD is primarily determined by the dipole−dipole interaction of the amide chromophores.[35] The TDC spectral curve is similar to those obtained previously with empirical dipoles only,[12,34] and the double negative minimum qualitatively best corresponds to the experimental α-helical profile. On the basis of these results, we can conclude that the α-helical spectra of short helices (2−5 amino acids), if experimentally realizable, are different from the longer ones, but in an aqueous environment, the spectral shapes stabilize relatively quickly with respect to the α-helix length. However, the intensity magnitude keeps gradually increasing with peptide length.

The saturation behavior with respect to the peptide length is even more apparent from the dependence of the $n-\pi^*$ band intensities and wavelengths plotted in Figure 6. The TDC model

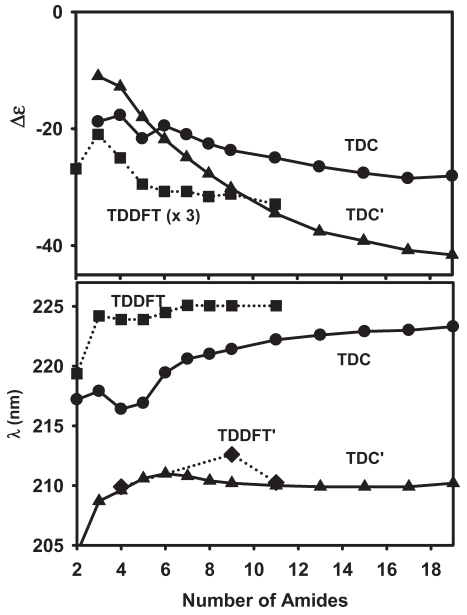


Figure 6. Dependence of wavelengths and CD intensities of α-helical negative $n-\pi^*$ bands on the chain length, obtained by B3LYP/6-311++G**/CPCM and TDC computations.

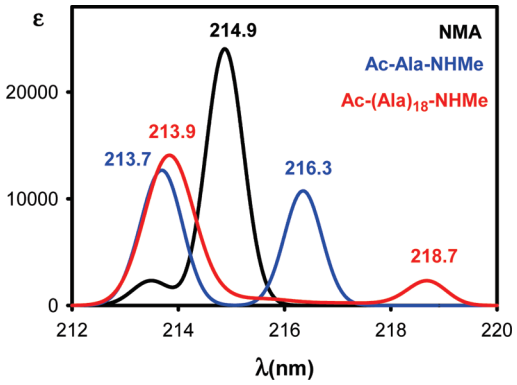


Figure 7. Absorption spectra of the $n-\pi^*$ band modeled by TDC, plotted with 1 cm^{-1} bandwidth, to emphasize the different band splitting in Ac-Ala-NH-Me and Ac-$(Ala)_{18}$-NH-Me. BPW91/Sadlej-pVTZ/CPCM NMA dipoles and transition energies were used.

results in two bands in this region, and its convergence is clearly much slower than that of the B3LYP/6-311++G**/CPCM calculation, which stabilizes already at $N \sim 6$. Experimentally, as the concentration and measured intensities often contain a large error (typically 5−10%), a change of the CD shape during the peptide lengthening obtained by TDC may be more indicative than the intensity.

The slower TDC length convergence compared to TDDFT is probably in better agreement with available experimental data on very short peptides,[12] see also Figure 8 below. This, along with the better agreement of the TDC overall CD bandshapes with experiment, would suggest that in principle the more advanced TDDFT does not work so well here. However, at present there is not sufficient data to make definite conclusions; for example, the experimentally observed behavior can be caused by a larger flexibility of the short peptides or by environmental factors, which is difficult to include in the model. In spite of the differences, the TDC and TDDFT computations agree in main features, for example, both predict a small dip at the frequency and CD intensity dependencies (Figure 6) for the number of amides equal to five ($N = 4$).

The $\pi-\pi^*$ transitions have the electric transition dipole moments oriented predominantly perpendicularly to the helical axis, whereas the $n-\pi^*$ dipoles are approximately parallel, in agreement with experimental observations.[5] The electronic transitions exhibit features of the helical symmetry for relatively short peptides, already with $N \sim 5$. The delocalized exciton modes across the peptide chain are also partially responsible for the CD enhancement in longer helices.[73] However, the enhancement is relatively minor, if compared with the oriented DNA molecules, where sometimes ~10 times higher CD intensity (normalized to a single chromophore) is observed for longer polymers if compared to shorter segments.[74]

Nevertheless, the α-helical exciton modes in peptides behave similarly to those in DNA, which can be demonstrated with the TDC model in Figure 7. Here, the spectra were plotted with narrower widths (1 nm) to highlight the details of the band splitting. For the longest-wavelength $n-\pi^*$ band, for example, the central wavelength λ_0 splits to $\lambda_0 \pm V$ for the dimer, where $V \sim 1.3$ nm corresponds to the interaction potential between two neighboring amide groups (for this wavelength, $V[\text{eV}] \sim 0.026V[\text{nm}]$). For longer chains, the exciton bands can split up to $\lambda_0 \pm 2V$.[19,73] In our case (Figure 7), the exciton splitting corresponds to $V \sim (218.7-213.9)/2 = 1.2$ nm, that is, almost to the theoretical value. Note, however, that these spectra were generated with the $n-\pi^*$ transitions only. Analysis of the TDC wave function revealed that the $n-\pi^*$ intensity diminishes due to the coupling with the lower-wavelength transitions, such as $\pi-\pi^*$, and also the splitting pattern is in general more complicated.

The magnitude of the length variation predicted by the TDC and TDDFT methods seems to be consistent with the experimental model, as can be seen in Figure 8. Around the peptide length of six amide units the band positions and spectral shapes are already nearly stabilized, and resemble typical long peptide or protein α-helices. Nevertheless, additional amino acid residues still cause an increase in the signal magnitude, both in the $n-\pi^*$ and $\pi-\pi^*$ spectral regions. The $\pi-\pi^*$ band is clearly more sensitive: both the experiment and the TDC model indicate much greater CD intensity changes caused by the added helix length than for the $n-\pi^*$ band. TDDFT predicts a small decrease in the $\pi-\pi^*$ band intensity, but captures much better the shape of the negative ($n-\pi^*$) CD signal, while TDC overestimates the intensity around 210 nm. The calculated overall $\pi-\pi^*/n-\pi^*$ intensity ratio seems to be lower than in experiment; this corresponds to the computational error discussed above. Using the TDC model, we found that the $\pi-\pi^*$ and $n-\pi^*$ transitions are strongly coupled and a small change in the model parameters can lead to large intensity redistributions.

Effect of Inter-Helical Interactions. To estimate, at least qualitatively, the influence of other molecular parts (tertiary structure), we computed at the CIS/3-21G level spectra of the dimer, $(\text{Ac-Ala}_{12}\text{-NHMe})_2$. The system was in part inspired by recent folding studies on the viral P22 protein,[75] where it is important to distinguish the influence of the environment on the CD spectra from the changes in secondary structure. Five orientation angles (180, 0, 20, 90, and 160°) and three distances (9, 10.5, and 15 Å) were investigated, as indicated in Figure 9.

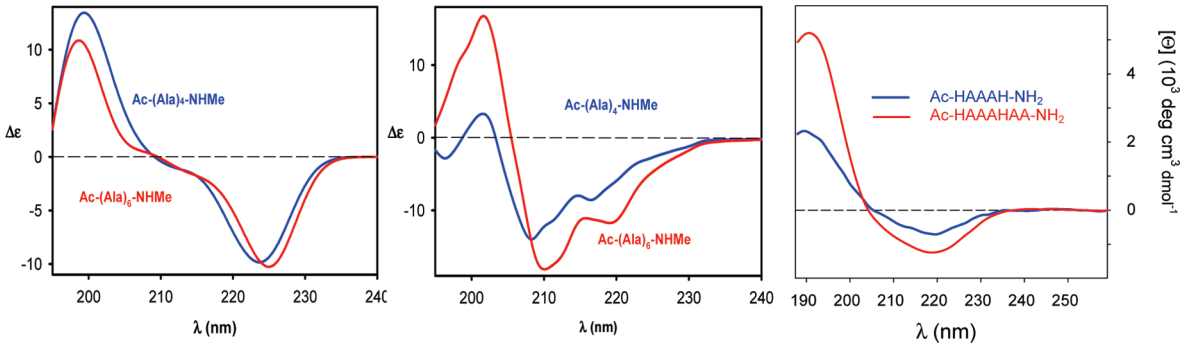


Figure 8. Comparison of CD spectra of the indicated shorter and longer oligopeptides obtained by the TDDFT (left) and TDC (middle) calculations with the experiment (right). For TDDFT, the B3LYP/6-311++G**/CPCM method was used, TDC was performed with TDDFT dipoles and energies obtained by BPW91/Sadlej-pVTZ/CPCM on 90 NMA/water clusters and averaged, the experimental spectra were obtained as a difference induced in the signal by addition of the Co^{2+} ions. See Figure S6 for the original experimental spectra.

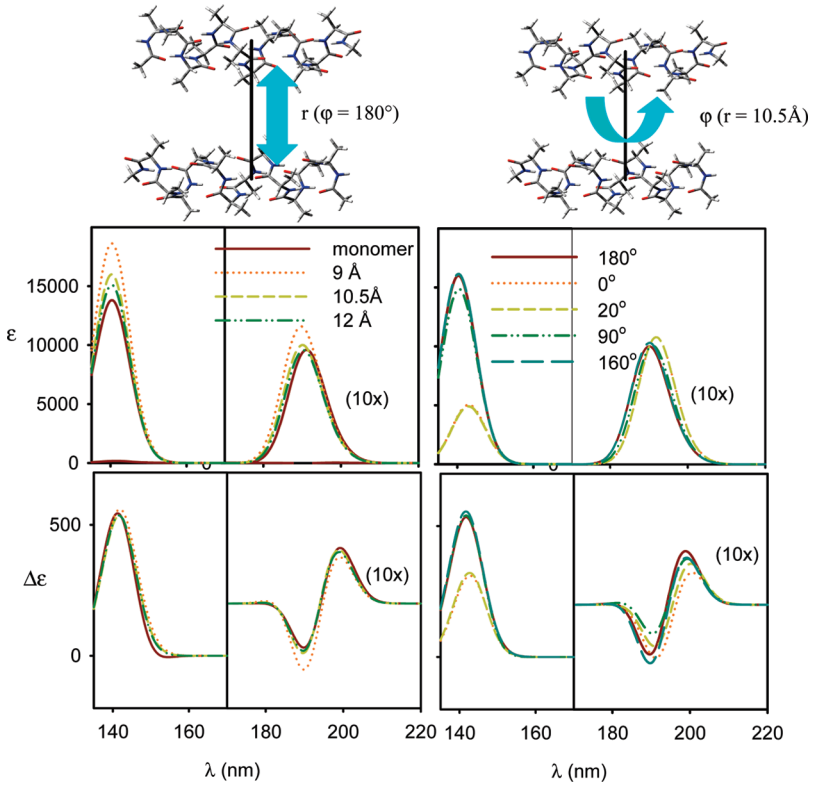


Figure 9. CIS/3-21G absorption and CD spectra of a helical dimer calculated for various interhelical distances (for antiparallel helices, left) and orientations (at a distance of 10.5 Å, right).

The CIS approach was adopted as it provided better spectral shapes than TDDFT with the smaller basis set. Obviously, the CIS/3-21G level spectra (Figure 9) deviate from the experiment, in particular, in the $\pi-\pi^*$ transition energies calculated at ~145 nm. However, we analyzed the CIS and DFT transitions by looking at the relevant orbitals and smaller systems (NMA, pentaamide) and verified that the main features of the electronic states are similar, in spite of large frequency differences between these two quantum approaches. In the penta-amide, CIS and DFT provided the same helical symmetry of the $\pi-\pi^*$ and $n-\pi^*$ bands. Finally, we have also compared the CIS/3-21G and the TDC dimer calculations (SI, Figure S7).

In the simulated absorption and CD (Figure 9), the influence of the helical interaction on the spectra is clearly apparent. For the antiparallel orientation, which corresponds to the most frequent arrangement of helices in living organisms,[71] the UV absorption is surprisingly more dependent on the distance than the CD. For the $n-\pi^*$ band, calculated at this level at 195 nm, the absorption and CD intensities at the separation of 9 Å are by about 10% larger than for the monomer.

The orientation of the two α-helices (right-hand side of Figure 9) has an even larger effect than the distance variation; in the extreme case of 0 or 20°, the $\pi-\pi^*$ absorption and CD drops to 25 and 50%, respectively. The higher sensitivity of the $\pi-\pi^*$ signal is also consistent with the dependence on the

dx.doi.org/10.1021/jp110418w |J. Phys. Chem. A 2011, 115, 1734–1742

peptide length (Figure 8). The dimer spectra simulated by the TDC (Figure S7) are quite similar to those of the monomer. This can be explained by a lack of the static polarization effects absent in the TDC model. The dipole−dipole interaction of TDC quickly diminishes with the distance. Nevertheless, the CIS computation suggests that the influence of the α-helix interaction with another protein segment has a rather negligible influence on CD shape, if compared to the factors discussed above, such as the hydration, flexibility, and length dependence. On the other hand, it may contribute to the overall CD intensity. Similarly as for the length dependence, it can be expected that the actual effects of other intermolecular interactions in solution will be smaller than predicted by the vacuum CIS computation, due to the motional and solvent averaging discussed above.

CONCLUSIONS

We have employed several computational models to better understand various factors contributing to the CD spectra of α-helical peptides. Although the computational accuracy was limited by the large size of the systems, the results convincingly indicated that the quantum chemical techniques can capture the most important effects, including peptide dynamics, hydration, and length dependence. Yet a quantitative modeling of CD peptide spectra remains a challenge for the future.

The conformational fluctuations and hydration caused significant changes in the UV absorption and CD intensities. Therefore, a large number of geometries are needed for the adequate averaging in the combined MM/QM approach. The TDDFT method yields similar dependence of the spectra on the peptide length as the simplified models relying on empirical parameters. Exciton-like electronic states delocalized over more amide groups are needed to reproduce the typical α-helical CD profile, although the characteristic shape can be discerned already for relatively short sequences. The predicted length dependence was consistent with the experimental spectra for the short peptides containing two histidine residues, with the α-helix stabilized by the cobalt ion. A limited influence of the side chains and other protein parts were predicted by the calculations. The TDDFT results were nevertheless found to be extremely sensitive to the basis set and the solvent model. This sensitivity could be partially explained by the helix geometry leading to the nearly perpendicular orientation of the electric and magnetic transition dipoles. The highest level of theory (6-311++G**/CPCM) provided the most realistic spectral shapes. CD spectral shape agreeing more with the experiment, in particular in the $\pi-\pi^*$ transition region, was currently obtained by the TDC model. On the other hand, unlike TDDFT, TDC may miss important long-range effects, as demonstrated from simulations of the UV absorption and CD spectra for α-helical dimers.

ASSOCIATED CONTENT

Supporting Information. Further computational tests and complete experimental spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: kaminskj@gmail.com; jkubelka@uwyo.edu; bour@uochb.cas.cz.

ACKNOWLEDGMENT

This work was supported by the Grant Agency of the Academy of Sciences (M200550902, IAA400550702), the Grant Agency of the Czech Republic (P208/10/P356, P208/11/0105), National Science Foundation (CAREER 0846140), Luna, Metacentrum, and University of Tromsø computer facilities.

REFERENCES

(1) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(2) Drenth, J. *Principles of Protein X-ray Crystallography*; Springer-Verlag: Berlin, 1994.
(3) Sreerama, N.; Woody, R. W. Circular dichroism of peptides and proteins. In *Circular Dichroism Principles and Applications*; Nakanishi, K., Berova, N., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; p 601.
(4) Johnson, W. C. Analyzing the CD of proteins for secondary structure. In *Molecular Chirality in Chemistry and Life Sciences*; CNR: Pisa, Italy, 1997; Vol. Book of abstracts, p 15.
(5) Greenfield, N. J. *Anal. Biochem.* **1996**, *235*, 1.
(6) Snatzke, G. Circular dichroism: An introduction. In *Circular Dichroism*; Nakanishi, K., Berova, N., Woody, R. W., Eds.; VCH: New York, 1994; p 1.
(7) Andersen, N. H.; Liu, Z.; Prickett, K. S. *FEBS Lett.* **1996**, *399*, 47.
(8) Sudha, T. S.; Vijayakumar, E. K. S.; Balaram, P. *Int. J. Pept. Protein Res.* **2009**, *22*, 464.
(9) Dukor, R. K.; Keiderling, T. A. *Biopolymers* **1991**, *31*, 1747.
(10) Stapley, B. J.; Creamer, T. P. *Protein Sci.* **1999**, *8*, 587.
(11) Naganathan, A. N.; Doshi, U.; Fung, A.; Sadgi, M.; Muñoz, V. *Biochemistry* **2006**, *45*, 8466.
(12) Chin, D. H.; Woody, R. W.; Rohl, C. A.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *11*, 15416.
(13) Dang, Z.; Hirst, J. D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3619.
(14) Kirkwood, J. G. *J. Chem. Phys.* **1937**, *5*, 479.
(15) Schellman, J. A. *Acc. Chem. Res.* **1968**, *1*, 144.
(16) Condon, E. U.; Altar, W.; Eyring, H. *J. Chem. Phys.* **1937**, *5*, 753.
(17) Rosenfeld, L. *Z. Phys.* **1928**, *52*, 161.
(18) Fitts, D. D.; Kirkwood, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **1956**, *42*, 33.
(19) Moffit, W. *J. Chem. Phys.* **1956**, *25*, 467.
(20) Moffit, W. *Proc. Natl. Acad. Sci. U.S.A.* **1956**, *42*, 736.
(21) Kendrew, J. C.; Dickerson, R. E.; Strandberg, B. E.; Hart, K. G.; Davies, D. R.; Phillips, D. C.; Shore, V. C. *Nature* **1960**, *185*, 422.
(22) Specht, K. M.; Nam, J.; Ho, D. M.; Berova, N.; Kondru, R. K.; Beratan, D. N.; Wipf, P.; Pascal, R. A.; Kahne, D. *J. Am. Chem. Soc.* **2001**, *123*, 8961.
(23) Butz, P.; Tranter, G. E.; Simmons, J. P. *Phys. Chem. Commun.* **2002**, *5*, 91.
(24) Furche, F.; Ahlrichs, R.; Wachsmann, C.; Weber, E.; Sobanski, A.; Vögtle, F.; Grimme, S. *J. Am. Chem. Soc.* **2000**, *122*, 1717.
(25) Bouř, P. *Chem. Phys. Lett.* **2001**, *345*, 331.
(26) Šebek, J.; Kejík, Z.; Bouř, P. *J. Phys. Chem. A* **2006**, *110*, 4702.
(27) Jiang, J.; Abramavisius, D.; Bulheller, B. M.; Hirst, D. M.; Mukamel, S. *J. Phys. Chem. B* **2010**, *114*, 8270.
(28) Bayley, P. M.; Nielsen, E. B.; Schellman, J. A. *J. Phys. Chem.* **1969**, *73*, 228.
(29) Woody, R. W. *J. Chem. Phys.* **1968**, *49*, 4797.
(30) Besley, N. A.; Hirst, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 936.
(31) Gilbert, A. T. B.; Hirst, J. D. *J. Mol. Struct.: THEOCHEM* **2004**, *675*, 53.
(32) Rogers, D. M.; Hirst, J. D. *Chirality* **2004**, *16*, 234.
(33) Hirst, J. D.; Colella, K.; Gilbert, A. T. B. *J. Phys. Chem. B* **2003**, *107*, 11813.
(34) Woody, R. W. *Monatsh. Chem.* **2005**, *136*, 347.
(35) DeVoe, H. *J. Chem. Phys.* **1964**, *41*, 393.
(36) DeVoe, H. *J. Chem. Phys.* **1965**, *43*, 3199.
(37) Applequist, J. *J. Chem. Phys.* **1979**, *71*, 4332.

(38) Lowe, S. L.; Pandey, R. R.; Czlapinski, J.; Kie-Adams, G.; Hoffmann, M. R.; Thomasson, K. A.; Pierce, K. S. *J. Pept. Res.* **2002**, *61*, 189.
(39) Bulheller, B. M.; Hirst, D. M. *Bioinformatics* **2009**, *25*, 539.
(40) Fujimoto, K. J. *J. Chem. Phys.* **2010**, *133*, 124101.
(41) Dračínský, M.; Bouř, P. *J. Chem. Theory Comput.* **2010**, *6*, 288.
(42) Bouř, P.; Maloň, P. *MCM Molecular Graphics*; Academy of Sciences: Prague, 1995−2009.
(43) Richardson, J. S.; Richardson, D. C. Principles and patterns of protein conformation. In *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G. D., Ed.; Plenum: New York, 1989; p 1.
(44) Ahlrichs, R.; Bar, M.; Baron, H.-P.; Bauernschmitt, R.; Bocker, S.; Ehrig, M.; Eichkorn, K.; Elliot, S.; Furche, F.; Haase, F.; Haser, M.; Horn, H.; Huber, C.; Huniar, U.; Kattannek, M.; Kolmel, C.; Koolwitz, M.; May, K.; Ochsenfeld, C.; Ohm, H.; Schafer, A.; Schneider, U.; Treutler, O.; von Arnim, M.; Weigend, F.; Weis, P.; Weiss, H. *Turbomole*, version 5; Quantum Chemistry Group, University of Karlsruhe: Karlsruhe, 1998.
(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. J. A.; 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 C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(46) Furche, F.; Alhrichs, R. *J. Chem. Phys.* **2004**, *121*, 12772.
(47) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *J. Phys. Chem.* **1992**, *96*, 135.
(48) Klamt, A.; Jonas, V.; Burger, T.; Lohrentz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074.
(49) Cossi, M.; Barone, V. *J. Chem. Phys.* **1998**, *109*, 6246.
(50) Ponder, J. W. *Tinker*, Software Tools for Molecular Design, 4.2 ed.; Washington University School of Medicine: Saint Louis, 2000.
(51) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(52) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
(53) Holzwarth, G.; Chabay, I. *J. Chem. Phys.* **1972**, *57*, 1632.
(54) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098.
(55) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995.
(56) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113.
(57) Krantz, B. A.; Sosnick, T. R. *Nat. Struct. Biol.* **2001**, *8*, 1042.
(58) Caricato, M.; Mennucci, B.; Tomasi, J.; Ingrosso, F.; Cammi, R.; Corni, S.; Scalmani, G. *J. Chem. Phys.* **2006**, *124*, 124520.
(59) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999.
(60) Dračínský, M.; Kaminský, J.; Bouř, P. *J. Phys. Chem. B* **2009**, 14698.
(61) Šebek, J.; Kapitán, J.; Šebestík, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2009**, *113*, 7760.
(62) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126.
(63) Klamt, A. COSMO and COSMO-RS. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley and Sons: Chichester, 1998; Vol. 1, p 604.
(64) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman and Co.: New York, 1993.
(65) Fairman, R.; Chao, H. G.; Mueller, L. J.; Lavoie, T. B.; Shen, L.; Novotny, J.; Matsueda, G. R. *Protein Sci.* **1995**, *4*, 1457.
(66) Bouř, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45.
(67) Cristancho, D.; Seminario, J. M. *J. Chem. Phys.* **2010**, *132*, 065102.
(68) Filippi, C.; Umrigar, C. J.; Gonze, X. *J. Chem. Phys.* **1997**, *107*, 9994.
(69) Lehtonen, O.; Sundholm, D.; Send, R.; Johansson, M. P. *J. Chem. Phys.* **2009**, *131*, 024301.
(70) Skála, L. P.; Pančoška, P. *Chem. Phys.* **1988**, *125*, 21.
(71) Fasman, G. D. *Prediction of Protein Structure and the Principles of Protein Conformation*; Plenum: New York, 1989.
(72) Volosov, A. *Int. J. Quantum Chem.* **1988**, *34*, 471.
(73) Andrushchenko, V.; Bouř, P. *J. Comput. Chem.* **2008**, *29*, 2693.
(74) Lerman, L. S. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 1886.
(75) Amunson, K.; Ackels, L.; Kubelka, J. *J. Am. Chem. Soc.* **2008**, *130*, 8146.