HAL
open science

Characterization of β-turns by electronic circular dichroism spectroscopy: A coupled molecular dynamics and time-dependent density functional theory computational study

Mattia Migliore, Andrea Bonvicini, Vincent Tognetti, Laure Guilhaudis, Marc Baaden, Hassan Oulyadi, Laurent Joubert, Isabelle Ségalas-Milazzo

To cite this version:

Mattia Migliore, Andrea Bonvicini, Vincent Tognetti, Laure Guilhaudis, Marc Baaden, et al.. Characterization of β-turns by electronic circular dichroism spectroscopy: A coupled molecular dynamics and time-dependent density functional theory computational study. Physical Chemistry Chemical Physics, 2020, 22 (3), pp.1611-1623. 10.1039/C9CP05776E . hal-02397135

HAL Id: hal-02397135

https://hal.science/hal-02397135v1

Submitted on 22 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Characterization of β-turns by electronic circular dichroism spectroscopy: A coupled molecular dynamics and time-dependent density functional theory computational study

Mattia MIGLIORE,[a+] Andrea BONVICINI,[a+] Vincent TOGNETTI,[a]* Laure GUILHAUDIS,[a]* Marc BAADEN,[b] Hassan OULYADI,[a] Laurent JOUBERT,[a] Isabelle SEGALAS-MILAZZO[a]

[a] Normandy Univ., COBRA UMR 6014 & FR 3038, Université de Rouen, INSA Rouen, CNRS, 1 rue Tesniére 76821 Mont St Aignan, Cedex, France

[b] Laboratoire de Biochimie Théorique, CNRS, UPR9080, Univ Paris Diderot, Sorbonne Paris Cité, PSL Research University, 13 rue Pierre et Marie Curie, 75005 Paris, France

Abstract

Electronic circular dichroism is one of the most used spectroscopic techniques for peptide and protein structural characterization. However, while valuable experimental spectra exist for α-helix, β-sheet and random coil secondary structures, previous studies showed important discrepancies for β-turns, limiting their use as reference for structural studies. In this paper, we simulated circular dichroism spectra for the most encountered β-turns in peptides, namely types I, II, I' and II'. In particular, by combining classical molecular dynamics simulations and state-of-the-art quantum time-dependent density functional theory (with polarizable embedding multiscale model) computations, two common electronic circular dichroism patterns were found for couples of β-turn types (type I/type II' and type II/type I'). These patterns were subsequently rationalized based on the exciton coupling theory, paving the way for a possible experimental characterization of β-turns based on circular dichroism spectroscopy.

Keywords: secondary structures, β-turns, electronic circular dichroism, molecular dynamics, time-dependent density functional theory, polarizable embedding multiscale model

* Corresponding authors: vincent.tognetti@univ-rouen.fr, laure.guilhaudis@univ-rouen.fr

[+] A. B. and M. M. contributed equally to the work

1. Introduction

β-turns (also known as tight-turns, β-bends or reverse turns), recognized in 1968 by Venkatachlam,[1] are motifs that allow to reverse the polypeptide chain direction in proteins and peptides. They constitute, after α-helices and β-sheets, the third most common secondary structures present in these macromolecules.[2] Their structural role notably occurs in globular protein folding,[3–7] in molecular recognition processes between proteins, and in interactions between peptide substrates and receptors.[8] In particular, β-turns have been reported to play an essential role in recognition phenomena between peptide-activated G protein-coupled receptors (GPCRs) and their endogenous peptide ligands.[9,10] Moreover, the structural characterization of β-turns is mandatory for a full understanding of the molecular interactions taking place between the partners and to aid the design of pseudo-peptidic analogs.[11,12]

From a structural point of view, β-turns involve four sequential amino-acids (denoted *i*, *i+1*, *i+2*, *i+3*, following the ordering from the N- to the C- terminal residues), and they can be classified into different types based on the Φ and Ψ dihedral angles of their central residues (*i+1* and *i+2*). They are thus composed by three peptides bonds and are stabilized by an intramolecular hydrogen bond between the H_N atom in residue *i+3* and the CO group of residue *i*, as depicted in Figure 1a.

Originally, six types of β-turns were defined (I, II, III and their mirror images I', II' and III').[1] Over the last 50 years, with the increasing number of solved protein structures, additional types were identified (IV, V, VI, VII and VIII), some subtypes were introduced (IV_1, IV_2, IV_3, IV_4, VIa1, VIa2, VIb), while some types were eliminated (III, III', V and VII) and incorporated in others categories (for example III into I, III' into I').[1,2,6,13–19] Although a new nomenclature has recently been proposed,[19] we decided, in order to allow direct comparisons with the dense literature on such topic, to keep the classification proposed by De Brevern in 2016.[18] As shown in Table 1, it is composed of 12 types of β-turns.

It should be emphasized that, as β-turns are flexible secondary structures, each Φ and Ψ dihedral angle of the *i+1* and *i+2* residues is allowed to deviate up to ±30° compared to its canonical value.[18] Last but not least, while α-helices and β-sheets are ordered and repetitive secondary structures, β-turns are ordered but non-repetitive motifs. For all these reasons, the establishment of reference electronic circular dichroism (ECD) spectra for β-turns presents more difficulties than those encountered for the two other major secondary structures (α-helix and β-sheet) and for the random coil conformation.[20–23] Accordingly, it still remains an open question in the field of structural biology, which may be tackled within either experimental or theoretical perspectives.

It should be mentioned that other experimental techniques, such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, may be used to determine the secondary and tertiary structures of peptides and proteins. However, although the first method allows identifying all secondary structural motifs present in the molecule in solid state, its use is intrinsically limited by the ability of the molecule to crystallize. This is a particular issue for peptides whose flexibility often hinders crystallization.[24,25]

On the other hand, NMR gives access to the solution structure of peptides or proteins, for which helices and sheets are easily distinguished, whereas turns are more difficult to identify.[26] Indeed, this technique only allows to characterize four types of β-turns (i.e. I, I', II and II') among the twelve described in Table I, and the distinction between these four types is often not absolute.[26] Moreover, this interpretation is made more complicated in the case of small size peptides, which frequently exhibit more than one conformation in solution.[27–29]

Besides, it should be pointed out that both methods are time and sample consuming compared to electronic circular dichroism, which requires a small amount of sample and is easy to implement from an experimental viewpoint. Circular dichroism spectroscopy is thus the privileged technique used to screen the secondary structure of polypeptides. Nevertheless, the lack of

reference spectra for β-turns makes the estimations of polypeptide secondary structure content rather unreliable when the molecule contains a large amount of this kind of motif.[22,30,31]

From a theoretical point of view, the first work on ECD spectroscopy of β-turns was actually performed by Woody in 1974,[32] who identified via semi-empirical calculations (an improved version of the Kirkwood-Moffitt-Tinoco coupled oscillators theory based on the electrostatic interaction between chromophores[33], also known as the matrix method[34], which includes the description of charge transfer transitions between amide groups[35]) the fundamental role of the n → π* and π_{nb} → π* transitions for types I, II and III β-turns. More specifically, these calculations indicated that all analysed types were generally characterized by one negative n → π* band at ~225 nm, followed by one positive and one negative π_{nb} → π* bands at ~205 and 190 nm. Such an ECD pattern has been denoted as class B spectra (see Table 2 for all the ECD patterns found in this work).

Such findings were then challenged by the theoretical work by Sathyanarayana and Applequist[36] (still at the semi-empirical level based on the DeVoe theory, a classical approach of dispersive oscillators with complex frequency-dependent polarizabilities), focusing only on the π_{nb} → π* (NV1) transition, thus omitting the n → π* transitions.[33] Indeed, their computed ECD spectra of Ace-X-X-NHMe peptides (with X=Ala, Gly) showed, on the one hand, that types I and III share similar spectra with a negative band (~205 nm) and a positive band (~190 nm), but that, on the contrary, a mirror spectrum (with a maximum at 205 nm and a minimum at 190 nm) was found for type II.

On the other hand, the available experimental ECD reference spectra for β-turns were derived using various approaches. The first one consists in subtracting the factored α-helix, β-sheet and random coil contributions from the globular protein circular dichroism spectra.[37,38] This leads to a general ECD β-turn spectrum, which does not take into account the various possible types.

Thus, it remains intricate to correlate this obtained ECD spectrum with a precise β-turn conformation.

A second approach based on a comparative study of theoretical and experimental peptide ECD spectra was performed by Brahms and collaborators in 1977.[39] Using a poly(Ala_2-Gly_2) peptide, they obtained a single ECD curve similar to the ones calculated by Woody[25] and thus considered it as representative of both types I (or III) and II. However, no physical evidence has been shown to indicate which type of β-turn(s) may be present in this peptide, although the authors stated that a type II should be expected.[40]

The third and most used approach relies on the comparative structural studies of peptides by ECD, NMR and X-ray crystallography.[28,29,41–44] In this approach, cyclic or modified aminoacid peptides were used to restrain the flexibility of the molecules. For instance, ECD spectra, composed by a negative band at 227-222 nm and a positive one at 200-205 nm, were observed for cyclic hexapeptides (cyclo(Gly/Orn-L-Pro-D-Phe/Cha)$_2$) adopting a type II β-turn geometry.[41]

Although ECD experiments were not recorded below 190 nm, the obtained pattern resembles class B of Woody's classification (Table II). Similarly, Bandekar et al. observed[44] a class B ECD spectra, with a minimum below 188 nm, a maximum at 202-204 nm and another minimum around 228-230 nm, for the cyclo(L-Ala-D-Ala-Aca) peptide forming a type II β-turn. In the same study, a class C pattern (a maximum near 190 nm and two distinct minima at 206-208 and 218-222 nm) was observed for the cyclo(L-Ala-L-Ala-Aca) peptide that exhibits a type I (and III) β-turn.

A few years later, experimental ECD spectra from non-cyclic model peptides (Z-Aib-L-Pro-Aib-L-Pro-OMe (**1**), Piv-L-Pro-Aib-NHMe (**2**), Piv-L-Pro-D-Ala-NHMe (**3**) and Piv-L-Pro-L-Val-NHMe (**4**)) with various β-turn types (types I, II or III) were reported by Crisma et al.[35]

They observed two different ECD patterns for type I: a class C for peptide **1** and a class B for peptide **4**. These discrepancies led them to assign the class C pattern to a type III structure, rather than type I. Type II (peptides **2** and **3**) was characterized by a ECD spectrum having two positive bands at ~230 nm and ~200 nm, a feature not predicted by Woody's calculations.

It should be emphasized that, in all these studies, the type of turn was characterized in solid-state from X-ray analyses, and that the authors assumed that the same folding also persists in solution. This assumption was based on NMR analyses, performed in solvents different from those used for ECD spectroscopy, and it thus might be arguable.[28,45] As a consequence, the discrepancies observed in these experimental studies might be linked to the presence of additional conformations in the solvents used for recording ECD spectra.

Interestingly, such ambiguities were not removed by subsequent ECD spectra reported on pseudohexapeptides that adopt a conformational mixture of type I and type II β-turns (determined by NMR experiments) in different solvents.[28,29,43,46,47] In one of these studies,[40] two different patterns have been attributed to type I: one presenting a negative band between 225-230 nm and a positive band at ~195 nm, a second characterized by a minimum at ~205 nm and a maximum at ~195 nm, while one spectrum has been identified for type II featuring a positive band at ~200 nm.

It can thus be concluded from this fast literature survey that, because of all of the inconsistencies found in theoretical and experimental works and owing to the fact that ECD is one of the most used experimental methods to study peptides and proteins secondary structures, the establishment of reliable reference ECD spectra in solution remains highly desirable.

For such reasons, we investigated in this paper if β-turns have characteristic ECD patterns and if ECD spectroscopy can distinguish between the four principal β-turn types described in solution by NMR spectroscopy:[26] i.e., types I, I', II and II'. In particular, we used state-of-the-art

advanced quantum chemical computations, based on the time-dependent density functional theory (TD-DFT) formalism, coupled with classical molecular dynamics (MD) simulations on "minimal β-turn models" in aqueous solutions. In order to take into account solvent effects in the computed ECD spectra (both the work of Woody[32] and Sathyanarayana and Applequist[36] were actually performed in vacuum), we used the polarizable embedding (PE) multiscale scheme,[48,49] as described in the following section.

2. Computational details

2.1. A minimal peptide model for β-turns

Our "minimal β-turn model" can be described as a di-amino acid peptide with N-terminal acetyl and C-terminal N-Me amide capping groups (Ace_i-X_{i+1}-X_{i+2}-$NHMe_{i+3}$), as represented in Figure 1a. For each conformation, the initial "canonical" Φ_{i+1}, Ψ_{i+1}, Φ_{i+2}, and Ψ_{i+2} dihedral angles (see Table 1) were built using the UCSF Chimera molecular modelling software.[50]

2.2. Classical molecular dynamics simulations

Each β-turn conformation was explored employing *restrained* classical MD simulations using the GROMACS 5.4.1 software[51] and the OPLS-AA force field.[52] In particular, all MD processes (energy minimization, equilibration, and production) were performed in all cases with energy penalty functions consisting of additional quadratic potentials placed on the Φ_{i+1}, Ψ_{i+1}, Φ_{i+2}, and Ψ_{i+2} dihedral angles with a force constant set to 10000 kJ/mol/rad^2 (except for energy minimization step where it was reduced to 1000 kJ/mol/rad^2, in order to allow the molecule to reach the conformational minimum more easily, but always constrained to be close to a specific β-turn conformation) applied when the dihedral angles deviations with respect to the canonical values (Table 1) were larger than ±10°, ±20° or ±30°.

For each β-turn type, the minimal peptide model was solvated in a water box (~3x3x3 nm^3) using the TIP3P solvent water model[53] (~700 water molecules) using the GROMACS solvation implementation. The solvation box was neutralized with a counter-ion in case of charged amino-acids (note that we considered the protonation state of amino acid side chains at physiological pH 7.4).

Once the equilibrium geometry was obtained from the energy minimization step, the temperature of the system was equilibrated in the NVT ensemble at 300 K for 1 ns using the Berendsen thermostat (with a τ_t = 0.1 ps time-step). After this NVT equilibration, a pressure equilibration was carried out in the NPT ensemble (with P =1 bar) using still the Berendsen thermostat and with the Parrinello-Rahman barostat (with a τ_p = 0.1 ps time-step).[54] Then, we performed the 10 ns production run, still in the NPT ensemble. From each MD production, the minimal peptide model was centred in the box, and a total of 100 snapshots were extracted every 100 ps for the simulation of absorption and ECD spectra.

2.3. Quantum chemistry calculations

Absorption and ECD spectra were simulated by linear-response TDDFT with the DALTON 2018 quantum chemistry software.[55] The CAM-B3LYP range-separated hybrid exchange-correlation functional[56] was used in conjunction with Pople's 6-31+G(d) and 6-311++G(d,p) basis-sets.[57]

Environment effects were taken into account by using the Polarizable Embedding (PE) multiscale model of Kongsted and co-workers,[48] which was already used in the determination of ECD spectra of stereoisomers of terrein molecule.[49]

Basically, the PE model divides the total system in two sub-systems: a "QM" region (usually the molecule or a part of the molecule containing the chromophore), which is described by a quantum mechanical method (Kohn-Sham (KS)-DFT in our case), and a "MM" region (usually

the rest of the molecule and/or the solvent molecules) which is described classically using point charges (and/or multipoles) that can be polarized by the electron density of the QM part. The effective Hamiltonian operator for the PE model will thus describe the interaction between the two sub-systems and can be written according to:

$$\hat{H}^{eff} = \hat{v}^{KS} + \hat{v}^{emb} = \hat{v}^{KS} + \hat{v}^{es} + \hat{v}^{ind} \quad (1)$$

where $\hat{v}^{KS}$ is the conventional vacuum KS operator (defined by a specific DFT exchange-correlation functional), while $\hat{v}^{emb}$ is the embedding operator which is further divided into two contributions: $\hat{v}^{es}$ that describes the electrostatic interaction between fixed permanent charges and the QM region, and $\hat{v}^{ind}$ that describes the polarization of the environment (induced dipoles) due to the QM part (which directly depends on the electron density of the former). The solution of the quantum-mechanical and classical polarization equations are solved by a self-consistent scheme in which the induced dipoles at each SCF iteration are updated until complete relaxation of both sub-systems.[58]

Consequently, a main advantage of the PE model with respect to some other electrostatic embedding QM/MM schemes (e.g. COBRAMM[59,60]) stems from the fact that the mutual polarization between the QM region and the environment is taken into account. Moreover, the PE model is extended to calculation of excited state properties within the quantum-mechanical response theory[61] and can be used in the simulation of linear (e.g. electronic absorption[48] and ECD[49] spectroscopies) and even non-linear (e.g., two-[62,63] and three-photon[64] absorption spectroscopies) response properties. Further details about the mathematical form of $\hat{v}^{es}$, $\hat{v}^{ind}$ and their incorporation within the response theory can be found in refs. 48,58,61.

The minimal peptide model (see Figure 1) was thus treated as the QM region (also called "core"). However, in order to check the effects of hydrogen bonds between the minimal model and solvent water molecules on ECD spectra, water molecules within 0 (no water molecules

added), 5 and 6 Å from the centre of mass of the minimal peptide model were also added to the QM sub-system. The remaining water molecules (and ions when present) were added to the polarizable MM region using the standard embedding parameters (charges and isotropic polarizabilities) provided by the "solvent embedding potential" (SEP).[65] For each frame, the potential files were constructed using the PyFraME library.[58,66]

Then, for each snapshot, the total simulated absorption spectra ($\varepsilon_{\text{total}}$, expressed in $M^{-1}cm^{-1}$) was created by summation of normalized Gaussian function (ε_{0i}) centred around each individual electronic transition, using the following equation (eq. 5 in ref. [67]):

$$\varepsilon_{total} = \sum_{i=1}^{n} \varepsilon_{0i} = \sum_{i=1}^{n} \frac{\sqrt{\pi} e^2 N_A}{1000\ ln10\ c^2 m_e} \frac{f_{0i}}{\Delta}\ exp\left[-\left(\frac{E - \Delta E_{0i}}{\Delta}\right)^2\right], \qquad (2)$$

where n is the total number of computed excited states (here fixed at 20) and ε_{0i} is the gaussian broadening function applied to the i-th transition, e is the electron charge in esu units, N_A is the Avogadro number, c stands for the speed of light in vacuum (in cm s^{-1}), m_e is the electron mass (in kg), Δ is half the bandwidth at $1/e$ of peak height expressed in cm^{-1} (set to 0.3 eV but expressed in cm^{-1}), ΔE_{0i} is the transition energy from the ground state to the i-th excited state, and f_{0i} is the corresponding oscillator strength for the i-th transition.

Similar to absorption spectra, the total simulated ECD spectra (molar ellipticities, $\Delta\varepsilon_{\text{total}}$, expressed in $M^{-1}cm^{-1}$ units) for each snapshot was created by summation of normalized Gaussian functions ($\Delta\varepsilon_{0i}$) placed around each individual electronic transition, using the following equation (taken from ref. [68], eq. 8):

$$\Delta\varepsilon_{total} = \sum_{i=1}^{n} \Delta\varepsilon_{0i} = \sum_{i=1}^{n} \frac{\Delta E_{0i}}{2.296\times10^{-39}\sqrt{\pi}\Delta} R_{0i}\ exp\left[-\left(\frac{E - \Delta E_{0i}}{\Delta}\right)^2\right], \qquad (3)$$

where, as before, n is the total number of computed excited states (20) and $\Delta\varepsilon_{0i}$ is the gaussian broadening function applied to the i-th ECD transition, ΔE_{0i} is the transition energy from the ground state to the i-th excited state, Δ is half of the bandwidth at $1/e$ peak height expressed in energy units (set to 0.3 eV if not otherwise stated), and R_{0i} is the corresponding rotatory strength (in 10^{-40} esu^2 cm^2 units) for the i-th transition (expressed in the velocity form, in order to ensure its origin-independence).

Both absorption and ECD spectra were subsequently averaged for the 100 snapshots extracted from the restrained MD simulation.

3. Results and discussion

3.1. Protocol validation

First, we will discuss the validation of our computational protocol on the minimal model Ace_i-A_{i+1}-A_{i+2}-$NHMe_{i+3}$ in type I β-turn conformation. In such cases of aliphatic side-chains, the main chromophoric units are the peptide bonds (see Figure 1b). There are three bonds of that type in the model peptide, and their interactions between their excited states are expected to be sensitive to different β-turn conformations.

However, such sequences, Ace_i-X_{i+1}-X_{i+2}-$NHMe_{i+3}$, are too short peptides to lead to strong structuration in a specific β-turn or even in a bend conformation without suitable modifications[69] as for example cyclization[70], chemical design of residues[71,72] and/or capping units in the sequence.[73] Without the use of any restraint, such peptides would be mainly unfolded along the MD simulation, and the β-turns types we would like to investigate would be only very little populated. This would imply that only very few snapshots would be informative, and that, in consequence, the simulated averaged ECD spectra would not be statistically relevant. This issue

was then circumvented by including dihedral restraints on each Φ and Ψ dihedral angles for the two central residues (*i+1* and *i+2*).

Although this approach precludes a straightforward thermodynamical study of these minimal β-turn models, it conversely constitutes a prolific source of conformations for each type of β-turns (generated within the ±10°, ±20° or ±30° allowed dihedral angle deviations with respect to the canonical values gathered in Table 1), which were subsequently used in the ECD spectra calculations.

Figures S1-S4 in the Supporting Information File (SIF) reports the Φ and Ψ distributions produced by the classical MD simulations for the two central residues (*i+1* and *i+2*). They show that the range of allowed angles (±10° in this case) around the canonical values was almost uniformly explored, excluding the presence of conformational subgroups within each type of β-turns.

In case of peptides with a charged residue (*i.e.* Ace-K-A-NHMe, Ace-A-K-NHMe, Ace-D-A-NHMe and Ace-A-D-NHMe), the system was neutralized adding a single counter-ion in the simulation box (Na^+ or Cl^- for negative or positive amino acids, respectively). The effect of this simplification was validated in the positively charged Ace-K-A-NHMe case (type I) by comparing the previous simulations with those neutralized with 150 mM of NaCl (a total of 3 Na^+ and 4 Cl^- ions added) that represents a common used salt concentration for molecular dynamics simulations of proteins.[74–76] Noteworthy, no relevant differences were found between the simulations for the Φ and Ψ dihedral angles distributions as shown in SIF (Figure S5), validating the chosen classical MD protocol.

Shifting now to the quantum chemistry part, we assessed the basis-set size effect by using two different Pople's basis-sets: 6-31+G(d) and the larger 6-311++G(d,p) one. As shown in Figure 2a for the A_{i+1}-A_{i+2}-$NHMe_{i+3}$ peptide in type I conformation, the less time-consuming 6-

31+G(d) basis-set can be safely considered as suitable for the semi-quantitative analysis we are aiming at.

The coupling between the QM and MM parts also deserves to be scrutinized. In particular, the impact of hydrogen bonds between the peptide model and the solvent molecules on average ECD spectra was checked including the closest water molecules in the QM part of the QM/MM system, since this radius parameter was shown in some cases to be critical.[77] The selection of these water molecules was based on the distance to the peptide centre of mass: from 0 (no solvent molecules to the QM-part), up to 5 and 6 Å. Due to timing reasons (the inclusion of several water molecules in the QM region considerably increases the computing time), this test was carried out only for type I β-turn conformation. It can be deduced from Figure 2b that the insertion of explicit solvent molecules in the QM part has not a significant impact on the average ECD spectra. Consequently, all water molecules were then treated as part of the MM sub-system.

Finally, Figure 2c confirms that the three allowed deviations (±10°, ±20°, ±30°) with respect to canonical dihedral angles afford comparable patterns for the average ECD spectra. In the following, only the ±10° case will thus be further discussed in details, consistently with the previous study on β-turns by Woody.[32]

3.2. Electronic circular dichroism spectra

Once more, Ace_i-A_{i+1}-A_{i+2}-$NHMe_{i+3}$ in types I, I', II or II' will be our working case for the analysis of spectroscopic properties. The corresponding average ECD spectra are plotted in Figure 3. In the four cases, there are two ECD bands with similar intensities and opposite signs approximately around 180 nm. In particular, types I and II' possess a common pattern: the first band at minor energies (i.e. high wavelengths) is negative in sign followed by a positive one at higher energies.

On the other hand, types I' and II reveal the opposite pattern: the first band at lower energies is positive in sign followed by a negative one at higher energies. It should be noticed that such a discrimination is not possible from the simple absorption spectra (see Figure S6 in SIF) since all types present a common pattern: a very weak band (ε ~500 $M^{-1}cm^{-1}$) at 220 nm, corresponding to an electric dipole-forbidden n → π* transition, followed by an intense band (ε > 20000 $M^{-1}cm^{-1}$) at ~175 nm, ascribable to a π_{nb} → π* (NV1) electric dipole-allowed transition. This makes absorption spectroscopy fully inadequate for β-turn characterization.

Then, to deepen this shape analysis, we calculated the ratio (*κ*, unitless) between the negative and positive ECD bands intensities, as well as the distance (Δλ, in nm) between the corresponding extrema, as defined by equations 4 and 5:

$$\Delta\lambda = arg\ min\big(\Delta\varepsilon(\lambda)\big) - arg\ max\big(\Delta\varepsilon(\lambda)\big) \quad (4)$$

$$\kappa = min\big(\Delta\varepsilon(\lambda)\big)/max\big(\Delta\varepsilon(\lambda)\big) \quad (5)$$

For example, for Ace_i-A_{i+1}-A_{i+2}-$NHMe_{i+3}$ in type I conformation, the ratio between the first negative and the second positive ECD bands is distributed around 1.0 (see Figure S7 in SIF), and Δλ between the positive and negative ECD bands maxima shows a bi-modal distribution at ~15 nm (majority) and another one at ~ -15 nm (see Figure S8). However, we noticed that the first mode (corresponding to positive values) is prevalently populated. Note that the percentage of ECD spectra that respect the "first negative", types I/II', or "first positive", types I'/II are indicated in Figure 3. Similar conclusions were obtained for the other β-turn conformations (i.e. types I', II and II').

We now inspect the impact of substitution in these "first negative/positive" patterns by modifying the amino-acids in *i+1* or *i+2* positions, leading to the following generic sequence: Ace_i-X_{i+1}-X_{i+2}-$NHMe_{i+3}$. Amino-acids are usually classified into six classes:[78] hydrophobic, polar uncharged, positively charged, aromatic, and special cases. We thus chose one or two

representative amino-acids for each of these subclasses, as detailed in Table 3. For each of them, the influence of the substitution position was investigated. For instance, we considered both Ace$_i$-G-A-NHMe$_{i+3}$ and Ace$_i$-A-G-NHMe$_{i+3}$ (where G denotes glycine). In total, 15 additional peptides were simulated for each of the four retained β-turn structures, resulting in 60 computed spectra.

In a first instance, we now consider non-aromatic side chains. From Figure 4, it can be inferred that non-aromatic amino-acids have an effect on the energetic position and on the intensity of the two opposite-in-sign ECD bands but that, globally, the patterns we found previously with the Ace$_i$-A$_{i+1}$-A$_{i+2}$-NHMe$_{i+3}$ minimal model for each kind of β-turns, were unchanged.

Figure 4 also indicates the percentage of frames whose ECD spectra fulfil either the "first negative" (types I and II') or the "first positive" (types I' and II) patterns previously identified. The mean values of respected averaged ECD pattern in each of the 13 sequences considered here are 66%, 83%, 67% and 78% for types I, I', II and II', respectively. For type II, only 9 sequences respect the mean pattern with percentages greater than 60%. However, the fact that the average ECD spectra for a β-turn conformation calculated using different sequences converge to the same pattern with highly superimposable profiles strongly supports the two ECD patterns found with the minimal peptide model Ace$_i$-A$_{i+1}$-A$_{i+2}$-NHMe$_{i+3}$ ("first negative" and "first positive").

Consequently, non-aromatic side chains do not seem to have a relevant impact on the averaged ECD spectra. This conclusion was actually already stated by Woody,[79] who underlined that among the amino acids with chromophoric side-chains (i.e. phenylalanine, tyrosine, tryptophan, disulfides, histidine, arginine, glutamine, asparagine, glutamic acid, and aspartic acid), only the aromatic and disulfides ones have electric or magnetic dipole transition moments large enough to make appreciable spectral contributions.[79]

In contrast, as stated before, amino-acids with aromatic side-chains (represented here by phenylalanine, F) greatly perturb the patterns found with the Ace_i-A_{i+1}-A_{i+2}-$NHMe_{i+3}$ minimal peptide model, as illustrated by Figure 5. This perturbation is in fact caused by the presence of other chromophores in the system with strong electric transition moments (phenyl in case of phenylalanine). Similar perturbations have been already studied and characterized in several experimental and theoretical works.[41,80–83] Indeed, such a perturbing chromophore is spatially quite close to the three peptide bonds and, at the same time, its excited states (^1L_a and the degenerate states ^1B_b and ^1B_a) are energetically close to the $\pi_{nb} \rightarrow \pi^*$ transitions of amide groups responsible for the exciton interactions.[84] As a consequence, these perturbations do not allow the identification of the previously found "first positive/negative" patterns.

This important shortcoming limits the validity of "first positive/negative" patterns as *unique* reference for types I, I', II and II' β-turns when at least one aromatic amino acid is included in the sequence of a short peptide. In such cases, ECD spectroscopy should certainly be complemented by other spectroscopic methods to univocally determine the conformation of the involved β-turn structures. However, it should be stressed that aromatic amino-acids (Phe, Tyr and Trp) are very **rarely** found (with a probability minor to 1%) in position i+1 and/or i+2 for these types of β-turns,[85] so that this limitation will be fully ineffective in most cases.

Finally, even if we are not dealing with *chirality* determinations (all the amino-acids are indeed L-amino-acids), we qualitatively tried to explain the origin of the "first negative" (characteristic for types I and II') and "first positive" (characteristic for type I' and II) ECD patterns by means of the *non-empirical* exciton chirality rule (ECR).[86] For this purpose, we made three strong approximations: i) the peptide bond is a suitable chromophore for ECR (usually, chromophores with strong $\pi_{nb} \rightarrow \pi^*$ absorption, as benzoate derivatives, are recommended for ECR[87]); ii) the direction of the electric transition dipole moment (TDM) for the $\pi_{nb} \rightarrow \pi^*$ (NV_1) transition for each of the three peptide bonds are assumed to be parallel to the peptide bond and co-planar to

the peptide group[88] (see Figure 6a); iii) only interactions and relative orientations between the electric TDMs between two *consecutive* chromophores are considered.

This last point implies that we have to analyse the relative orientations between the first and the second residues, and between the second and the third amide chromophores, through the definition of pseudo-dihedral angles "*p1*" and "*p2*", represented in Figure 6a. It also follows from this assumption that the interaction between the first and third peptide bonds is not considered. This can be crudely justified by observing that these bonds are far away, and by recalling that the exciton splitting intensity A is inversely proportional to R_{ij}^2, where $\boldsymbol{R}_{ij}$ denotes the interchoromophoric vector between chromophores i and j. More specifically, the amplitude of the exciton ECD couplet A (defined as the difference between the first ECD band at lower energies and the second ECD band at higher energies, i.e. $\Delta\varepsilon_1 - \Delta\varepsilon_2$) reads:[86,89]

$$A \propto \boldsymbol{R}_{ij} \cdot (\boldsymbol{\mu}_{i0a} \times \boldsymbol{\mu}_{j0a}) V_{ij} = D_{i0a} D_{j0a} R_{ij}^{-2} \boldsymbol{e}_{\boldsymbol{ij}} \cdot (\boldsymbol{e}_i \times \boldsymbol{e}_j) \{\boldsymbol{e}_i \cdot \boldsymbol{e}_j - 3(\boldsymbol{e}_i \cdot \boldsymbol{e}_{ij})(\boldsymbol{e}_j \cdot \boldsymbol{e}_{ij})\}, \qquad (6)$$

where $\boldsymbol{\mu}_{i0a}$ and $\boldsymbol{\mu}_{j0a}$ are the electric transition dipole moments, D_{i0a} and D_{j0a} denote the electric dipole strength (squares of $\boldsymbol{\mu}_{i0a}$ and $\boldsymbol{\mu}_{j0a}$) of chromophores i and j for the transition $0 \rightarrow a$ (the $\pi_{nb} \rightarrow \pi^*$ NV1 transition in this case), while $\boldsymbol{e}_{\boldsymbol{ij}}$, $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are the unit vectors of $\boldsymbol{R}_{ij}$, $\boldsymbol{\mu}_{i0a}$ and $\boldsymbol{\mu}_{j0a}$, respectively.

Considering type I and type II' (belonging to the "first negative" pattern), we found (see Figure 7) that the values for the pseudo-dihedral angles between consecutives amides bonds *p1* and *p2* are *both* negative (see Figure 6b for the adopted sign convention) along the whole *restrained* MD simulation. If the electric TDMs (for the $\pi_{nb} \rightarrow \pi^*$ transitions) are assumed to lie along the peptides bonds, a clear correspondence between the sign of these dihedral angles and the sign of the first ECD band at lower energies in the region of the $\pi_{nb} \rightarrow \pi^*$ transition is observed. On the other hand, if we consider type I' and type II ("first positive" pattern), the *p1* and *p2* values are

both positives (Figure 7), a fact that is reflected on the sign of the first (positive) ECD band at lower energies.

These observations can be qualitatively rationalized by the ECM[86] that states that the sign of the first band at lower energy is determined by the dihedral angles between the electric TDMs between two interacting chromophores. We have thus *"negative-exciton chirality"* when the dihedral angle between two electric TDMs is negative, resulting in a negative ECD band at lower energies followed by a positive band at higher energies of similar intensity (pattern for types I and II'), while there is *"positive-exciton chirality"* when the dihedral angle between two electric TDMs is positive, resulting in a positive ECD band at longer wavelengths followed by a negative band at higher energies of similar intensity (patterns for types I' and II).

4. Conclusions

In this paper, state-of-the-art time-dependent density functional theory calculations coupled with polarizable embedding model and with molecular dynamics simulations showed that two different electric circular dichroism patterns can be identified within four main β-turns types (namely I, I', II and II') in solution when no aromatic amino-acids are present in the peptide sequence. In particular, a "first negative" pattern can be associated with types I and II', while a "first positive" pattern is associated with types I' and II. This is in contrast with the seminal work by Woody[32] (in which types I and II presented the same pattern, coined class "B"), but in agreement with the work by Sathyanarayana and Applequist,[36] which are both semi-empirical calculations. Besides, these patterns were qualitatively rationalized by the exciton chirality method. These results thus pave the way towards a better theoretical and experimental discrimination of β-turns in peptides.

Acknowledgments

We gratefully acknowledge LABEX SynOrg (Ajouter n° projet), the Région Normandie and the European Union (Ajouter n° projet) for funding. The "Centre Régional Informatique et d'Applications Numériques de Normandie – CRIANN" (projects n° 2018019 and n° 2018012) and "Institut du Développement et des Ressources en Informatique Scientifique - IDRIS" (DARI project n° 100825) are thanked for computational support. A. B. warmly thanks Prof. Jacob Kongsted and Peter Reinholdt (University of Southern Denmark, SDU, Denmark) for initiating him to the polarizable embedding model.

References

1 C. M. Venkatachalam, Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units, *Biopolymers*, 1968, **6**, 1425–1436.

2 J. S. Richardson, The anatomy and taxonomy of protein structure, *Adv. Protein Chem.*, 1981, **34**, 167–339.

3 P. N. Lewis, F. A. Momany and H. A. Scheraga, Folding of Polypeptide Chains in Proteins: A Proposed Mechanism for Folding, *Proc. Natl. Acad. Sci.*, 2006, **68**, 2293–2297.

4 I. D. Kuntz, Protein folding, *J. Am. Chem. Soc.*, 1972, **94**, 4009–4012.

5 J. L. Crawford, W. N. Lipscomb and C. G. Schellman, The Reverse Turn as a Polypeptide Conformation in Globular Proteins, *Proc. Natl. Acad. Sci.*, 1973, **70**, 538–542.

6 P. Y. Chou and G. D. Fasman, β-turns in proteins, *J. Mol. Biol.*, 1977, **115**, 135–175.

7 G. D. Rose, L. M. Glerasch and J. A. Smith, Turns in Peptides and Proteins, *Adv. Protein Chem.*, 1985, **37**, 1–109.

8 L. L. Videau, W. B. Arendall and J. S. Richardson, The cis-Pro touch-turn: A rare motif preferred at functional sites, *Proteins Struct. Funct. Genet.*, 2004, **56**, 298–309.

9 J. D. A. Tyndall, B. Pfeiffer, G. Abbenante and D. P. Fairlie, Over One Hundred Peptide-Activated G Protein-Coupled Receptors Recognize Ligands with Turn Structure, *Chem. Rev.*, 2005, **105**, 793–826.

10 G. Ruiz-Gómez, J. D. A. Tyndall, B. Pfeiffer, G. Abbenante and D. P. Fairlie, Update 1 of: Over One Hundred Peptide-Activated G Protein-Coupled Receptors Recognize Ligands with Turn Structure, *Chem. Rev.*, 2010, **110**, PR1–PR41.

11 M. E. Johnson and M. Kahn, in *Biotechnology and Pharmacy*, eds. J. M. Pezzuto, M. E. Johnson and H. R. Manasse, Springer Netherlands, Dordrecht, 1993, pp. 366–378.

12 P. Grieco, A. Carotenuto, R. Patacchini, C. A. Maggi, E. Novellino and P. Rovero, Design, Synthesis, Conformational Analysis, and Biological Studies of Urotensin-II Lactam Analogues, *Bioorg. Med. Chem.*, 2002, **10**, 3731–3739.

13 P. N. Lewis, F. A. Momany and H. A. Scheraga, Chain reversals in proteins, *Biochim. Biophys. Acta - Protein Struct.*, 1973, **303**, 211–229.

14 R. Huber and W. Steigemann, Two cis-prolines in the Bence-Jones protein Rei and the cis-pro-bend, *FEBS Lett.*, 1974, **48**, 235–237.

15 C. M. Wilmot and J. M. Thornton, Analysis and prediction of the different types of β-turn in proteins, *J. Mol. Biol.*, 1988, **203**, 221–232.

16 E. G. Hutchinson and J. M. Thornton, A revised set of potentials for β-turn formation in proteins, *Protein Sci.*, 1994, **3**, 2207–2216.

17 K.-C. Chou, Prediction of Tight Turns and Their Types in Proteins, *Anal. Biochem.*, 2000, **286**, 1–16.

18 A. G. De Brevern, Extension of the classical classification of β-turns, *Sci. Rep.*, 2016, **6**, 33191.

19 M. Shapovalov, S. Vucetic and R. L. Dunbrack, A new clustering and nomenclature for beta turns derived from high-resolution protein structures, *PLOS Comput. Biol.*, 2019, **15**, e1006844.

20 G. Holzwarth and P. Doty, The Ultraviolet Circular Dichroism of Polypeptides 1, *J. Am. Chem. Soc.*, 1965, **87**, 218–228.

21 R. Townend, T. F. Kumosinski, S. N. Timasheff, G. D. Fasman and B. Davidson, The circular dichroism of the β structure of poly-l-lysine, *Biochem. Biophys. Res. Commun.*, 1966, **23**, 163–169.

22 J. Reed and T. A. Reed, A Set of Constructed Type Spectra for the Practical Estimation of Peptide Secondary Structure from Circular Dichroism, *Anal. Biochem.*, 1997, **254**, 36–40.

23 P. K. Sarkar and P. Doty, The optical rotatory properties of the beta-configuration in

polypeptides and proteins., *Proc. Natl. Acad. Sci.*, 1966, **55**, 981–989.

24 A. McPherson, *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press, 1999.

25 M. W. Parker, Protein Structure from X-Ray Diffraction, *J. Biol. Phys.*, 2003, **29**, 341–362.

26 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, 1986.

27 K. D. Kopple, T. J. Schamper and A. Go, Conformation of cyclic peptides. VIII. Cyclic hexapeptides containing the L-pro-D-phe sequence, *J. Am. Chem. Soc.*, 1974, **96**, 2597–2605.

28 A. Perczel, M. Hollósi, B. M. Foxman, G. D. Fasman and G. D. Fasman, Conformational Analysis of Pseudocyclic Hexapeptides Based on Quantitative Circular Dichroism (CD), NOE, and X-ray Data. The Pure CD Spectra of Type I and Type II β-Turns, *J. Am. Chem. Soc.*, 1991, **113**, 9772–9784.

29 L. M. Gierasch, C. M. Deber, V. Madison, C.-H. Niu and E. R. Blout, Conformations of (X-L-Pro-Y)2 cyclic hexapeptides. Preferred β-turn conformers and implications for .beta. turns in proteins, *Biochemistry*, 1981, **20**, 4730–4738.

30 L. Whitmore and B. A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases, *Biopolymers*, 2008, **89**, 392–400.

31 N. J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, *Nat. Protoc.*, 2006, **1**, 2876–2890.

32 R. W. Woody, in *Peptides, Polypeptides, and Proteins*, eds. E. R. Blout, F. A. Bovey, M. Goodman and N. Lotan, John Wiley & Sons, New York, 1974, pp. 338–350.

33 R. W. Woody, in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed. G. D. Fasman, 1996, pp. 25–67.

34 P. M. Bayley, E. B. Nielsen and J. A. Schellman, Rotatory properties of molecules

containing two peptide groups: theory, *J. Phys. Chem.*, 1969, **73**, 228–243.

35 R. W. Woody, Improved Calculation of the nπ* Rotational Strength in Polypeptides, *J. Chem. Phys.*, 1968, **49**, 4797–4806.

36 B. K. Sathyanarayana and J. Applequist, Theoretical π-π* absorption and circular dichroic spectra of β-turn model peptides, *Int. J. Pept. Protein Res.*, 1986, **27**, 86–94.

37 C. T. Chang, C. S. C. Wu and J. T. Yang, Circular dichroic analysis of protein conformation: Inclusion of the β-turns, *Anal. Biochem.*, 1978, **91**, 13–31.

38 L. A. Compton and W. C. Johnson, Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication, *Anal. Biochem.*, 1986, **155**, 155–167.

39 S. Brahms, J. Brahms, G. Spach and A. Brack, Identification of β-turns and unordered conformations in polypeptide chains by vacuum ultraviolet circular dichroism, *Proc. Natl. Acad. Sci.*, 1977, **74**, 3208–3212.

40 S. Brahms and J. Brahms, Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism, *J. Mol. Biol.*, 1980, **138**, 149–178.

41 C. A. Bush, S. K. Sarkar and K. D. Kopple, Circular Dichroism of β Turns in Peptides and Proteins, *Biochemistry*, 1978, **17**, 4951–4954.

42 M. Crisma, G. D. Fasman, H. Balaram and P. Balaram, Peptide models for β-turns. A circular dichroism study, *Int. J. Pept. Protein Res.*, 1984, **23**, 411–419.

43 M. Hollósi, K. E. Kövér, S. Holly, L. Radics and G. D. Fasman, β-Turns in bridged proline-containing cyclic peptide models, *Biopolymers*, 1987, **26**, 1555–1572.

44 J. Bandekar, D. J. Evans, S. Krimm, S. J. Leach, S. Lee, J. R. McQuie, E. Minasian, G. Nemethy, M. S. Pottle, H. A. Scheraga, E. R. Stimson and R. W. Woody, Conformations of cyclo(L-alanyl-L-alanyl-ε-aminocaproyl) and of cyclo(L-alany1-D-alanyl-ε-aminocaproyl); cyclized dipeptide models for specific types of β-bends, *Int. J. Pept.*

Protein Res., 1982, **19**, 187–205.

45 B. Drouillat, C. Peggion, B. Biondi, K. Wright, F. Couty, M. Crisma, F. Formaggio and C. Toniolo, Heterochiral Ala/(αMe)Aze sequential oligopeptides: Synthesis and conformational study, *J. Pept. Sci.*, 2019, **25**, e3165.

46 A. Perczel, M. Hollósi, G. Tusnády and G. D. Fasman, Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins., *Protein Eng.*, 1991, **4**, 669–79.

47 A. Perczel and G. D. Fasman, Quantitative analysis of cyclic β-turn models, *Protein Sci.*, 1992, **1**, 378–395.

48 J. M. Olsen, K. Aidas and J. Kongsted, Excited states in solution through polarizable embedding, *J. Chem. Theory Comput.*, 2010, **6**, 3721–3734.

49 M. S. Nørby, J. M. H. Olsen, C. Steinmann and J. Kongsted, Modeling Electronic Circular Dichroism within the Polarizable Embedding Approach, *J. Chem. Theory Comput.*, 2017, **13**, 4442–4451.

50 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis., *J. Comput. Chem.*, 2004, **25**, 1605–12.

51 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

52 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.

53 P. Mark and L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, *J. Phys. Chem. A*, 2001, **105**, 9954–9960.

54 M. Parrinello and A. Rahman, Polymorphic transitions in single crystals: A new

molecular dynamics method, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

55 K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen, J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I. M. Høyvik, M. F. Iozzi, B. Jansík, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawlowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Sałek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski and H. Ågren, The Dalton quantum chemistry program system, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2014, **4**, 269–284.

56 T. Yanai, D. P. Tew and N. C. Handy, A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP), *Chem. Phys. Lett.*, 2004, **393**, 51–57.

57 R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, *J. Chem. Phys.*, 1980, **72**, 650–654.

58 C. Steinmann, P. Reinholdt, M. S. Nørby, J. Kongsted and J. M. H. Olsen, Response properties of embedded molecules through the polarizable embedding model, *Int. J. Quantum Chem.*, 2018, **119**, e25717.

59 P. Altoè, M. Stenta, A. Bottoni and M. Garavelli, A tunable QM/MM approach to

chemical reactivity, structure and physico-chemical properties prediction, *Theor. Chem. Acc.*, 2007, **118**, 219–240.

60 O. Weingart, A. Nenov, P. Altoè, I. Rivalta, J. Segarra-Martí, I. Dokukina and M. Garavelli, COBRAMM 2.0 — A software interface for tailoring molecular electronic structure calculations and running nanoscale (QM/MM) simulations, *J. Mol. Model.*, 2018, **24**, 271.

61 J. M. H. Olsen and J. Kongsted, in *Advances in Quantum Chemistry*, 2011, pp. 107–143.

62 L. J. Nåbo, N. H. List, C. Steinmann and J. Kongsted, Computational Approach to Evaluation of Optical Properties of Membrane Probes, *J. Chem. Theory Comput.*, 2017, **13**, 719–726.

63 N. H. List, H. J. A. Jensen and J. Kongsted, Local electric fields and molecular properties in heterogeneous environments through polarizable embedding, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10070–10080.

64 A. Bonvicini, P. Reinholdt, V. Tognetti, L. Joubert, D. Wüstner and J. Kongsted, Rational design of novel fluorescent analogues of cholesterol: a “step-by-step” computational study, *Phys. Chem. Chem. Phys.*, 2019, **21**, 15487-15503.

65 M. T. P. Beerepoot, A. H. Steindal, N. H. List, J. Kongsted and J. M. H. Olsen, Averaged Solvent Embedding Potential Parameters for Multiscale Modeling of Molecular Properties, *J. Chem. Theory Comput.*, 2016, **12**, 1684–1695.

66 J. M. H. Olsen, PyFraME: Python tools for Fragment-based Multiscale Embedding calculations (version 0.1.0), DOI:10.5281/ZENODO.293765.

67 Gaussian Inc, Creating UV/Visible Plots from the Results of Excited States Calculations.

68 P. J. Stephens and N. Harada, *Chirality*, 2010, **22**, 229–233.

69 M. S. Searle, D. H. Williams and L. C. Packman, A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native β-hairpin, *Nat. Struct. Biol.*, 1995, **2**, 999–1006.

70 Y. Mirassou, C. M. Santiveri, M. J. Pérez de Vega, R. González-Muñiz and M. A. Jiménez, Disulfide Bonds versus Trp…Trp Pairs in Irregular β-Hairpins: NMR Structure of Vammin Loop 3-Derived Peptides as a Case Study, *ChemBioChem*, 2009, **10**, 902–910.

71 T. Tancredi, A. Pastore, S. Salvadori, V. Esposito and P. A. Temussi, Interaction of sweet proteins with their receptor, *Eur. J. Biochem.*, 2004, **271**, 2231–2240.

72 M. S. Searle, Peptide models of protein β-sheets: design, folding and insights into stabilising weak interactions, *J. Chem. Soc. Perkin Trans. 2*, 2001, 1011–1020.

73 B. L. Kier, I. Shu, L. A. Eidenschink and N. H. Andersen, Stabilizing capping motif for -hairpins and sheets, *Proc. Natl. Acad. Sci.*, 2010, **107**, 10466–10471.

74 G. T. Ibragimova and R. C. Wade, Importance of Explicit Salt Ions for Protein Stability in Molecular Dynamics Simulation, *Biophys. J.*, 1998, **74**, 2906–2911.

75 K. El Hage, F. Hédin, P. K. Gupta, M. Meuwly and M. Karplus, Valid molecular dynamics simulations of human hemoglobin require a surprisingly large box size, *Elife*, DOI:10.7554/eLife.35560.

76 Y. Li, R. Sun, H. Liu and H. Gong, Molecular dynamics study of ion transport through an open model of voltage-gated sodium channel, *Biochim. Biophys. Acta - Biomembr.*, 2017, **1859**, 879–887.

77 J. Sjöqvist, M. Linares, K. V. Mikkelsen and P. Norman, QM/MM-MD Simulations of Conjugated Polyelectrolytes: A Study of Luminescent Conjugated Oligothiophenes for Use as Biophysical Probes, *J. Phys. Chem. A*, 2014, **118**, 3419–3428.

78 D. L. Nelson and M. M. Cox, *Lehninger Principles Of Biochemistry 5/E*, NA, 2008.

79 N. Sreerama and R. W. Woody, in *Methods in Enzymology, Numerical Mehthods part D*, 2004, pp. 318–351.

80 C. Krittanai and W. C. Johnson, Correcting the Circular Dichroism Spectra of Peptides for Contributions of Absorbing Side Chains, *Anal. Biochem.*, 1997, **253**, 57–64.

81 A. Chakrabartty, T. Kortemme, S. Padmanabhan and R. L. Baldwin, Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities, *Biochemistry*, 1993, **32**, 5560–5565.

82 R. W. Woody, Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins, *Eur. Biophys. J.*, 1994, **23**, 253–262.

83 M. C. Manning and R. W. Woody, Theoretical study of the contribution of aromatic side chains to the circular dichroism of basic bovine pancreatic trypsin inhibitor, *Biochemistry*, 1989, **28**, 8609–8613.

84 N. Berova, P. L. Polavarapu, K. Nakanishi and R. W. Woody, *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012, vol. 1.

85 K. Guruprasad and S. Rajkumar, β- and γ-turns in proteins revisited: A new set of amino acid turn-type dependent positional preferences and potentials, *J. Biosci.*, 2000, **25**, 143–156.

86 N. Harada, K. Nakanishi and N. Berova, in *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012, pp. 115–166.

87 N. Harada and K. Nakanishi, *Circular Dichroic Spectroscopy: Exciton Coupling in Organic Stereochemistry*, University Science Books, 1983.

88 N. Sreerama and R. W. Woody, in *Circular dichroism: principles and applications*, Wiley-VCH, 2000, p. 877.

89 N. Harada, S.-M. L. Chen and K. Nakanishi, Quantitative definition of exciton chirality and the distant effect in the exciton chirality method, *J. Am. Chem. Soc.*, 1975, **97**, 5345–5352.

TABLE 1: β-turns classification based on the Φ_{i+1}, Ψ_{i+1}, Φ_{i+2} and Ψ_{i+2} dihedral angles canonical values according to De Brevern (ref. [18]). Types in bold are those considered in this work.

Type	Φ_{i+1}	Ψ_{i+1}	Φ_{i+2}	Ψ_{i+2}
I	**-60°**	**-30°**	**-90°**	**0°**
I'	**+60°**	**+30°**	**+90°**	**0°**
II	**-60°**	**+120°**	**+80°**	**0°**
II'	**+60°**	**-120°**	**-80°**	**0°**
IV_1	-120°	+130°	+55°	+41°
IV_2	-85°	-15°	-125°	+55
IV_3	-71°	-30°	-72°	-47°
IV_4	-97°	-2°	-117°	-11°
VIa1	-60°	+120°	-90°	0°
VIa2	-120°	+120	-60°	0°
VIb	-135°	+135°	-75°	160°
VIII	-60°	-30°	-120°	120°

TABLE 2: Classification scheme for calculated ECD spectra by Woody (ref. [32]). The classes were established in function of the number and of the position of maxima and minima.

Class	Sign	Wavelength range [nm]
A	-	210-220
	+	195-200
	-	<190
B	-	>220
	+	200-210
	-	<190
C	-	200-210
	+	180-195
D	-	>225
	+	210-220
	-	190-200
	+	<190

TABLE 3: Amino-acids classification based on the physic-chemical properties of their side-chain. Amino-acids in bold indicate the representative member chosen for the present study.

Amino-acid family	Components
Hydrophobic	**Ala**, **Val**, Ile, Leu, Met
Polar uncharged	**Ser**, Thr, Asn, Gln
Positively charged	**Lys**, Arg, His
Negatively charged	**Asp**, Glu
Special cases	**Cys**, **Gly**, Pro
Aromatic	**Phe**, Tyr, Trp

Figure captions

Figure 1. a) Chemical structure of the Ace-A-A-NHMe peptide used in this work as minimal model (H in white, C in grey, N in blue, O in red). b) Direction of the n → π* magnetic dipole transition moment and NV1 ($\pi_{nb} \rightarrow \pi^*$) electric dipole transition moment with respect to the peptide bond.

Figure 2. a) Averaged ECD spectra of type I β-turn using different basis-sets. b) Effect of explicit quantum chemical water molecules on the averaged ECD spectra of type I using the 6-31+G(d) basis-set. c) Averaged ECD spectra for type I using various deviations for the Φ_{i+1}, Ψ_{i+1}, Φ_{i+2} and Ψ_{i+2} dihedral angles (with respect to the canonical values collected in Table 1). CAM-B3LYP level of theory.

Figure 3. Ace-A-A-NHMe averaged ECD spectra for the four selected β-turns types (types I, I', II and II'). CAM-B3LYP/6-31+G(d) level of theory.

Figure 4. Effect of residue substitutions in position *i+1* (straight lines) or *i+2* (dotted lines) on averaged ECD spectra compared to the Ace-A-A-NHMe peptide (black lines) for each amino-acid family (see Table 2). CAM-B3LYP/6-31+G(d) level of theory.

Figure 5. Effect of an aromatic residue (Phe) substitution in position *i+1* (red lines) or *i+2* (blue lines) on averaged ECD spectra compared to Ace-A-A-NHMe (black lines). CAM-B3LYP/6-31+G(d) level of theory.

Figure 6. a) The minimal β-turn model with the electric transition dipole moment TDM for the $\pi_{nb} \rightarrow \pi^*$ (NV_1) transition, approximated to lie along the C-N peptide bond (H in white, C in grey, N in blue, O in red). *p1* (C_i-N_{i+1}-C_{i+1}-N_{i+2}) and *p2* (C_{i+1}-N_{i+2}-C_{i+2}-N_{i+3}) are the pseudo-dihedral angles between consecutives peptides bonds. b) Definitions of negative

(counterclockwise) and positive (clockwise twist) dihedral angles between atoms 1, 2, 3 (hidden in green) and 4.

Figure 7. Distribution of pseudo-angles "*p1*" (green) and "*p2*" (violet) over the molecular dynamics production for Ace-A-A-NHMe. OPLS-AA and TIP3P force fields used.

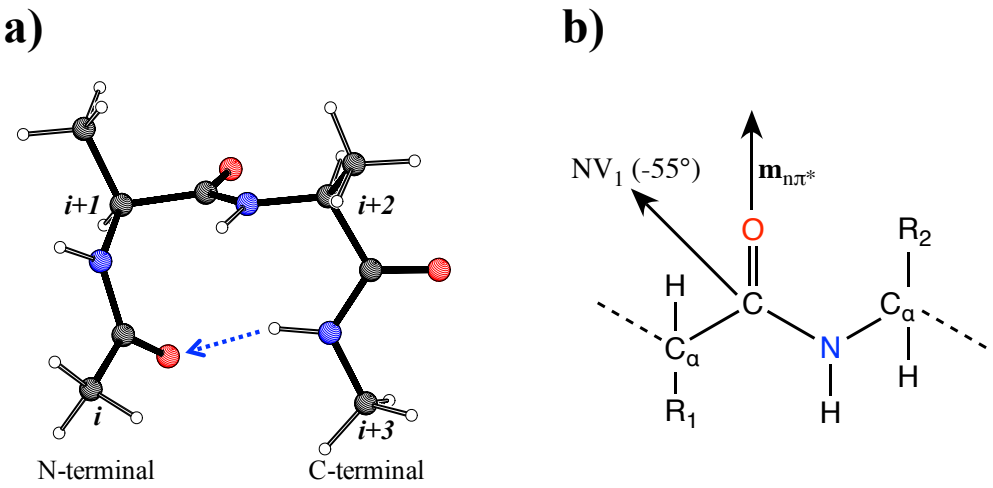

Figure 1

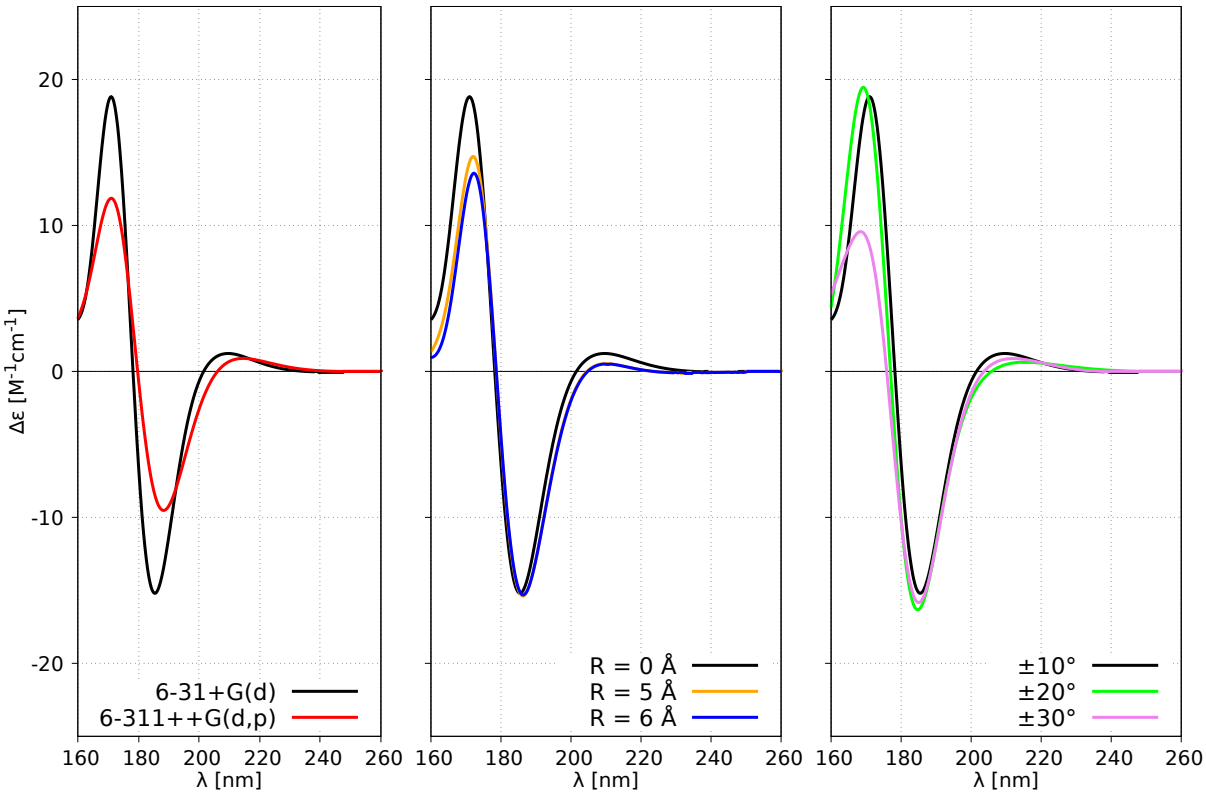
20
10
0
-10
-20
Δε [M-1cm-1]
6-31+G(d)
6-311++G(d,p)
R = 0 Å
R = 5 Å
R = 6 Å
±10°
±20°
±30°
160 180 200 220 240 260
λ [nm]

Figure 2

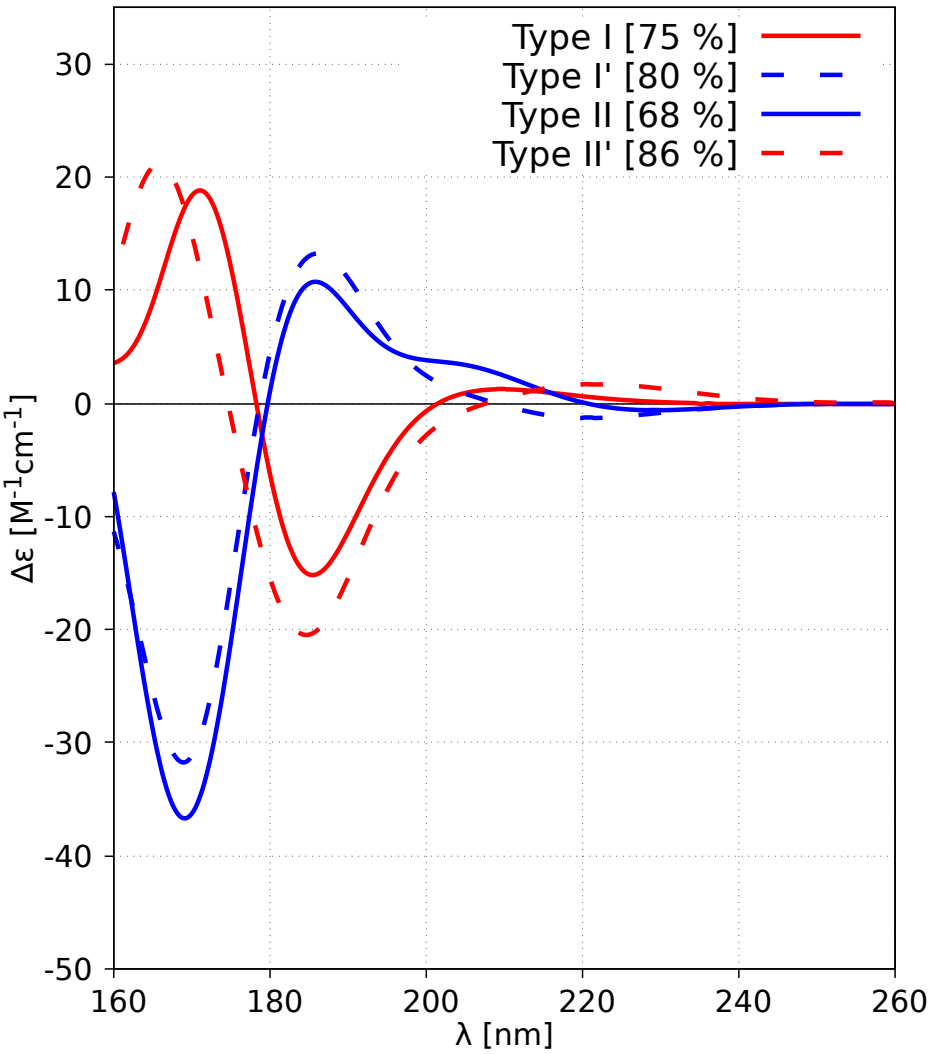
Type I [75 %]
Type I' [80 %]
Type II [68 %]
Type II' [86 %]
30
20
10
0
-10
-20
-30
-40
-50
Δε [M-1cm-1]
160
180
200
220
240
260
λ [nm]

Figure 3

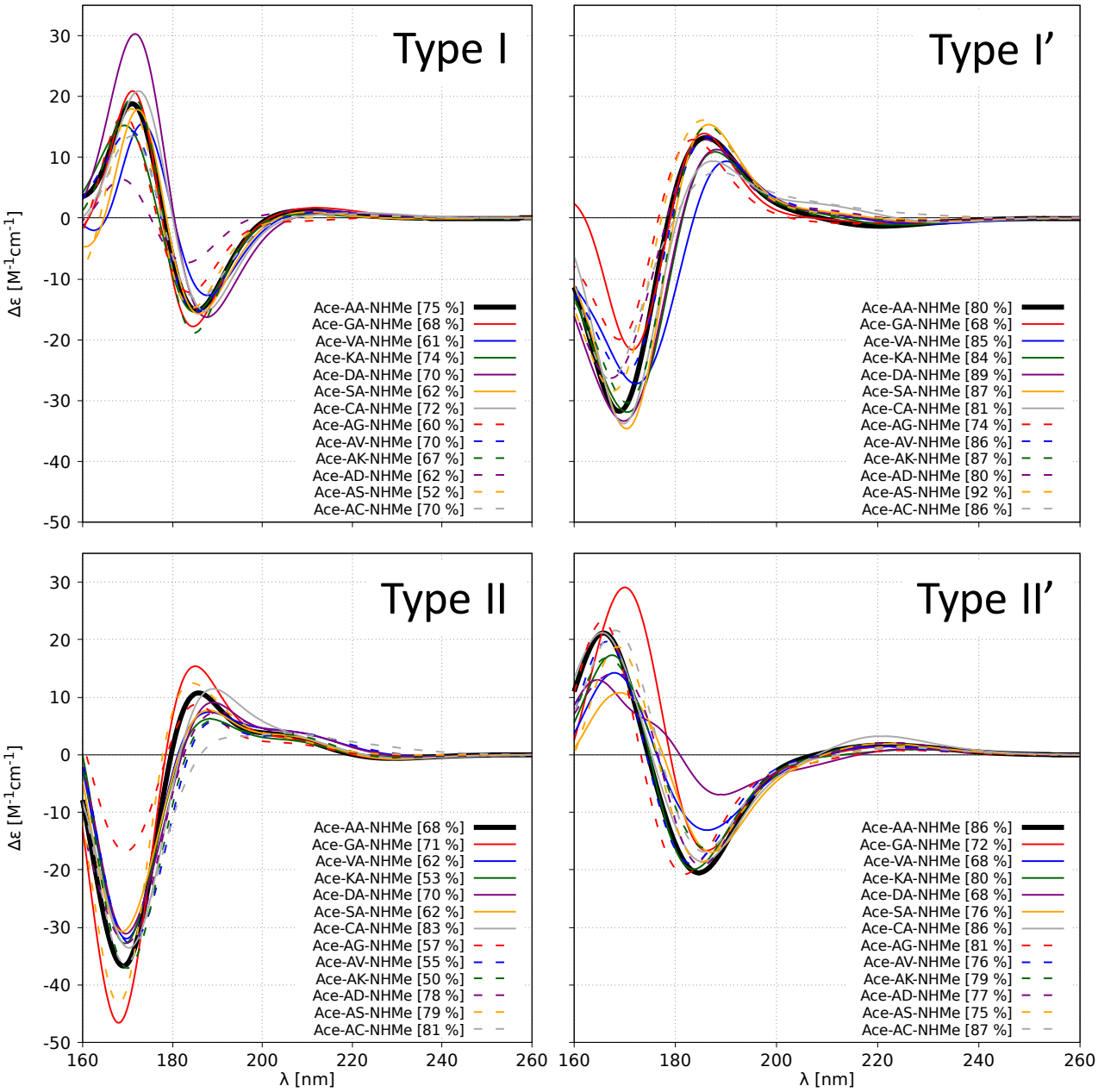
Type I
Ace-AA-NHMe [75 %]
Ace-GA-NHMe [68 %]
Ace-VA-NHMe [61 %]
Ace-KA-NHMe [74 %]
Ace-DA-NHMe [70 %]
Ace-SA-NHMe [62 %]
Ace-CA-NHMe [72 %]
Ace-AG-NHMe [60 %]
Ace-AV-NHMe [70 %]
Ace-AK-NHMe [67 %]
Ace-AD-NHMe [62 %]
Ace-AS-NHMe [52 %]
Ace-AC-NHMe [70 %]
Type I'
Ace-AA-NHMe [80 %]
Ace-GA-NHMe [68 %]
Ace-VA-NHMe [85 %]
Ace-KA-NHMe [84 %]
Ace-DA-NHMe [89 %]
Ace-SA-NHMe [87 %]
Ace-CA-NHMe [81 %]
Ace-AG-NHMe [74 %]
Ace-AV-NHMe [86 %]
Ace-AK-NHMe [87 %]
Ace-AD-NHMe [80 %]
Ace-AS-NHMe [92 %]
Ace-AC-NHMe [86 %]
Type II
Ace-AA-NHMe [68 %]
Ace-GA-NHMe [71 %]
Ace-VA-NHMe [62 %]
Ace-KA-NHMe [53 %]
Ace-DA-NHMe [70 %]
Ace-SA-NHMe [62 %]
Ace-CA-NHMe [83 %]
Ace-AG-NHMe [57 %]
Ace-AV-NHMe [55 %]
Ace-AK-NHMe [50 %]
Ace-AD-NHMe [78 %]
Ace-AS-NHMe [79 %]
Ace-AC-NHMe [81 %]
Type II'
Ace-AA-NHMe [86 %]
Ace-GA-NHMe [72 %]
Ace-VA-NHMe [68 %]
Ace-KA-NHMe [80 %]
Ace-DA-NHMe [68 %]
Ace-SA-NHMe [76 %]
Ace-CA-NHMe [86 %]
Ace-AG-NHMe [81 %]
Ace-AV-NHMe [76 %]
Ace-AK-NHMe [79 %]
Ace-AD-NHMe [77 %]
Ace-AS-NHMe [75 %]
Ace-AC-NHMe [87 %]
Δε [M-1cm-1]
λ [nm]

Figure 4

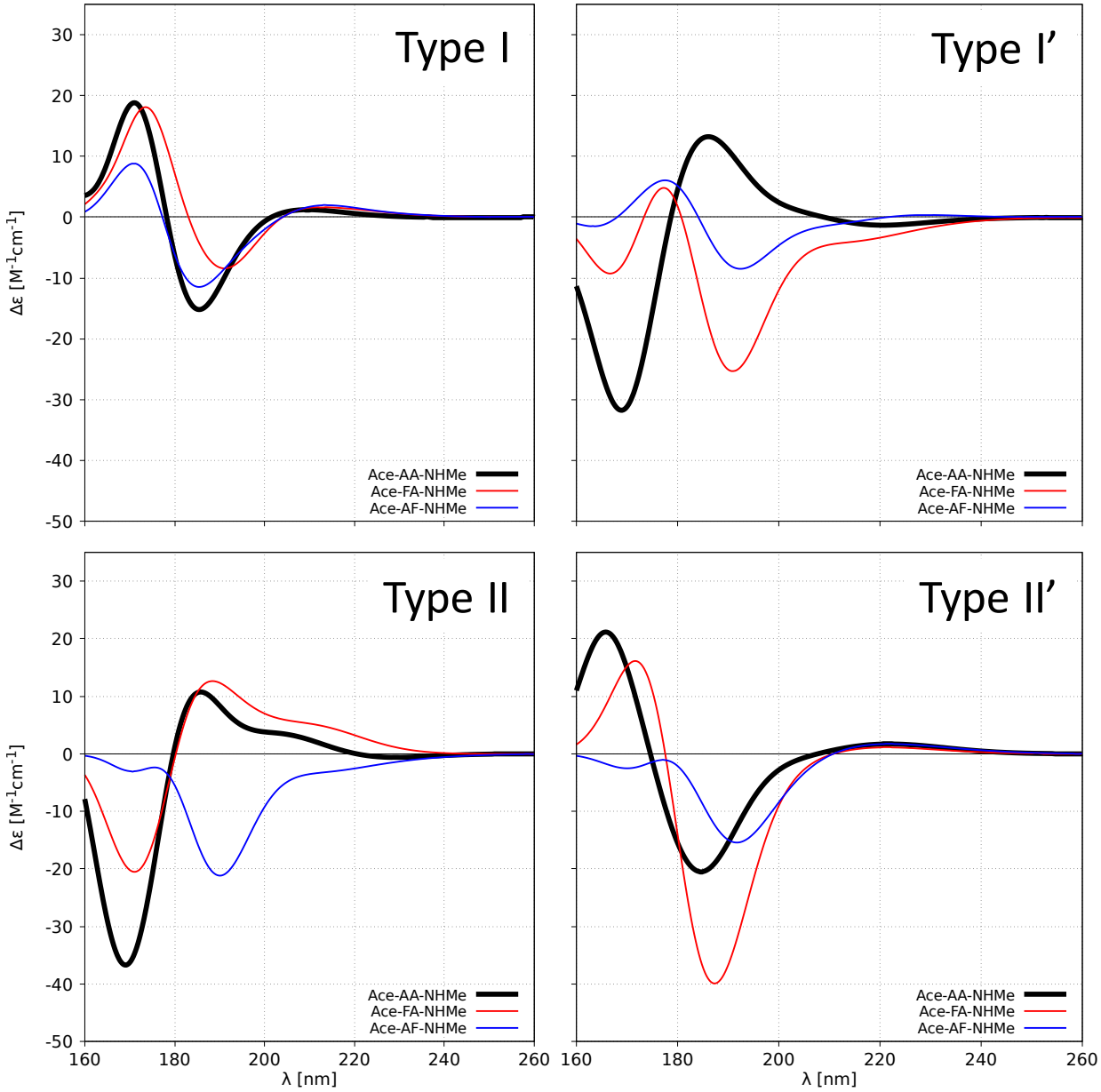
Type I
Type I'
Type II
Type II'
Ace-AA-NHMe
Ace-FA-NHMe
Ace-AF-NHMe
Δε [M-1cm-1]
λ [nm]
30
20
10
0
-10
-20
-30
-40
-50
160
180
200
220
240
260

Figure 5

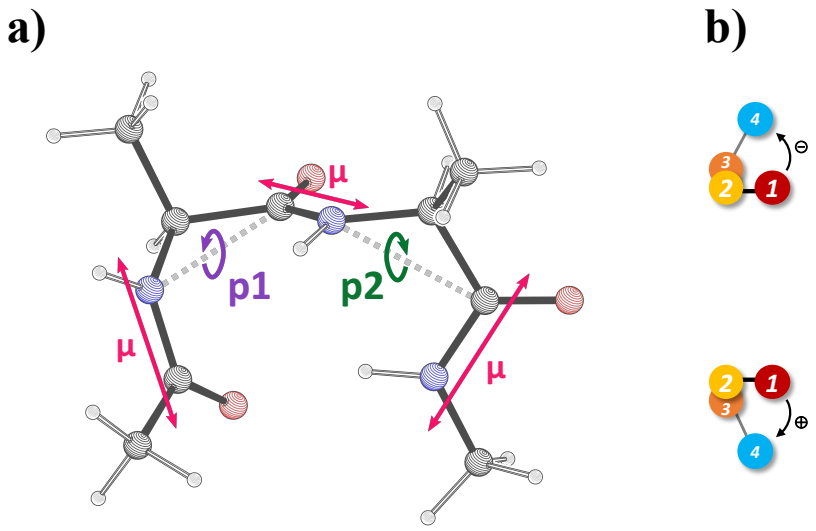

Figure 6

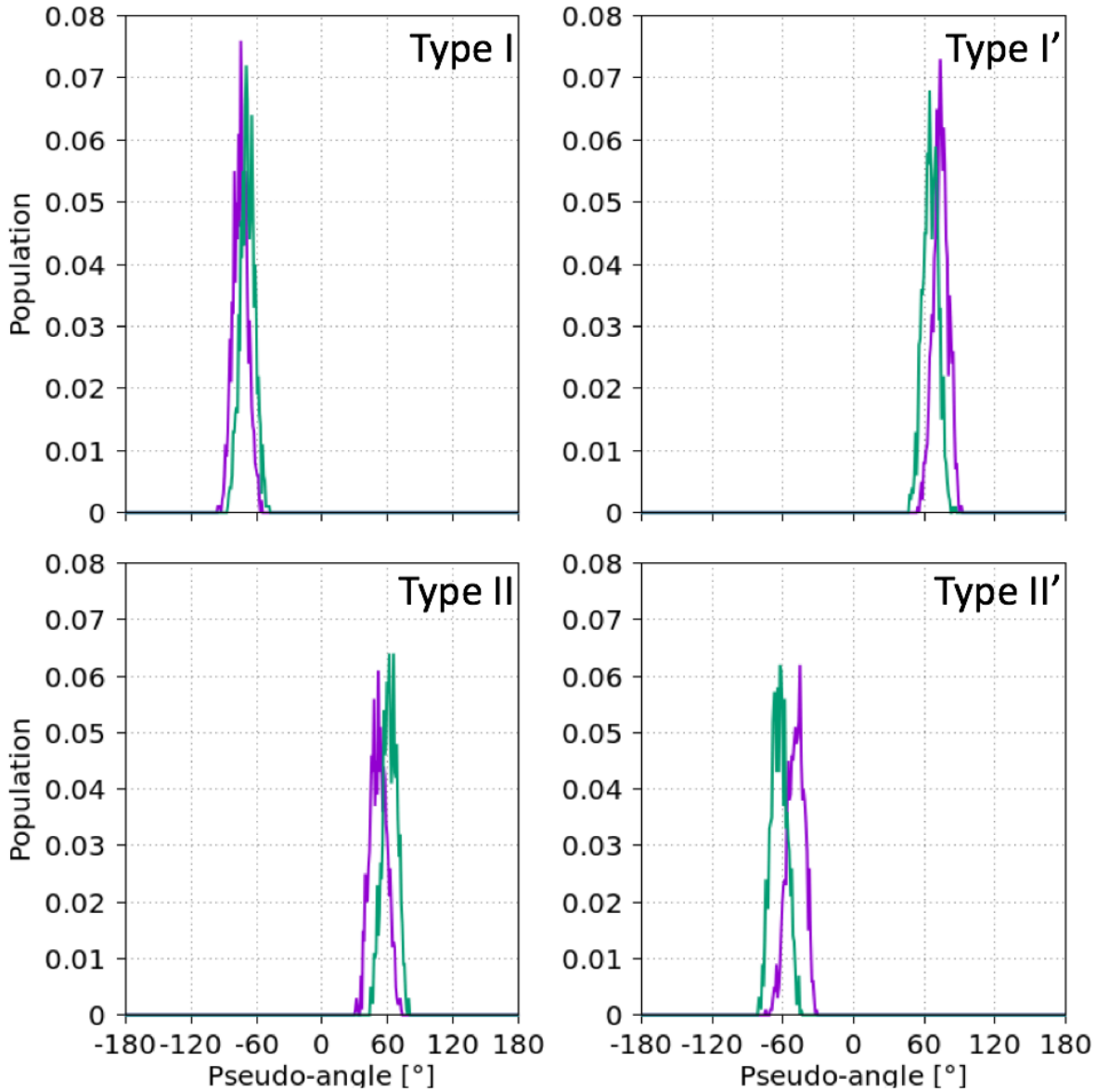
Type I
Type I'
Type II
Type II'
Population
Pseudo-angle [°]
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-180 -120 -60 0 60 120 180

Figure 7