JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Calculation of the CD Spectrum of a Peptide from Its Conformational Phase Space: The Case of Met-enkephalin and Its Unnatural Analogue

Zlatko Brkljača,[†] Karmen Čondić-Jurkić,[†,‡] Ana-Sunčana Smith,*[,†] and David M. Smith*[,‡,§]

[†]Institute for Theoretical Physics and Excellence Cluster: Engineering of Advanced Materials, FA University Erlangen—Nürnberg, Nägelsbachstraße 49b, Erlangen, 91052, Germany

[‡]Division of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička 54, 10000, Zagreb, Croatia

[§]Center for Computational Chemistry, FA University of Erlangen—Nürnberg, Nägelsbachstraße 25, Erlangen, 91052, Germany

Supporting Information

ABSTRACT: We have investigated the conformational phase spaces of both Met-enkephalin and Ada-enkephalin in 2,2,2-trifluoroethanol in order to connect them to their respective CD spectra. To this end, we have characterized the conformational preferences of the zwitterionic and neutral forms of Met-enkephalin and of both the *R*- and the *S*-epimers of Ada-enkephalin, as obtained by classical molecular dynamics. The CD spectrum for each peptide was subsequently obtained with a procedure of successive averaging, which accounts for the behavior of the solvent, the side chains, and the backbone variations of the peptides. To make an appropriate comparison with experimental results, we have produced composite spectra that account for the appropriate contributions of the zwitterionic and neutral forms of the peptides as well as the expected epimeric ratio. Such a procedure results in theoretically obtained CD spectra that show significant promise in terms of reproducing their experimentally measured counterparts.

INTRODUCTION

Opioid peptides belong to a group of biologically relevant signaling molecules that inhibit neurotransmitter release upon activation of the appropriate opioid receptor.[1] These peptides are known to play a significant role in regulating dopamine release[2] and drug-induced reward[3] as well as in pain mediation, opiate dependence, and euphoria.[2] Possible keys to their activity are the aromatic residues found in positions 1 and 4 in their sequences.[2]

The most ubiquitous opiod peptide is Met-enkephalin (Tyr-Gly-Gly-Phe-Met, Met, Figure 1), also known as the opioid growth factor (OGF). Met-enkephalin plays a major role in cell proliferation and tissue organization during development, as well as in cellular renewal, wound healing, and angiogenesis. However, particularly interesting is its capacity to inhibit tumor-cell division, which is done in a receptor-mediated fashion.[4−7] While it is known that Met-enkephalin binds to three classical receptors, namely, the μ, κ, and δ receptors, it was shown that ligands selective for these three receptors do not affect cell growth either *in vitro* or *in vivo*.[4] Subsequently, it was found that the Met-enkephalin also binds to an OGF-receptor (OGFr) positioned on the nuclear envelope.[8] OGFr, previously known as the ζ-receptor, has no significant homologies to classical opioid receptors. Nevertheless, it is precisely the Met-enkephalin-OGFr axis that was found to be responsible for the antitumor activity.[6−8] This implies that Met-enkephalin must be transported through the cell membrane in order to interact with the OGFr.

Following these discoveries, various unnatural analogues of Met-enkephalin-containing lipophilic moieties were synthesized[4] and tested for improved or novel activity profiles toward malignant diseases. The analogues either had backbone modifications or were covalently conjugated with a lipid or a polymeric tail. However, analogues in which both tyrosine and phenylalanine residues had undergone modifications expressed smaller affinity for the OGFr and in turn had diminished antitumor activity.[8] Other types of substitutions promoted the passive or active cellular uptake of peptides without greatly affecting the affinity, thus resulting in increased activity.[4,5]

An interesting group of peptides in this respect are those containing the adamantyl functional group. For example, the incorporation of (*R*,*S*)-(1-adamantyl)glycine (Aaa) in position 2 of the Met sequence results in Ada-enkephalin (Tyr-Aaa-Gly-Phe-Met, Ada, Figure 1).[4] This peptide showed a 10-fold increase in the antitumor activity when compared with Met-enkephalin in the case of HEp-2 cells (human larynx carcinoma) and a 2-fold increase with SW-620 tumor cells (colon carcinoma).[4] This efficiency may arise from the capability of the hydrophobic Aaa to induce both β-turn and/or γ-turn[9] spatial conformations, which then increase the hydrophobicity of Ada-enkephalin and its membrane permeability.[10] These results indicate that the limiting step in antitumor activity may be the diffusion of peptides through the membrane, which in turn may be related to the structure of peptides in nonpolar environments.

The structure of Met-enkephalin has been studied both in water and in nonpolar solvents. In water, both NMR results and molecular dynamics simulations show Met-enkephalin to possess an essentially random distribution of conformers, with no unique native structure.[2] On the other hand, NMR experiments in binary bilayered mixed micelles showed relatively well-defined structures.[11]

Received: December 2, 2011
Published: April 2, 2012

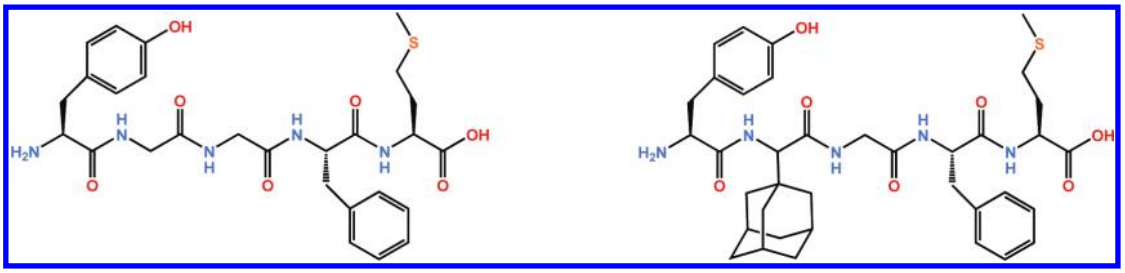


Figure 1. Met-enkephalin (Met, left) and Ada-enkephalin (Ada, right).

These latter results, however, have not thus far been reproduced theoretically.

As an alternative to NMR, information about the secondary structures of proteins and peptides can be obtained by measuring circular dichroism (CD) spectra. In the case of Met-enkephalin and Ada-enkephalin, such experiments were performed in the low-polarity solvent 2,2,2-trifluoroethanol (TFE). The spectrum of Ada-enkephalin displayed a deep negative minimum at 200 nm, which was assigned as representing a high population of ordered conformers, with likely folded backbones. Met-enkephalin was assumed to remain unfolded because its spectrum remained positive for all wavelengths between 190 and 240 nm.[5] Such conclusions support the hypothesis that the antiproliferative activity is connected to the folding ability of investigated peptides.[5]

Ada-enkephalin has two epimers, for which different antitumor capacities have been measured. The *S*-epimer (containing *S*-Aaa) penetrates rapidly to the cellular interior, while the *R*-epimer (containing *R*-Aaa) does not even bind to the outer membrane and therefore does not enter the cell.[4] This points to possible differences in the physical and structural properties of the epimers. However, the discussed CD measurements of Ada-enkephalin were conducted on a 1:1 mixture of the two epimers. Hence, the origin of the dominant minimum in the spectrum cannot be clearly associated with certain structural preference. Furthermore, the side-chain chromophores present in the two aromatic residues (Tyr1 and Phe4), present in all peptides, may have as strong a contribution to the CD spectra as the backbone structure itself.[12−14]

The detailed understanding of the above results requires the establishment of a direct link between the conformational phase space of the peptide and its CD spectrum. This task was successfully accomplished for molecules with well-defined secondary structure, after the development of semiempirical methods for the calculation of CD spectra.[15−21] However, attempts to calculate the CD spectra of small and flexible peptides by semiempirical methods[22,23] have proven to be significantly more challenging. On one hand, peptides typically have an ill-defined secondary structure yielding a very wide phase space. Since the spectrum emerges as an ensemble average, the capacity to sample and manipulate this wide phase space significantly affects the result. On the other hand, the entire peptide is typically exposed to the solvent. Even if the individual solvent molecules are not chiral, they will surround the chiral peptides in an asymmetric manner and thus contribute to the CD spectra.[24] This subtle but strong effect appears very challenging for semiempirical methods.[20] Hence, TDDFT calculations become a method of choice for the CD spectra of small molecules. Indeed, experimental CD spectra of short constrained α-helices were reproduced successfully with these more sophisticated methods, after taking the solvent effects into account.[25]

In the context of comparisons with experiments, an additional difficulty is the potential coexistence of the zwitterionic and neutral forms of the peptide. These forms may have very different phase spaces but will both contribute to the measured spectrum with statistical weights given by their concentration ratio, which is often not known. As a consequence, the experimental CD spectra of highly flexible peptides consisting of both aromatic and nonaromatic constituents have not yet been successfully reproduced by theoretical means.

In this work, we aim to address the described difficulties by combining a number of state-of-the-art computational methods. We first investigate the conformational phase spaces of Met-enkephalin and the Ada-enkephalin epimers in TFE, in both the neutral and zwitterionic forms. Inspired by previous work on this topic on molecules in a constrained conformation, we develop a procedure to account for solvent effects for flexible molecules. On the basis of these phase spaces, we calculate the CD spectra by means of TDDFT calculations. After estimating the concentration ratio of the two forms, we construct the analogue of measured CD spectra of these two peptides. Our methodology is summarized in a flowchart shown in the Supporting Information (Scheme S1). Our results show that even though the errors of the calculations are not negligible, the main features of the experimental CD spectra are reasonably well reproduced in our calculations. Consequently, the link between the CD spectra and the conformational phase space of flexible peptides can be established for the first time.

METHODS

Preparation. All peptides were modeled using the Duan et al. *ff03*[26] and general AMBER force fields,[27] as part of the AMBER 10 simulation package.[28] Specifically, we constructed the zwitterionic forms of (*i*) Met_Z, Met-enkephalin (TyrGlyGlyPheMet); (*ii*) *R*-Ada_Z, the *R*-epimer of Ada-enkephalin (Tyr*R*-AaaGlyPheMet); and (*iii*) *S*-Ada_Z, the *S*-epimer (Tyr*S*-AaaGlyPheMet). Furthermore, we considered the neutral forms of (*iv*) Met_N (Tyr_0GlyGlyPheMet_0), (*v*) *R*-Ada_N (Tyr_0*R*-AaaGlyPheMet_0), and (*vi*) *S*-Ada_N (Tyr_0S-AaaGlyPheMet_0). Here, Tyr_0 and Met_0 represent the neutral forms of the N- and C-terminal tyrosine and methionine residues, respectively.

The missing parameters for Tyr_0, Met_0, and Aaa amino acids were obtained in a manner consistent with the *ff03* force field. Specifically, the charges were obtained from a restrained (RESP) two-conformer (extended and α helix) fit to the electrostatic potential. The potential was obtained from the B3LYP/cc-pVTZ//HF/6-31G(d,p) quantum mechanical method with the IEFPCM (ε = 4.335) model representing a (low-polarity) polarizable organic continuum.[26] Prior to calculating the electrostatic potential, all structures were subject to constrained optimizations using the HF/6-31G(d,p) level of theory. The dihedral angles (Φ,Ψ) were fixed at (−60°, −40°) and (−120°, 140°) for the α helix and the extended conformations, respectively. All aforementioned calculations were performed using the Gaussian 03 package.[29]

Starting from an initially linear conformation, the initial folded structures for the dynamics with explicit solvent were obtained through a combination of minimizations and dynamics simulations using the generalized Born representation of implicit solvent available in AMBER. All subsequent explicit simulations were performed with periodic boundary conditions, whereas the long-range electrostatic interactions beyond an 8 Å cutoff were calculated using the particle mesh Ewald method.[28]

A cubic box containing 512 molecules of TFE was built using Pacmol.[30] Parameters for the TFE molecule were obtained from the R.E.DD.B. (RESP and ESP database,[31] parameter code name W-15). The initial size of the box was chosen such that the density was 1.393 g/cm^3, which is typical for TFE at T = 300 K. The resulting box was equilibrated under NPT conditions and used as a template to solvate the folded peptides in regular truncated octahedron unit cells. All forms of Met and Ada were solvated with 226 and 347 molecules of TFE (box sizes were 37.4 Å and 42.6 Å), respectively.

The minimization of this complete periodic system consisted of several steps: (*i*) minimization with the fully restrained solvent, (*ii*) 50 ps of Langevin dynamics, in the NVT ensemble at 300 K, with the collision frequency of 1 ps^{-1} and a time step of 2 fs, restraining both the peptide and the solvent, (*iii*) minimization with reduced positional restraints on the solvent and the peptide, (*iv*) minimization with no positional restraints, (*v*) a further 50 ps of Langevin dynamics at 300 K (NVT ensemble) with no positional restraints. Finally, (*vi*) to equilibrate the density to its stable value of 1.37 g/cm^3, a final 50 ps Langevin dynamics simulation was performed at 300 K with a constant pressure of 1 atm (NPT).

REMD Production. The configuration space of each peptide was generated by combining two replica exchange molecular dynamics (REMD) simulations performed in AMBER 10.[28] The starting conformations of the second run were chosen to be as different as possible from that of the first run. In each run, 16 replicas of the fully equilibrated system were set at different temperatures (T = 275, 283, 291, 299, 307, 316, 326, 335, 345, 355, 365, 375, 386, 397, 408, 420 K), with the latter chosen such that the exchange probability between replicas with adjacent temperatures was about 20%. The exchanges were attempted every 500 steps. In between, all replicas were subject to Langevin dynamics in the NVT ensemble with a 2 fs time step and a collision frequency of 1 ps^{-1}. The 16 replicas were propagated for a total of 176 ns (11 ns per replica). The coordinates of the entire system were saved every 0.5 ps. The first 2.5 ns were omitted from the analysis. Consequently, 17 000 structures from replicas at T = 299 K were extracted from each REMD run, giving rise to a total of 34 000 conformations comprising the conformational space of each of the six peptides.

Aspects of the convergence of the REMD simulations, including the structural ensembles (Figure S1),[32] the temperature sampling (Figure S2), and the success rate of the temperature exchanges (Table S1), are addressed in the Supporting Information.

Principal Component Analysis (PCA) and Clustering. Principal component analysis was performed using the PCA tool from the AmberTools software package.[33] The conformational spaces were projected on the first two principal components determined on the basis of the positions of the α-carbon atoms of the peptide backbone. Thereby, the PCA analysis was performed simultaneously on all six conformational spaces. Such a procedure permits the construction of a collective conformational phase space for all of the investigated peptides. Subsequent projections were performed on the first two principal components (PC1 and PC2).

The clustering of the conformations was performed using the *K*-means algorithm,[34] also from AmberTools.[33] In the analysis of the conformational phase space, the α-carbon atoms were again used, whereas for CD spectra calculations, further subclustering was performed on the basis of all atoms. Representative structures for each cluster were determined as the structure with the lowest RMS deviation from the relevant centroid.

Circular Dichroism (CD) Calculations. Using the TDDFT implementation available in Gaussian 09,[35] it is possible to calculate the excitation energies and rotatory strengths for the electronic transitions of the investigated peptides. The CD spectra can be constructed as a convolution of overlapping Gaussian functions, each representing an electronic transition:

$$\Delta\varepsilon(E) = \frac{1}{2.297 \times 10^{-39}} \frac{1}{\sqrt{2\pi\sigma^2}} \times \sum_i^A \Delta E_i R_i \, e^{-(E-\Delta E_i)^2/2\sigma^2}$$

Here, A represents the total number of calculated excited states and σ is the width of the Gaussian functions. ΔE_i and R_i are the excitation energies and rotatory strengths for the transition i, respectively.[36] In calculations involving individually and explicitly solvated structures, we chose A = 72 and σ = 0.15 eV. In calculations involving the averaged solvent charge field (see below), we used values of A = 100 and σ = 0.19 eV.

CD spectra were calculated using a QM/MM approach, in which the entire peptide was treated quantum mechanically (TDDFT), whereas the effect of the solvent was included through the introduction of partial charges at the positions of the solvent nuclei, as obtained from the molecular dynamics simulations. The spectra were obtained with the B3LYP functional, which unless otherwise specified was used in conjunction with the 6-31G(d) basis set. The convergence testing in Figures 2 and 3 was performed with the STO-3G basis set, whereas the 6-311G(d) basis set was used to demonstrate that the default choice of 6-31G(d) was sufficient to obtain converged spectra (see Figure S4 in the Supporting Information).

Effects of the Solvent on a Single Conformation. The effect of differing solvent configurations on the CD spectra is clearly demonstrated in Figure 2a. Therein, each spectrum corresponds to the same conformation of *S*-Ada$_N$, each with a unique solvent configuration. This result confirms the marked sensitivity of both ΔE_i and R_i to the presence of an asymmetric solvent structure.[24] Clearly, in order to obtain a reliable and reproducible CD spectrum for a single peptide conformation, one needs to average over many configurations of the solvent.[37] In the current application, we generated the required configurations in the following manner. First, we extracted and minimized a selected snapshot from the REMD trajectory. The resulting configuration was then subjected to constant temperature molecular dynamics (NVT) with strong positional restraints on the peptide. After an equilibration period, the solvent coordinates were extracted every 10 ps and, after alignment with respect to the minimized peptide conformation, used as a set of independent solvent configurations.

The CD spectrum of the selected peptide conformation was then calculated under the influence of each of the obtained solvent configurations (Figure 2a). The resulting spectra were combined to give the solvent-averaged spectrum for the relevant peptide conformation. The convergence of such an average is shown in Figure 2b using 5, 10, 30, 50, 80, and 100 solvent configurations.

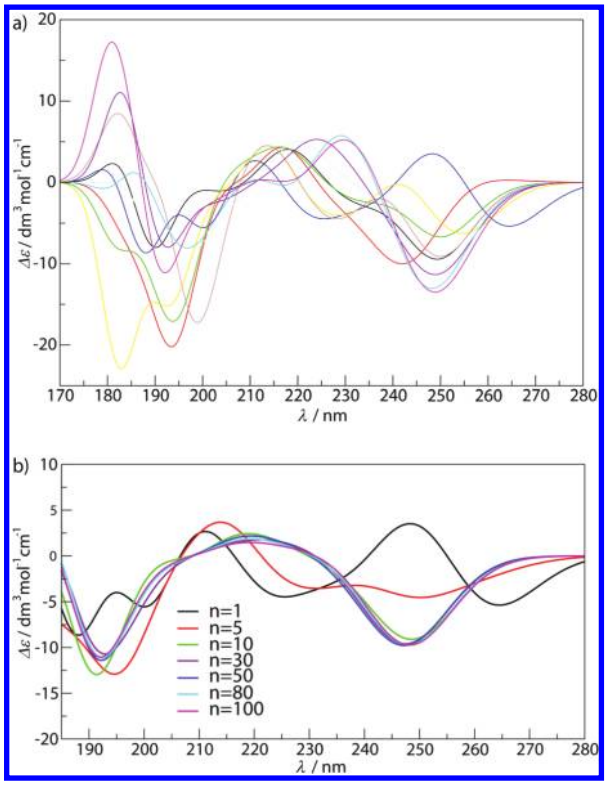


Figure 2. (a) CD spectra of a single *S*-Ada$_N$ conformation with different configurations of TFE. (b) Averaged spectra after finding a mean of *n* spectra shown in the upper panel. To demonstrate that this convergence persists across different structures of different peptides, we have included the corresponding data for a closed structure of Met$_Z$ in Figure S3 of the Supporting Information.

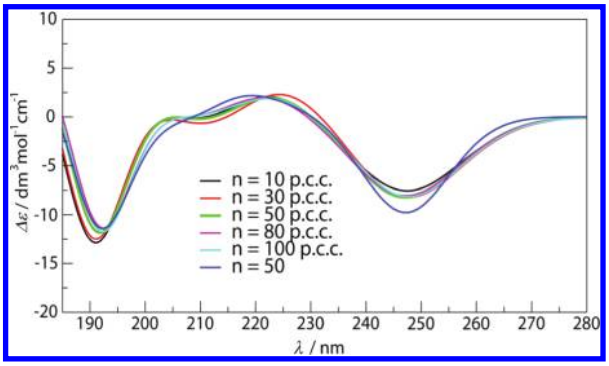


Figure 3. The same *S*-Ada$_N$ conformation from Figure 2 is used to test the convergence of CD spectra when the ASEC approach is applied. The convergence of the CD spectrum again occurs when 50 point-charge configurations (p.c.c.) are employed in order to build the ASEC. For comparison, the spectrum obtained by averaging over 50 simulations (Figure 2b, blue curve) is also plotted (blue curve).

It can be seen that while reasonable convergence can already be achieved with 30 solvent configurations, the use of 50 is preferable. Such a procedure is, however, prohibitively expensive for general use, since many peptide conformations would need to be treated in this way in order to obtain a CD spectrum representative of the overall structural ensemble (see later).

To render an adequate treatment of the solvent more feasible, we have employed an approach in which we first combined the distinct solvent configurations and then evaluated their collective effect on the CD spectrum. Specifically, by superposing *n* solvent configurations (obtained as per the description above), each with its partial charges scaled by *n*, we generated a single averaged solvent configuration. The CD spectrum for the relevant peptide configuration was then calculated under the influence of this single configuration. Such a procedure can be considered as a simplified version of the Averaged Solvent Electrostatic Potential (ASEP) developed by Aguilar and co-workers,[38] which closely resembles the Average Solvent Electrostatic Configuration (ASEC) approach suggested by Coutinho et al.[39] The resulting dependence of the spectrum on the number of solvent configurations used to construct the average configuration is shown in Figure 3. The comparison with the stepwise calculation (averaging over 50 spectra as in Figure 2b) shows that excellent agreement is obtained after including the contribution of 50 solvent configurations to the superposed average.[40]

To summarize the procedure for building the representative set of structures for the construction of the CD spectra, we have included a flowchart (Supporting Information, Scheme S1), which outlines the main steps involved and our choices of methods and parameters.

CONFORMATIONAL PHASE SPACE OF PEPTIDES

In order to explore the conformational phase spaces in a comparative fashion, all structures from all three forms of both peptides, associated with a temperature of 299 K, were combined into a single set of structures. This set was used to determine a common set of principal components, based on the positions of α-carbon atoms (see Methods). Each structure was then projected onto the first two principal components, which upon visual inspection could be associated with the end-to-end distance and the twist of the backbone, respectively. In this way, 2-D plots, where each conformation is represented by a single data point, were constructed for each peptide simulated. The associated free energy surfaces were then calculated from the local density of points in this representation. In addition, the trajectories of each peptide were clustered, also on the basis of their α-carbon positions, so as to typically produce five clusters. Exceptionally, if such a procedure would yield substantial overlaps in the space of the first two principal components, only four clusters were constructed.

Zwitterionic Forms. The results of the described procedure for the zwitterionic peptides confirm the somewhat intuitive expectation that Met$_Z$ has a less structured phase space than either *R*-Ada$_Z$ or *S*-Ada$_Z$ (Figure 4). Indeed, one can observe that an entire section of the phase space is inaccessible to the two Ada$_Z$ variants. This is presumably because the Aaa group cannot adjust favorably to the semiopen conformation of the backbone, typical for cluster III of Met$_Z$ (Figure 5). The more surprising difference, however, is that observed between the phase spaces of *S*-Ada$_Z$ and *R*-Ada$_Z$, with the latter being considerably more constrained. Interestingly, this effect emerges from differences in how the direct Coulomb interaction between the charged ends of the peptides, which is not screened efficiently by the low-polarity solvent, competes with the twisting of the backbone. The former phenomenon seems to have a dominant effect on the conformational phase space of all three zwitterionic forms. Consequently, the completely open structures, typical for clusters I and II, are scarce in comparison to folded (U-shaped) conformations (clusters IV and V) that together comprise about 70%, 80%, and 95% of all conformations of Met$_Z$, *S*-Ada$_Z$, and *R*-Ada$_Z$, respectively (see Figure 5 for details on cluster populations and representative structures).

From the point of view of the 2-D representations of the free energy, all folded structures appear at values of PC1 > 1, where

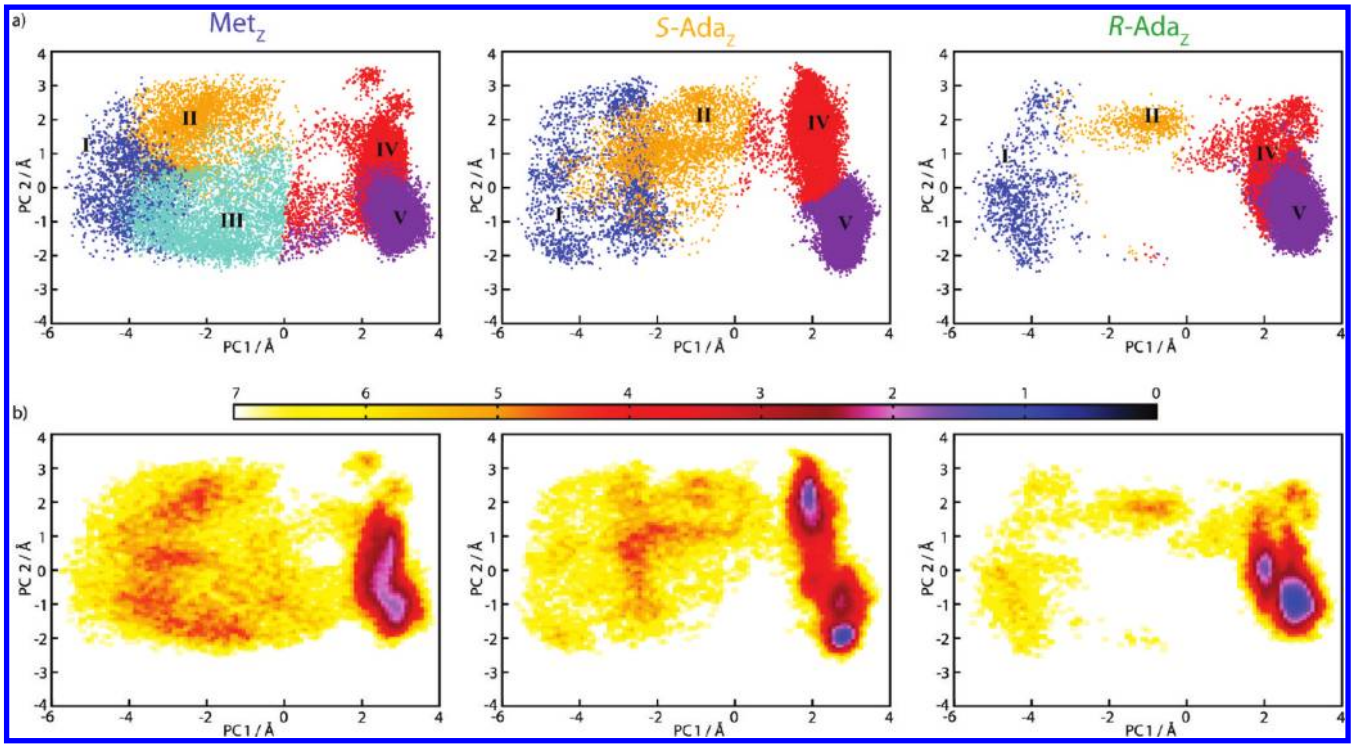


Figure 4. (a) Conformational phase space of Met$_Z$, *S*-Ada$_Z$, and *R*-Ada$_Z$, respectively, at T = 299 K, after clustering. Each cluster is represented by a Roman numeral. (b) Free energy surfaces for Met$_Z$, *S*-Ada$_Z$, and *R*-Ada$_Z$, respectively. The free energy surface is given by $W/kT = -\ln(N/N_{tot})$ where N_{tot} represents the total number of configurations of each individual peptide (N_{tot} = 34 000).

Peptide	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V
Met$_Z$	5.71	9.71	12.5	41.73	30.36
S-Ada$_Z$	7.51	9.95	-	41.90	38.78
R-Ada$_Z$	2.87	1.98	-	28.53	66.63

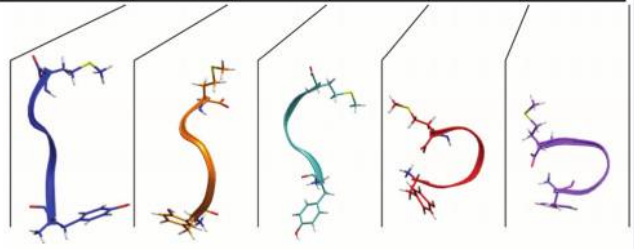


Figure 5. Cluster population percentages of Met$_Z$, *S*-Ada$_Z$, and *R*-Ada$_Z$. Ribbon renderings of five representative Met$_Z$ conformations (conformations most similar to the centroids of their respective clusters) are shown. Conformations were plotted using the VMD visualization software.[40]

the clustering algorithm found two distinct clusters (IV and V). Indeed, two separate deep minima are observed for both epimers of Ada$_Z$, albeit at different positions in the PC2 coordinate (Figure 4b). In the case of Met$_Z$, the two minima seem to merge into a single, considerably shallower and wider minimum. Closer inspection of the representative structures of clusters IV and V (Figures 6 and 7, respectively) shows that structures in cluster IV are dominated by the electrostatically driven formation of a salt bridge between the N- and C-termini of all three peptides. In the case of cluster V, the dominance of this bridge is somewhat reduced, and all three representative structures have a β-turn hydrogen bond between Phe4 and Tyr1.

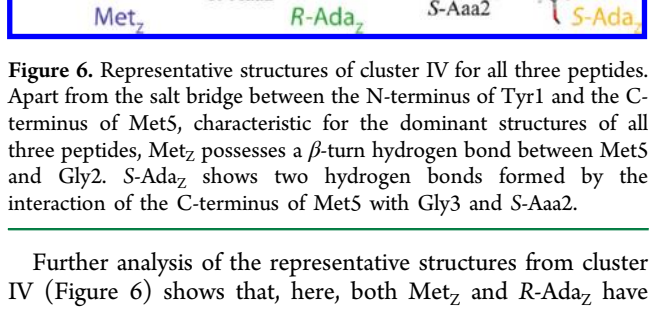


Figure 6. Representative structures of cluster IV for all three peptides. Apart from the salt bridge between the N-terminus of Tyr1 and the C-terminus of Met5, characteristic for the dominant structures of all three peptides, Met$_Z$ possesses a β-turn hydrogen bond between Met5 and Gly2. *S*-Ada$_Z$ shows two hydrogen bonds formed by the interaction of the C-terminus of Met5 with Gly3 and *S*-Aaa2.

Further analysis of the representative structures from cluster IV (Figure 6) shows that, here, both Met$_Z$ and *R*-Ada$_Z$ have U-shaped backbones, with the upper and lower parts of the backbones well aligned. In the case of Met$_Z$, the structure is maintained by two hydrogen bonds, out of which one, present in 17% of conformations, is between Met5 and Gly2 and induces β-turns of the type II′. Due to the lack of the backbone twist, the centroids of cluster IV of Met$_Z$ and *R*-Ada$_Z$ have roughly the same projections on PC2. However, as a consequence of the two hydrogen bonds at the C-terminus, the representative structure of *S*-Ada$_Z$ has a more twisted backbone. Consistently, the free energy minimum and the position of the centroid in the conformational phase space are shifted to considerably larger PC2 values.

The clusters V can be described, for Met$_Z$ and *R*-Ada$_Z$, as nearly circular islands occupying practically identical regions of the conformational phase space. They are both occupied with β-turns of the type II′ with about 50% of both Met$_Z$ and *R*-Ada$_Z$ structures possessing the identical β-turn hydrogen bond

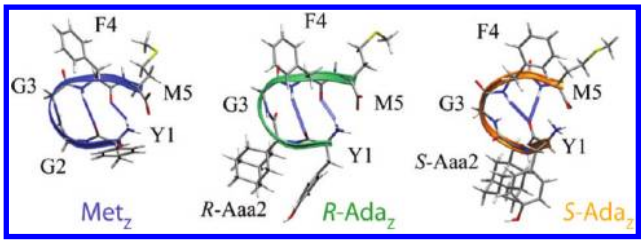


Figure 7. Representative structures of cluster V for all three peptides. Dominant structures of Met$_Z$ and *R*-Ada$_Z$ have a strong hydrogen bond between N-terminus of Tyr1 and Phe4 as well as a β-turn hydrogen bond between Phe4 and Tyr1. This same hydrogen bond exists also in the representative structure of *S*-Ada$_Z$ but is bifurcated with a second bond from Met5 to Tyr1.

between Phe4 and Tyr1, and a very similar backbone appearance (Figure 7). On the other hand, *S*-Ada$_Z$ possesses a bifurcated hydrogen bond of the $i + 4 \rightarrow i$ type, characteristic for α-helical structures, placing the free energy minimum lower in terms of its PC2 projection. However, the strong electrostatic interaction between the charged end groups prevents the formation of a pure α-helix structure. As a result, the conformations of *S*-Ada$_Z$ form mainly structures that resemble β-turns of type II.

Neutral Forms. The lack of the charged end groups in the neutral forms of the peptides results in more evenly populated conformational spaces. In addition, some new parts of the phase space, that are completely unoccupied by the zwitterionic forms become apparent. In the case of Met$_N$, all five clusters have similar populations (Table 1) and occupy areas of similar sizes in the

Table 1. Population Percentages for the Clusters Obtained for the Neutral Structures

peptide	cluster I	cluster II	cluster III	cluster IV	cluster V
Met$_N$	17.83	21.97	28.69	19.02	12.49
S-Ada$_N$	56.96	25.86	5.06	6.75	5.37
R-Ada$_N$	33.85	7.68	34.86	14.86	8.75

phase space (Figure 8). Consequently, the free energy plot of Met$_N$ contains a number of connected shallow minima. One of these can be associated with the subcluster Va in which helical structures of the 3_{10} type are most common (Figure 9a).

S-Ada$_N$ and *R*-Ada$_N$ are also found to visit virtually the entire phase space. However, in both cases, more pronounced minima in the free energy surface can be observed. Specifically, about 75% of *S*-Ada$_N$ conformations are extended, which provides a free energy centered around cluster I. Moreover, like in Met$_N$, a minimum can be associated with the subcluster Va for *S*-Ada$_N$. Due to different hydrogen bonding patterns (Figure 9b), true α-helices can be found for the latter peptide. In addition to these features, inspection of the middle panel in Figure 8b shows a wide minimum at a PC2 value of about 1, stretching over all negative values of PC1, as well as a minimum with folded structures close to the centroid of cluster IV (Figure 9c). Both of these minima are significant for *S*-Ada$_N$ but are occupied to a much lesser extent for Met$_N$ and *R*-Ada$_N$.

The most dominant feature of the *R*-Ada$_N$ conformational space is a broad minimum at low values of both PC1 and PC2, to which extended and semifolded structures (typical for clusters I and III, respectively) contribute. Indeed, these two clusters account for nearly 70% of the distribution for *R*-Ada$_N$. Another interesting part of the phase space is the far upper right-hand corner of the PC plot. For both *R*-Ada$_N$ and Met$_N$, this region, which belongs to cluster IV, is populated by quasi left helices (Figure 9c). In the case of *S*-Ada$_N$, however, the formation of a proper left-helix is prevented by the *S*-Aaa residue, and this particular section of the phase space is completely inaccessible. Similarly, *R*-Ada$_N$ does not exhibit a significant population in the vicinity of the Va region. Thereby, the *R*-adamantyl residue apparently retards the formation of both α- and 3_{10}-helices.

Summary. Our simulations show that in TFE, the zwitterionic peptides show a much larger propensity for folded structures than the neutral forms, which, in turn, exhibit a considerably higher proportion of extended structures. This appears to be the result of the dominant electrostatic force acting between the charged end groups in the zwitterionic forms. Since such a force is absent in the neutral forms, they are able to occupy the phase space in a less restricted manner. This result contrasts with the situation in water, where the polar solvent successfully screens the charged termini. As a result, in water, Met$_Z$ is equivalently abundant in all accessible parts of the conformational phase space (Figure S6 in Supporting Information),[41] in a similar way as Met$_N$ is in TFE.

CIRCULAR DICHROISM CALCULATIONS

Calculating CD Spectra of a Single Cluster. Under the assumption that the conformational phase space of each particular peptide is well represented by the previously described results of the REMD simulations, we proceed with calculating the CD spectra of a single cluster (see, also, the method flowchart, Scheme S1). This will be exemplified for the well-defined cluster V of *R*-Ada$_Z$ (see Figure 7). Despite the fact that the individual structures in this cluster have similar backbone conformations, their (solvent averaged) CD spectra (calculated as discussed in the Methods section) differ appreciably (Figure 10a). This suggests that the side-chains have considerable effects on the individual CD spectra. Consequently, it is clear that a single structure cannot represent the whole cluster, but a set of the representative conformations, which will be the basis for the calculation of the CD spectra, need to be constructed. The set is obtained by subclustering the initial backbone-based cluster (obtained as described in the Methods section), by applying the *K*-means clustering algorithm, taking into account the positions of all of the atoms in the peptide.

The size of the representative set is determined by monitoring the convergence of the spectra after averaging over a number of representative structures in the set (Figure 10b). In the specific case of cluster V or *R*-Ada$_Z$, subclustering was performed to give sets with 3, 9, 18, 27, and 36 structures. The average CD spectrum of the set was constructed by weighted averaging. Thereby, the weight of each individual spectrum is given by the population fraction of the respective subcluster in the context of the whole cluster. As can be seen from Figure 10b, the spectrum begins to converge for a set of nine representative conformations, and very good results are obtained for a set with 27 characteristic conformations, each representing a subcluster with, on average, 680 conformations. While further fragmentation of the phase space is possible, this tends to produce a large number of subclusters containing only four to five conformations, with badly defined centroids.

The consistency of our approach can be tested by calculating the CD spectra of two clusters belonging to different peptides, which overlap in the collective phase space. For this purpose, we identify the cluster V of Met$_Z$ that corresponds closely to the

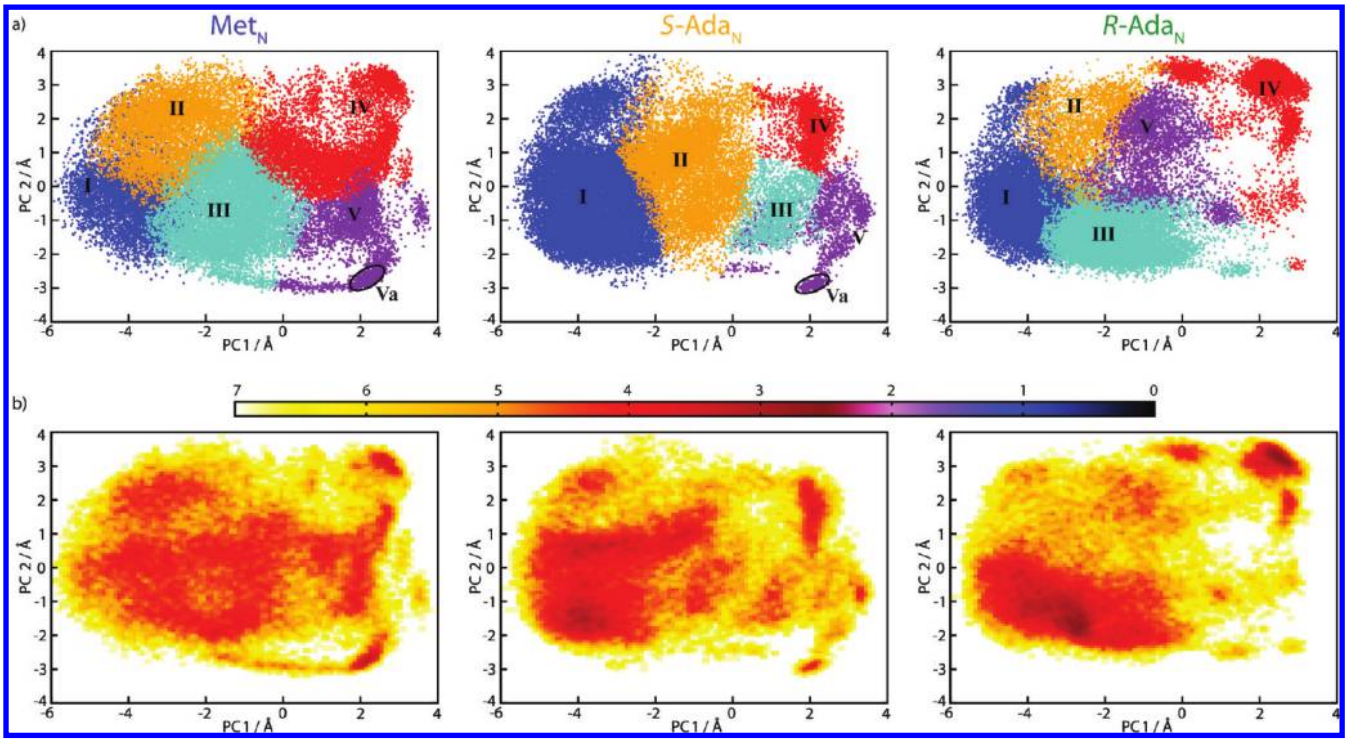


Figure 8. (a) Clustered conformational phase space spanned by 34 000 conformations and (b) the respective free energy plots of Met$_N$ (left), *S*-Ada$_N$ (middle), and *R*-Ada$_N$ (right) at T = 299 K. Each cluster is represented by a Roman numeral.

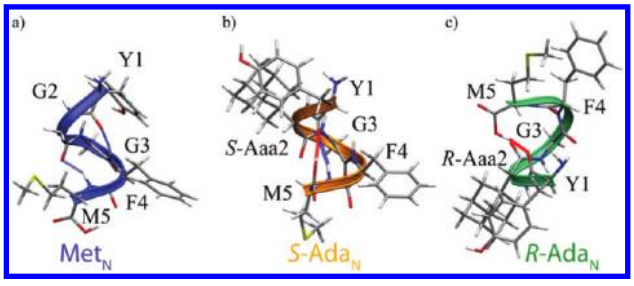


Figure 9. (a) An example of 3_{10} helical structure of Met$_N$, (b) α-helical structure of *S*-Ada$_N$, and (c) a conformation of cluster IV of *R*-Ada$_N$ that is an example of quasi left helical structures, which can be found in the upper right part of the phase space of *R*-Ada$_N$. The structures in a and b are both from the cluster Va. The Met$_N$ conformation possesses a hydrogen bond between Phe4 and Tyr1, and between Met5 and Gly3. The *S*-Ada$_N$ structure possesses a bifurcated hydrogen bond from Tyr1 to Met5, and to the C-terminal. The first bond is an α-helix bond, whereas the second further stabilizes the conformation.

cluster V in *R*-Ada$_Z$, both in terms of the positions of the centroids and in terms of very similar representative structures (Figure 7). We thus construct the CD spectrum of the cluster V of Met$_Z$ maintaining the same level of sampling discussed above (average subcluster size of 680 conformations). To obtain spectra of comparable intensity, both spectra are normalized by the total number of conformations in the respective clusters. The results of this calculation are shown in the upper panel of Figure 11, where the striking similarity between the average CD spectra of clusters V of Met$_Z$ and *R*-Ada$_Z$ is apparent.

Since cluster V is very compact and well-defined, we have extended our testing to a pair of similar but more dilute and spread clusters, which could be expected to be considerably more challenging for the method. We thus repeated the same procedure for clusters I of Met$_N$ and *R*-Ada$_N$ (lower panel in Figure 11). Again, a remarkable agreement of the two spectra was obtained.

Apart from confirming the reliability of our approach, these results show that even though structures with a very similar backbone shape can have very different spectra due to the side-chain conformations (Figure 10), this effect can be effectively averaged (Figure 11) in much the same way as the solvent can (Figures 2 and 3). Namely, the subclustering on the basis of all atoms can be regarded as sampling the side-chain conformations of structures that belong to a cluster with a well-defined backbone structure as determined by initial clustering over only the C$_\alpha$ atoms. This hierarchical sampling is very clear when the phase space is condensed (cluster V). However, in the sparse phase space, the backbone and side-chain sampling may be somewhat overlapping, without truly affecting the result (cluster I), if a sufficiently large set of representing structures is considered.

Calculating the CD Spectra of the Entire Phase Space. In order to build the CD spectra emerging from the entire phase space, we have opted to retain the above sampling frequency to arrive at an average of 680 conformations within a given subcluster. In the context of a phase space built from 34 000 structures, this amounts to a total of 50 subclusters and a corresponding set of 50 characteristic structures. The final combination of the corresponding spectra can be considered an averaging of the different backbone structures of the principal clusters (Figures 4 and 8) to produce the overall spectrum of a given peptide. Under such circumstances, we used A = 100 and σ = 0.135 eV (see Methods). To test the consistency of this methodology in its entirety, we calculated the overall CD spectrum of R_4-Ada$_N$ (*R*-Tyr-*R*-Aaa-Gly-*R*-Phe-*R*-Met), which is the enantiomer of *S*-Ada$_N$. The phase space and the CD spectrum for R_4-Ada$_N$ were generated according to the procedure outlined above. In the ideal case, due to the mirror symmetry, the resulting spectrum should be identical but opposite that of *S*-Ada$_N$. However, because of errors which

dx.doi.org/10.1021/ct200868y | J. Chem. Theory Comput. 2012, 8, 1694–1705

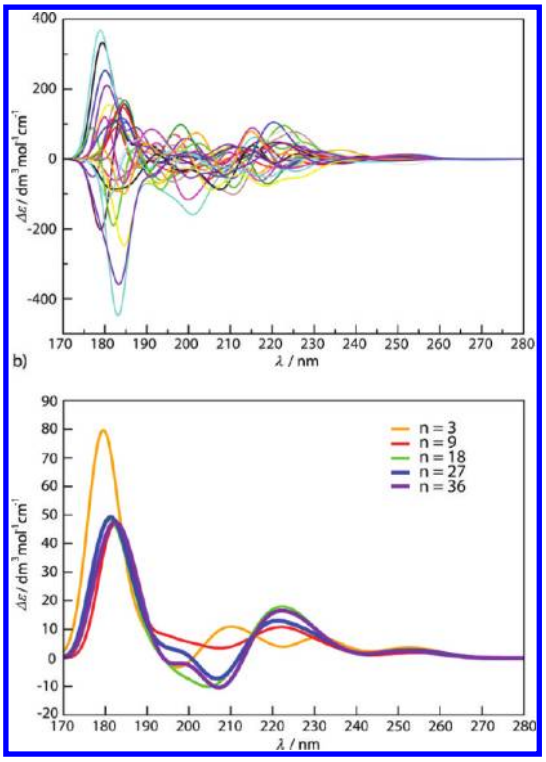


Figure 10. (a) Single CD spectra of 27 representative conformations of *R*-Ada$_Z$ obtained after partitioning cluster V into 27 subclusters. (b) Convergence can be seen when 27 representative conformations are used in order to obtain the average CD spectrum of cluster V (n is the number of subclusters into which the cluster V is partitioned and from which the representative structures are used to build the average CD spectrum).

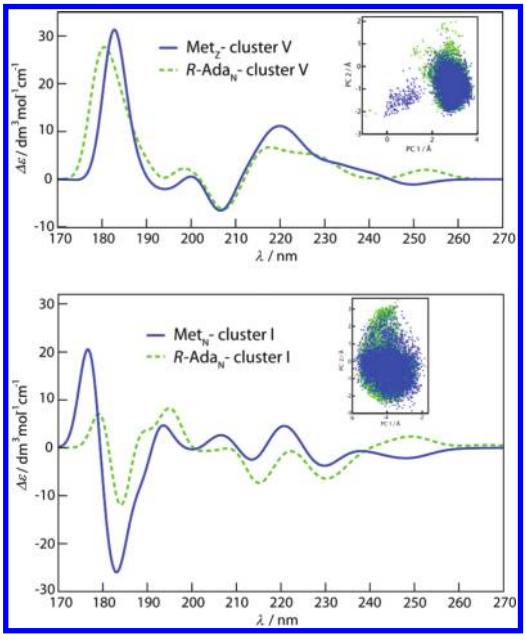


Figure 11. Average CD spectra corresponding to clusters V of *R*-Ada$_Z$ and Met$_Z$ and clusters I of Met$_N$ and *R*-Ada$_N$, shown in the upper and lower panel, respectively. The overlay of the clusters in the phase space is plotted in the insets.

accumulate in each step, from generating the phase space with a classical force field, over the sampling and the building of a subset of solvated representative structures, to

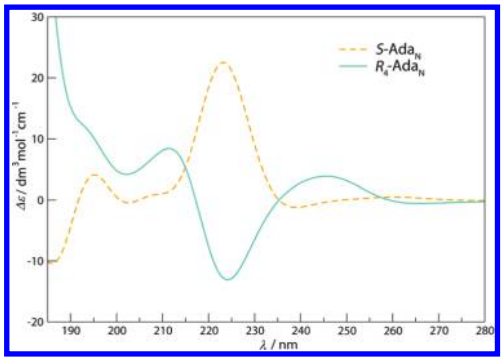


Figure 12. CD spectra of R_4-Ada$_N$ and *S*-Ada$_N$.

the TDDFT calculations, the sum of the two spectra is not exactly zero, and deviations of up to 20% have been observed (Figure 12). Nevertheless, the main features of both the R_4-Ada$_N$ and the *S*-Ada$_N$ spectra are identical and opposite, with main peaks appearing at exactly the same locations. Notably, these spectra emerged from two completely independent calculations performed on two strongly related but distinct peptides. Since the calculations on the neutral forms, with their less restricted phase spaces (Figure 8) are most prone to errors, this good correspondence supports the applicability of our method. This analysis is also supportive of the procedure of averaging over the different backbone conformations, present in each cluster, even when individual clusters have disparate spectra (e.g., Figure 11).

Figure 13a shows that the CD spectra of the zwitterionic peptides all exhibit negative minima. Specifically, Met$_Z$ possesses a

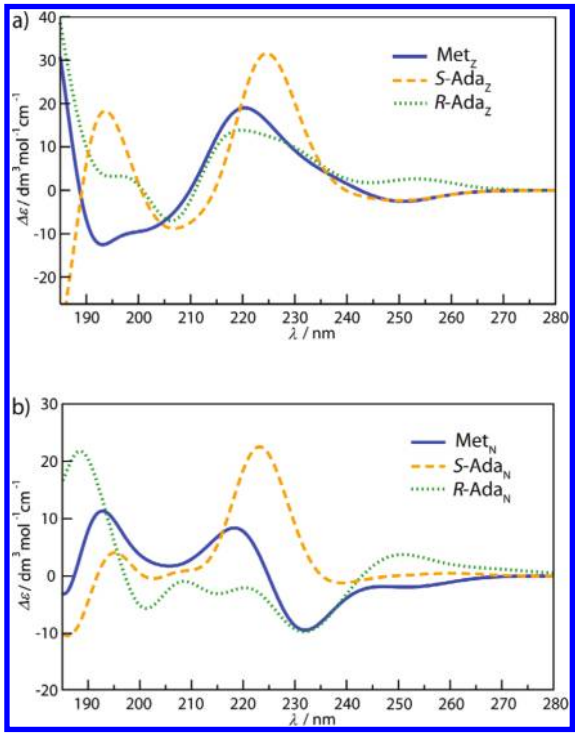


Figure 13. CD spectra of zwitterionic and neutral forms, shown in the upper and lower panels, respectively.

deep and wide minimum between 190 and 210 nm, followed by a positive peak at about 220 nm. This latter signal is associated with

dx.doi.org/10.1021/ct200868y | *J. Chem. Theory Comput.* 2012, 8, 1694−1705

the folded conformations residing in cluster V, which dominates the spectrum. Consistently with the similarities of the phase spaces, *R*-Ada$_Z$ has a spectrum with very similar peak positions as Met$_Z$, but of different intensities (Figure 13a).

The spectrum of Met$_N$ (Figure 13b) is very different from the spectrum of Met$_Z$. It remains positive for transitions between 185 and 225 nm, with peaks at 193 and 220 nm. The spectrum of *R*-Ada$_N$ becomes mostly negative after the peak at 185 nm. Interestingly, *S*-Ada shows a very pronounced peak at 225 nm, in both its zwitterionic and neutral forms. Also, both the *S*-Ada CD spectra have a peak at 195 nm, although of different intensities. The only difference between the two is the negative minimum in the spectrum of *S*-Ada$_Z$ (208 nm) that is not apparent in the spectrum of *S*-Ada$_N$.

Because the final spectra are the result of averaging 50 individual spectra, which are quite distinct (see, for example, Figure 10a), it is generally difficult to assign the spectral features in Figure 13 to any specific electronic transitions. Nevertheless, we have examined the underlying transitions and presented them for a set of five Met$_N$ structures (see Tables S2–S6 and the associated material, in the Supporting Information).

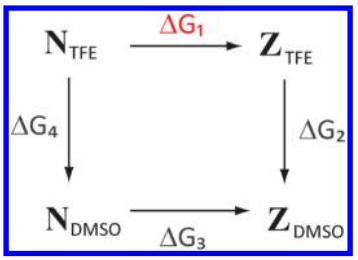
From this analysis, we are able to comment, for example, on some aspects of the π–π^* states of the aromatic chromophores. The ^1L_b states of Tyr and Phe are calculated to appear at approximately 250 and 233 nm, respectively. These values are some 25 nm lower than the corresponding experimental values.[42,43] The ^1L_a states of the same side chains are calculated to lie at approximately 213 and 208 nm, respectively. Both wavelengths are within about 15 nm of their corresponding experimental values.[43] Even though these states can carry significant rotatory strengths at the level of an individual structure (Tables S2–S6), the average rotatory strengths across the entire phase space are close to zero (Table S7). In this respect, such states can be expected to have only a very small net effect on the final CD spectra. This result, which is particularly apparent in the wavelength region around 250 nm, highlights the importance of averaging over a large number of individual structures.

Comparison of Theoretical and Experimental Spectra. In order to compare the calculated CD spectra of the individual peptides to those measured for Met and Ada in TFE,[5] two further steps are required. The first concerns the estimation of the equilibrium ratio of the neutral to zwitterionic forms (N/Z) of the individual peptides in TFE.

We base our calculations of the N/Z ratio in TFE on the measurements of the analogous ratio for Leu-enkephalin in DMSO, which was experimentally determined to be 3:2.[44] Our own simulations of Leu-enkephalin in TFE indicate that its conformational behavior is practically identical to that of Met, for both the neutral and zwitterionic forms (Figure S7). On this basis, we believe that 3:2 is also a good approximation of the N/Z ratio of Met in DMSO. Using this as a starting point, we have elected to estimate the N/Z ratio of Met in TFE by means of a free energy cycle describing the transfer from one solvent to the other (Scheme 1).

The desired free energy difference (for $[N/Z]_{TFE}$) emerges as $\Delta G_1 = \Delta G_3 + \Delta G_4 - \Delta G_2$, where ΔG_3 is assigned as 1 kJ mol^{-1}, based on $[N/Z]_{DMSO}$ = 3:2. We have estimated the (vertical) transfer free energies (ΔG_2 and ΔG_4), via the

Scheme 1. The Free Energy Cycle Used to Estimate the Equilibrium Ratio between the Neutral (N) and Zwitterionic (Z) Forms in TFE (ΔG_1) Based on the Experimentally Obtained Ratio (ΔG_3) for the Analogous Equilibrium for Leu-enkephalin in DMSO[44]



Zwanzig free energy perturbation scheme:

$$\Delta G(\mathrm{Met}_{TFE} \rightarrow \mathrm{Met}_{DMSO}) = -RT \ln \left\langle \exp\left(\frac{-(E_{DMSO} - E_{TFE})}{RT}\right)\right\rangle_{TFE} = -RT \ln \sum_{i=1}^{50} w_i \exp\left(\frac{-(E_{i,DMSO} - E_{i,TFE})}{RT}\right)$$

More specifically, we have evaluated the requisite differences in the solvation free energies by solving the linearized Poisson–Boltzmann equation within the Adaptive Poisson–Boltzmann Solver (APBS) program package.[45] Therein, the peptides were assigned the same *ff03*-derived parameters as described in the Methods section. TFE and DMSO were represented as continuous media with dielectric constants of 26.8 and 47.2, respectively, and the nonpolar solvation contributions were assumed to be equal in the two solvents. The individual calculations were performed for the same 50 representative structures used to obtain the CD spectra for each peptide. Similarly, we used the previously determined population fractions of the corresponding subclusters to represent the statistical weights (w_i) in the above expression.

Even though we expect this perturbation scheme to be most accurate for Met-enkephalin, we also performed the same calculations for both epimers of Ada-enkephalin (with ΔG_3 = 1 kJ mol^{-1}). For all three peptides, we find that the zwitterionic forms are present in fractions of 8–18% in TFE. Considering the uncertainties associated with this calculation, we estimate the fraction for all peptides to be very similar and equal to about 10% ($[N/Z]_{TFE}$ = 9:1). This is qualitatively consistent with the expectation that the zwitterionic form should be less favored in the less polar solvent. Consequently, we have used this ratio for weighting the contributions of neutral and zwitterionic forms to all composite spectra. One should note that increasing the contribution of the zwitterionic form up to 20% does not significantly affect the final CD spectra (Figure S8).

The second aspect that needs to be taken into account before making the final comparison between the published experimental and theoretical spectra is the fact that both epimers of Ada were present (in an equimolar ratio) under the experimental conditions. Consequently, the theoretical spectra were combined according to the expressions Met = 0.9Met$_N$ + 0.1Met$_Z$ and Ada = 0.9(*S*-Ada$_N$ + *R*-Ada$_N$)/2 + 0.1(*S*-Ada$_Z$ + *R*-Ada$_Z$)/2, for the purposes of comparison to experimental results.

The resulting final spectra are shown in Figure 14a and compared to the experimental spectra[5] in Figure 14b. The comparison immediately reveals that a number of features of the experimental

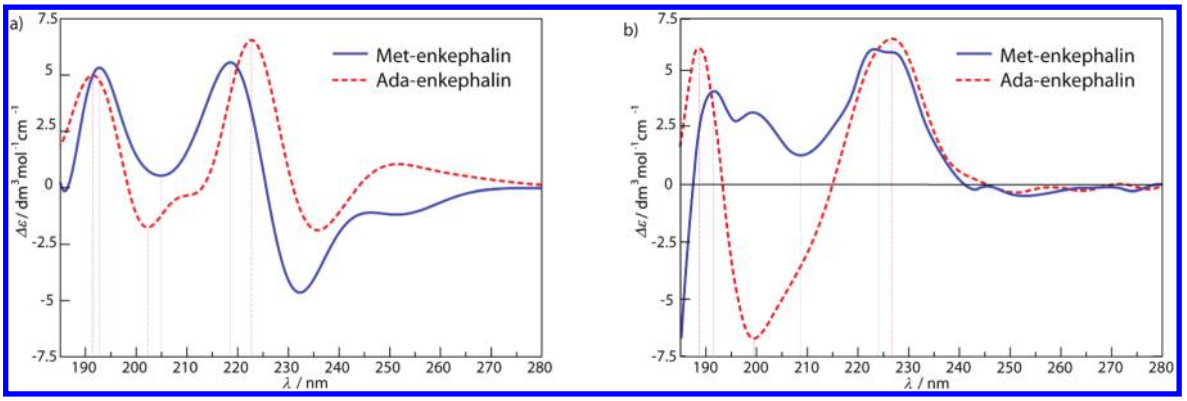


Figure 14. (a) Calculated CD spectra for Met-enkephalin and Ada-enkephalin. (b) Experimentally obtained CD spectra of Met-enkephalin and an epimeric mixture of Ada-enkephalin in TFE.[5]

spectra have been correctly reproduced. In the range of 185−230 nm, both the experimental and the theoretical CD spectra of Ada-enkephalin have a pronounced maximum at ∼190 nm, followed by a wide minimum between 200 and 215 nm. The strong peak at ∼224 nm obtained from calculations corresponds to the peak present at 227 nm in the experimental CD spectrum. Importantly, the analysis of the components contributing to the spectra shows that the minimum at 200 nm does not appear because of enhanced contribution of strongly folded structures of Ada, as was suggested previously.[5] Instead, we find that the strongest contribution at these wavelengths comes from *R*-Ada$_N$, which has a relatively disperse phase space.

A qualitatively good agreement was also obtained in the case of Met-enkephalin, where the correspondence between the theoretical and the experimental spectra is found in both peaks at ∼192 nm and ∼220 nm. Likewise, between the peaks, the spectra remain positive with a minimum appearing at 205 nm theoretically and 208 nm experimentally.

Between 230 and 240 nm, a relatively strong minimum appears in our calculations, even though the experimental spectra have no such particular features in this range. Taking into account the individual contributions of the forms, it is possible to see that it predominantly arises from the neutral forms of Met and *R*-Ada. It was not, however, possible to identify characteristic peptide structures responsible for this signal, as it was present in nearly all spectra of the relevant representative structures. A closer investigation of the excited states, however, has allowed us to pinpoint its origin.

We were able to identify an electronic transition of Met$_N$ consistently centered around 230 nm. Unlike the $\pi-\pi^*$ states discussed above, the rotatory strengths of this state did not average to zero over the conformational ensemble. Rather, the individual rotatory strengths are consistently negative, and the result is the anomalous minimum present in the final calculated spectra (Table S7). It is important to note that this is the only feature of the final spectra that we are able to unambiguously relate to a specific excited state in this way.

The state involves excitations originating from various occupied orbitals into the π^* orbital of the terminal COOH group. Among these are local promotions from the nonbonding orbital of the terminal carbonyl oxygen. Such $n-\pi^*$ states tend to appear around 205 nm in simple carboxylic acids.[46] In the present case, the mixing of less local contributions from the occupied π orbitals of adjacent amide bonds and the Phe side-chain, as well as nonbonding orbitals from nearby backbone oxygens, seems to lower the transition energy. At first sight, this might seem like a classic case of B3LYP artificially stabilizing so-called charge-transfer excitations.[47,48] In this light, we carried out selected testing using the Coulomb-attenuated CAM-B3LYP[49] functional, which is designed to correct the artificial stabilization.[50] Such testing, however, did not reveal any significant change in the position, rotatory strength, or nature of this state.

On this basis, it is difficult to unambiguously ascribe the problem as being related to charge-transfer.[51−53] It may be that both B3LYP and CAM-B3LYP give consistent yet erroneous wavelengths in this case or that there is a systematic error in the prediction of the rotatory strengths for this particular state. Due to the persistence of this feature across all structures, we find it unlikely that the error is related to the structural ensemble in terms of potential shortcomings in either the force-field or the sampling. Thus, we believe the discrepancy is due to a problem with TDDFT, although discerning the exact nature of this problem presents a significant challenge at this time.

Despite the apparently erroneous persistence of this minimum in the theoretical calculations, the spectra below 230 nm reproduce the main features of the experiments quite satisfactorily.

■ CONCLUSIONS

In this work, we investigated the conformational phase spaces of both Met-enkephalin and Ada-enkephalin in TFE in an attempt to calculate their respective CD spectra (see the flowchart in Scheme S1). Using replica exchange molecular dynamics, we first generated individual conformational phase spaces from 34 000 structures of the zwitterionic and neutral forms of Met-enkephalin and of both the *R*- and the *S*-epimers of Ada-enkephalin. Each phase space was characterized in detail by combining principal component analysis with clustering algorithms. For the calculation of the CD spectra, a characteristic subset of 50 representative conformations for each peptide was chosen, each representing a particular section of the phase space comprising about 700 structures. In agreement with previous studies, we have observed strong solvent effects on the spectra. These were taken into account by calculating an average field consisting of point charges originating from a number of solvent configurations around each peptide conformation in the subset. Subsequently, the CD spectrum of each solvated conformation was obtained using the QM/MM method incorporating TDDFT calculations with the B3LYP functional. The average CD spectra were obtained by taking the mean of all weighted single CD spectra, where the weighting factor corresponds to the population fraction of the given conformation. Qualitatively, our procedure can be considered as the sequential averaging over the solvent (for each representative conformation),

the side chains (inside the backbone-based clusters), and the backbone (combining the principal clusters).

The reproducibility and the internal errors of the method are best demonstrated in Figure 12. Thereby, spectra of the R_4-Ada$_N$ and S-Ada$_N$ enantiomers are shown after being constructed according to the flowchart (Scheme S1). In an ideal case, the two spectra should be mirror images of one another. Despite an overall good agreement, however, deviations of up to 20% are determined in some parts of the spectrum, with larger deviations at lower wave numbers.

In comparison with experiments, a larger absolute error has been found. Thereby, the uncertainties due to the classical force field, the sampling and clustering procedures as well as the reduction of the phase space onto a set of representative solvated structures, all contribute in a cumulative fashion. An additional difficulty is the need to construct the final spectra from different peptide forms. Despite the numerous complications, qualitative agreement with experimental results has been obtained below 230 nm, and the main features of the measured spectra have been reproduced by our calculations. The simulated spectra show a persistent minimum between about 230 and 240 nm that is not present in the experimental measurements and which seems to arise from the poor treatment of a particular excited state by TDDFT.

The method developed herein establishes a direct link between the conformational phase space of a small peptide in solution and its CD spectrum and provides an answer to the question, can one associate the CD spectra with structural features of peptides, as it is done for proteins? Our results show that similar parts of the phase space provide similar spectra as evidenced by the signals obtained from both the confined cluster V of Met$_Z$ and R-Ada$_Z$ and the sparse cluster I of Met$_N$ and R-Ada$_N$ (Figure 11). Consequently, if sampled correctly, similar structures will give similar CD spectra. However, we also show that different parts of the phase space may have similar CD signals, which can be seen by comparing the spectra of S-Ada$_N$ and S-Ada$_Z$ (Figure 13).

The emergent conclusion of our work is that the current state of the art theoretical methods show significant promise in reproducing experimentally observed CD spectra of flexible, solvated peptides. In the context of the methodology presented in this work, further development of the force fields and, particularly, functionals and basis sets for TDDFT calculations can only yield a more quantitative agreement. We show that theoretical calculations become necessary for interpretations of the experimental measurements, which cannot be performed in an unambiguous manner when attempting to determine the correspondence between the peptide structures in the underlying ensemble and the measured CD spectrum.

The comparison between the various peptides in TFE shows a trend in the extent of ordering such that Met < S-Ada < R-Ada, irrespective of the nature of the terminal groups (Figures 4 and 8). The surprising realization is that the two epimers of Ada actually have very different phase spaces. However, even though R-Ada phase spaces are more restricted than those of S-Ada (that are superficially more similar to Met), we have shown that R-Ada and Met actually have a much larger overlap of representative structures (Figures 6 and 7) and, indeed, their CD spectra are quite similar (Figure 13). The representative conformations for S-Ada, in both the neutral and zwitterionic forms, consistently exhibit a different hydrogen bonding pattern than those of the other two peptides, which is also evident in the contrasting CD spectra. These differences may be related to the enhanced capacity of S-Ada to penetrate cellular membranes and act as an antitumor agent.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1−S8, Tables S1−S16, Scheme S1, an excited state discussion, and force field parameter files for nonstandard residues. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: smith@physik.uni-erlangen.de, David.Smith@irb.hr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by Croatian Ministry of Science (Project No. 098-0982933-2937) and the EAM Cluster of Excellence at the Friedrich Alexander University Erlangen-Nürnberg. Z.B. acknowledges support from the Elite Network of Bavaria. A-S.S. thanks the Ruđer Bošković Institute for hospitality during preparation of this manuscript. We sincerely thank Drs. Kata Majerski and Štefica Horvat for helpful discussions.

■ REFERENCES

(1) Kendrew, J. C. In *The Encyclopedia of Molecular Biology*; Blackwell Science: Oxford, United Kingdom, 1994; pp 772−781.
(2) Sanbonmatsu, K. Y.; Garcia, A. E. *Proteins* **2002**, *46*, 225.
(3) Terenius, L. *Alcohol* **1996**, *13*, 31.
(4) Horvat, Š; Mlinarić-Majerski, K.; Glavaš-Obrovac, Lj.; Jakas, A.; Veljković, J.; Marczi, S.; Kragol, G.; Roščić, M.; Matković, M.; Milostić-Srb, A. *J. Med. Chem.* **2006**, *49*, 3136.
(5) Gredičak, M.; Supek, F.; Kralj, M.; Majer, Z.; Hollosi, M.; Šmuc, T.; Mlinarić-Majerski, K.; Horvat, Š *Amino Acids* **2009**, *38*, 1185.
(6) Cheng, F.; McLaughlin, P. J.; Verderame, M. F.; Zagon, I. S. *Mol. Cancer* **2008**, *7*, 5.
(7) Cheng, F.; McLaughlin, P. J.; Verderame, M. F.; Zagon, I. S. *Cancer Res.* **2007**, *67*, 10511.
(8) Zagon, I. S.; Verderame, M. F.; McLaughlin, P. J. *Brain Res. Rev.* **2002**, *38*, 351.
(9) Kuroda, Y.; Ueda, H.; Nozawa, H.; Ogoshi, H. *Tetrahedron Lett.* **1997**, *38*, 7901.
(10) Knipp, G. T.; Vander Velde, D. G.; Siahaan, T. J.; Borchardt, R. T. *Pharm. Res.* **1997**, *14*, 1332.
(11) Marcotte, I.; Separovic, F.; Auger, M.; Gagne, S. M. *Biophys. J.* **2004**, *86*, 1587.
(12) Andrew, C. D.; Bhattacharjee, S.; Kokkoni, N.; Hirst, J. D.; Jones, G. R.; Doig, A. J. *J. Am. Chem. Soc.* **2002**, *124*, 12706.
(13) Bhattacharjee, S.; Tóth, G.; Lovas, S.; Hirst, J. D. *J. Phys. Chem. B* **2003**, *107*, 8682.
(14) Tanaka, T.; Kodoma, T. S.; Morita, H. E.; Ohno, T. *Chirality* **2006**, *18*, 652.
(15) Besley, N. A.; Hirst, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 936.
(16) Hirst, J. D.; Colella, K.; Gilbert, A. T. B. *J. Phys. Chem. B* **2003**, *107*, 11813.
(17) Gilbert, A. T. B.; Hirst, J. D. *THEOCHEM* **2004**, *675*, 53.
(18) Rogers, D. M.; Hirst, J. D. *Chirality* **2004**, *16*, 234.
(19) Woody, R. W. *Monatsh. Chem.* **2005**, *136*, 347.
(20) Jiang, J.; Abramavisius, D.; Bulheller, B. M.; Hirst, J. D.; Mukamel, S. *J. Phys. Chem. B* **2010**, *114*, 8270.
(21) Abramavicius, D.; Jiang, J.; Bulheller, B. M.; Hirst, J. D.; Mukamel, S. *J. Am. Chem. Soc.* **2010**, *132*, 7769.
(22) Glättli, A.; Daura, X.; Seebach, D.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2002**, *124*, 12972.
(23) Daura, X.; Bakowies, D.; Seebach, D.; Fleischhauer, J.; van Gunsteren, W. F.; Krüger, P. *Eur. Biophys. J.* **2003**, *32*, 661.

(24) Fidler, J.; Rodger, P. M.; Rodger, A. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 235.
(25) Kaminský, J.; Kubelka, J.; Bouř, P. *J. Phys. Chem. A* **2011**, *115*, 1734.
(26) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J. W.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *23*, 1999.
(27) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(28) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*, University of California: San Francisco, CA, 2008.
(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., 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.
(30) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. *J. Comput. Chem.* **2009**, *30*, 2157.
(31) Dupradeau, F.-Y.; Cézard, C.; Lelong, R.; Stanislawiak, E.; Pecher, J.; Delepine, J. C.; Cieplak, P. *Nucleic Acids Res.* **2008**, *36*, 360.
(32) Garcia, A. E.; Herce, H.; Paschek, D. *Annu. Rep. Comput. Chem.* **2006**, *2*, 83.
(33) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. *J. Comput. Chem.* **2005**, *26*, 1668.
(34) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., III. *J. Chem. Theory Comput.* **2007**, *3*, 2312.
(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford CT, 2009.
(36) Diedrich, C.; Grimme, S. *J. Phys. Chem. A* **2003**, *107*, 2524.
(37) Jensen, L.; Swart, M.; van Duijnen, P. T.; Autschbach, J. *Int. J. Quantum Chem.* **2006**, *106*, 2479.
(38) Sanchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Comput. Chem.* **1997**, *18*, 313.
(39) Coutinho, K.; Georg, H. C.; Fonseca, T. L.; Ludwig, V.; Canuto, S. *Chem. Phys. Lett.* **2007**, *437*, 148.
(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(41) Su, L.; Cukier, R. I. *J. Phys. Chem. B* **2007**, *111*, 12310.
(42) Zhang, L.; Peslherbe, G. H.; Muchall, H. M. *Photochem. Photobiol.* **2006**, *82*, 324.
(43) *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; p 75.
(44) Gerothanassis, I. P.; Birlirakis, N.; Karayannis, T.; Tsikaris, V.; Sakarellos-Daitsiotis, M.; Sakarellos, S.; Vitouxb, B.; Marraud, M. *FEBS Lett.* **1992**, *298*, 188.
(45) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037.
(46) Limão-Vieira, P.; Giuliani, A.; Delwiche, J.; Parafita, R.; Mota, R.; Duflot, D.; Flament, J.-P.; Drage, E.; Cahillane, P.; Mason, N. J.; Hoffmann, S. V.; Hubin-Franskin, M.-J. *Chem. Phys.* **2006**, *324*, 339.
(47) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859.
(48) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943.
(49) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.
(50) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.
(51) Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C. *J. Chem. Theory Comput* **2009**, *5*, 2420.
(52) Jamorski, C.; Foresman, J. B.; Thilgen, C.; Lüthi, H.-P. *J. Chem. Phys.* **2002**, *116*, 8761.
(53) Shemesh, D.; Hättig, C.; Domcke, W. *Chem. Phys. Lett.* **2009**, *482*, 38.