RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Theoretical Study of Two-Photon Circular Dichroism on Molecular Structures Simulating Aromatic Amino Acid Residues in Proteins with Secondary Structures

Yuly Vesga,[1] Carlos Diaz,[1] Florencio E. Hernandez[1,2,*]

[1]Department of Chemistry and, [2]The College of Optics and Photonics, CREOL, University of Central Florida, P. O. Box 162366, Orlando, Florida 32816-2366, USA

** Corresponding Authors:* florencio.hernandez@ucf.edu

Abstract

Herein, we report on the calculation and the comparative analysis of the theoretical two-photon circular dichroism (TPCD) spectra of L-histidine (His), L-phenylalanine (Phe), and L-tyrosine (Tyr) simulating residues in proteins with secondary structures (α–helix, β-strand and random coil), down to the far-UV region (FUV). This work exposes unique signatures in the FUV for each conformer in each configuration. The outcomes of this research show how FUV-TPCD can be used to study peptide and protein structures in a region never evaluated before but packed with important structural information.

1. Introduction

Electronic circular dichroism (ECD) has played a major role in the study of the conformational and physical–chemical properties of optically active biomolecules such as peptides and proteins over the last several years[1, 2]. Determining relevant features of the structure of these biomolecules and the exchange between them and with their environment is possible by the presence of chromophores that can convey a clear ECD signal. Distinctive absorption bands (n-π* and π-π* transitions) in the near-ultraviolet (NUV) and far-ultraviolet (FUV) regions are provided by peptide bonds of secondary protein structures in α-helix, β-strand and random coil configurations[3-5]. Aromatic amino acid side-chains of peptides in the same spectral region produce a signal that offers important information about the rigidity and environment of these structures[6-8]. Even though ECD is a trustworthy technique for the study of proteins in the UV, this method presents some inevitable underlying weaknesses, i.e. the contribution of aromatic amino acid side-chain chromophores to ECD bands in the NUV inhibits the analysis of the secondary structure of the main-chain of peptides and proteins [9].

In order to help overcome this hindrance and estimate the contribution of aromatic side-chains to the UV region of the ECD spectra of proteins, in 2006 Kodama and co-workers [10] reproduced the ECD spectra of various side-chain conformations of model compounds. They used density functional theory (DFT) [11] and Ramachandran diagrams [12] to create the backbone dihedral angles ψ and φ of amino acid residues in proteins, optimizing the structures of model compounds in these biological molecules. Their calculations established that the ECD spectra of the aromatic residue models in α-helix, β-strand and random-coil configurations depend on the main-chain and the side-chain conformations [12]. Nevertheless, because some of the aromatic amino acids exhibit a complex ECD pattern in the NUV region, they concluded that it is very

challenging to distinguish different conformers in this region. With the purpose of surmounting this barrier, scientists have been attempting to access the FUV using Synchrotron radiation CD [9]. However, this method presents clear shortcomings in this spectral region because it is still based on the one-photon absorption (OPA) of optically active compounds. The OPA of amino acids typically occurs in the NUV (380–200 nm) to FUV (200–100 nm) region of the electromagnetic spectrum. As a result, the OPA of standard aqueous buffers in the same spectral region covers the small ECD signal and the scattering present at shorter wavelengths becomes an obstacle in heterogeneous samples. This has made further investigations of these essential biological systems more challenging.

To circumvent the characteristic limitations of ECD, Hernandez and co-workers recently proposed to utilize two-photon circular dichroism (TPCD)[13], which is a nonlinear spectroscopic technique analogous to ECD. First proposed theoretically [14-16] in the 1970s and invigorated in 2005 [17], TPCD facilitates the examination of chiral structures in the spectral region of the FUV that is inaccessible *via* ECD. TPCD is defined as the difference between the two-photon absorption (TPA) cross-sections for two photons of the same energy but opposite circular polarization (left and right)[14-16]. The measurement of this property was recently made possible with the establishment of the reliable double L-scan technique developed by Hernandez *et al* [18].

Since TPA processes are generated at longer wavelengths than those of OPA [19], the OPA in the TPA excitation region is usually insignificant and the scattering is greatly reduced. Furthermore, the existent quadratic dependence with the incident irradiance of TPA grants this excitation process more spatial resolution and penetration depth. This dependence also provides enhanced background discrimination and reduced photodamage to living specimens [20]. Vast technological applications in the fields of bioimaging [20, 21] and photodynamic therapy [22] have

been uncovered by these particular attributes. Moreover, TPCD is especially sensitive to small peptide structural distortions like variances in Ramachandran dihedral angles [12] and bond lengths of standard amino acids as demonstrated by Agren *et al.* [23], and even more recently by Hernandez and co-workers [24]. Thus, TPCD seems to be a harmonizing and favorable approach for the in-depth study and analysis of biomolecules *in vitro*.

As a continuation of our first work on TPCD of molecular structures simulating L-tryptophan residues in proteins, [24] our goal here is to determine the contribution of side- and main-chains conformation of aromatic amino acid residues to the TPCD spectra of proteins, in the NUV and FUV region. In order to accomplish this objective, we have embarked on the systematic and advanced conformational analysis of the theoretical TPA and TPCD spectra for a variety of side-chain conformations on model compounds in their secondary structure configuration, i.e. α-helix, β-strand and random coil.

In this article we report on the calculation and analysis of the theoretical TPA and TPCD of molecular structures imitating Phenylalanine (Phe), Tyrosine (Tyr), and Histidine (His) residues in proteins with secondary structure configurations. The comparative examination of the TPCD spectra of the various conformers in each configuration exposes unique fingerprints down to the FUV, a blind spectral region for ECD. Results exhibited in this article confirm the potential of FUV-TPCD to identify and study the structures of proteins in a region where intrinsic solvent absorption and sample scattering cover the ECD signal.

2. Theoretical and computational methods

The molecular structure of the L-stereoisomers of the main six Histidine residues (His11, His12, His21, His22, His31 and His32), the six primary Tyrosine residues (Tyr11, Tyr12, Tyr21, Tyr22, Tyr31 and Tyr32), and the three principal Phenylalanine residues (Phe1, Phe2 and Phe3), were optimized using Density Functional Theory (DFT) [11, 25], employing the Becke's three-parameter exchange, Lee, Yang and Parr correlation (B3LYP) hybrid functional [26-28] in combination with the 6-311G(d) basis set [29, 30], employing Gaussian 09 [31]. The establishment and optimization of all the His, Tyr and Phe conformers, in their corresponding α-helix, β-strand and random coil configuration, were achieved using initially the well-known Ramachandran dihedral angles for such structures [12]. These angles account for the relative configuration of the two groups about the C_α atom which are specified by ψ and ϕ (main-chain angles). The relative positioning of these two groups with respect to the C_α–C_β and C_β–C_γ bonds were estimated based on the side-chain angles χ_1 and χ_2 as shown in reference 24. Main- and side-chain angle values for all three amino acids were extracted from reference [10] (see Figure1 in reference 24).

TPA and TPCD for the lowest 80 electronic excited states of all optimized structures were calculated using Time-Dependent DFT (TD-DFT) [11, 25], using B3LYP and the Coulomb Attenuated Method variant of B3LYP (CAM-B3LYP) exchange correlation functionals (XCF), and the 6-311G(d) basis set [29, 30] using Dalton 2011[32]. All the calculations were completed in gas phase.

The TPA spectra were calculated using equation 1 [33-36].

$$\delta_{0f}^{TPA}(\omega) \approx 1.25273\times10^{-2}\times\omega^2\sum_f g\left(2\omega,\omega_{0f},\Gamma\right)\cdot\bar{\delta}_{0f}^{TPA}\left(\omega_{0f}\right). \quad (1)$$

Here $\overline{\delta}_{0f}^{TPA}\left(\omega_{0f}\right)$ is the orientational averaged two-photon probability for the degenerate case, ω is the excitation frequency and $g\left(2\omega,\omega_{0f},\Gamma\right)$ is a normalized Lorentzian lineshape function where Γ is the linewidth. The TPA spectra obtained from eq.(1) are in Göppert-Mayer units (GM), i.e. $10^{-50}cm^4.s.molec^{-1}.photon^{-1}$ when atomic units are used for the elements of eq. (1).

TPCD spectra were computed employing [14, 15, 36]:

$$\Delta\delta^{TPCD}\left(\omega\right)=\frac{4}{15}\frac{\left(2\pi\right)^3}{c_0^3\left(4\pi\varepsilon_0\right)^2}\times\omega^2\sum_f g\left(2\omega,\omega_{0f},\Gamma\right)\cdot R_{0f}^{TPCD}\left(\omega_{0f}\right), \quad (2)$$

$$\Delta\delta^{TPCD}\left(\omega\right)\approx 4.87555x10^{-5}\times\omega^2\sum_f g\left(2\omega,\omega_{0f},\Gamma\right)\cdot R_{0f}^{TPCD}\left(\omega_{0f}\right), \quad (3)$$

where $g\left(2\omega,\omega_{0f},\Gamma\right)$ is a normalized Lorentzian lineshape function (Γ) and $\Delta\delta^{TPCD}\left(\omega\right)$ from eq. (2) is obtained in GM units when the TPCD rotational strength $R_{0f}^{TPCD}\left(\omega_{0f}\right)$ and the circular frequency ω are in atomic units. $R_{0f}^{TPCD}\left(\omega_{0f}\right)$ is obtained, in atomic units, from:

$$R_{0f}^{TPCD}\left(\omega_{0f}\right)=-b_1B_1^{TI}\left(\omega_{0f}\right)-b_2B_2^{TI}\left(\omega_{0f}\right)-b_3B_3^{TI}\left(\omega_{0f}\right). \quad (4)$$

b_1, b_2 and b_3 are scalars defined by the experimental setup. In order to consider typical experimental conditions [13], we have used two left or two right circularly polarized photons propagating collinearly in the same direction, therefore b_1 = 6, b_2 = 2 and b_3 = -2 [14, 15, 36]. The molecular parameters B_1, B_2 and B_3 are acquired from the subsequent equations:

$$B_1^{TI}\left(\omega_{0f}\right)=\frac{1}{\omega^3}\sum_{\rho\sigma}M_{\rho\sigma}^{p,0f}\left(\omega_{0f}\right)P_{\rho\sigma}^{p*,0f}\left(\omega_{0f}\right), \quad (5)$$

$$B_2^{TI}\left(\omega_{0f}\right)=\frac{1}{2\omega^3}\sum_{\rho\sigma}T_{\rho\sigma}^{+,0f}\left(\omega_{0f}\right)P_{\rho\sigma}^{p*,0f}\left(\omega_{0f}\right), \quad (6)$$

$$B_3^{TI}\left(\omega_{0f}\right)=\frac{1}{\omega^3}\sum_{\rho\sigma}M_{\rho\sigma}^{p,0f}\left(\omega_{0f}\right)P_{\sigma\sigma}^{p*,0f}\left(\omega_{0f}\right). \quad (6)$$

Here, the generalized two-photon tensors $P_{\rho\sigma}^{p*,0f}\left(\omega_{0f}\right)$ and $T_{\rho\sigma}^{+,0f}\left(\omega_{0f}\right)$ depend on the electric transition dipole and quadrupole matrix elements in the velocity formulation, respectively, and $M_{\rho\sigma}^{p,0f}\left(\omega_{0f}\right)$ is controlled by the magnetic transition dipole matrix elements [14, 15, 36].

In order to achieve some initial validation of our theoretical approximation we performed linear calculations of the ECD spectra of three of the few amino acids that have been measured down to 150 nm, [37] i.e. L-Val, L-Leu and L-Ala (see SI). Even though we are mindful that this is not totally suitable to make a definite choice of a preferred XCF, in some cases it can be helpful [38].

Briefly, OPA and ECD for the lowest 80 electronic excited states were calculated on all optimized structures at the B3LYP/6-311G(d) and CAM-B3LYP/6-311G(d) levels of theory using Gaussian 09 [31] in gas phase (more details are given in SI). All convoluted spectra were calculated using Γ = 0.124 eV (FWHM) in agreement with reference 10. As determined in previous works [38], CAM-B3LYP predicts high energy, long range and charge transfer diffuse states better than B3LYP because the HOMO energy, determined with this XCF, is normally lower than for B3LYP [38]. Therefore, based on previous studies [24] we are convinced that CAM-B3LYP should predict more precisely the TPCD spectra of His, Tyr, and Phe residues in the FUV. Nevertheless, we are aware that using this specific XCF is not enough to perform the most

accurate calculation because the theoretical results are impacted by the strong state diffusion within the short wavelength region of the spectra, the high density of states and the prospect of intermediate state resonances. For that reason, one should consider complex damped response theory when working with molecules such as His, Tyr, and Phe [39, 40], which gives the same results as standard response theory if the response approach is able to converge all requested states but the damped response theory approach may prove to be a more computationally convenient way to obtain high-energy states. Unfortunately, this approach is available for TPA [41] but not for TPCD yet. On the other hand, because one cannot completely ignore the theoretical results achieved with B3LYP until the experimental TPCD spectra of these amino acids are measured in the FUV, in the supporting information (SI) we present all the theoretical TPA and TPCD spectra of His, Tyr, and Phe residues obtained with this XCF.

3. Results and discussion

In order to show the observed differences between configurations, the optimized stereochemical structures of His, Tyr, and Phe residues in random coil are presented in Figure 1. Similar types of structures for the α-helix and β-strand are shown in SI. All the optimized angles for each conformer are presented in Table I, II, and III. In these three tables, one can observe rather important differences in the main-chain angles for different configurations of His, Tyr, and Phe residues within each conformation. This indicates that, as it was concluded for Trp [24], the main-chain angles are unquestionably modified by the side-chain configuration and *vice versa*. It is also noticeable that the differences in ψ and ϕ are more significant in the β-strand conformation. All these distortions can be explained in terms of the allowable van der Waals distances which demarcate the areas of permissible ψ and ϕ [12]. In di-, tri-, and polypeptide chains, ψ and ϕ configurations always take place inside the outer boundaries of the minimum contact distance assumed, and a clustering configuration can emerge. Every time the C_α atom under analysis belongs to a glycyl residue the configuration is close to 180^0 and 0, and whenever there is a C_β atom the clustering is nearly 150^0 and 120^0 [12].

After having optimized the structures of all residues of the left-handed isomers of His, Tyr, and Phe, in each backbone configuration (α-helix, β-strand and random coil), we computed their TPA and TPCD spectra within the two-photon excitation wavelength range comprising 220 nm to 500 nm (in the OPA scale this range corresponds to a wavelength range from 110 nm to 250 nm). The convoluted spectra were calculated using CAM-B3LYP/6-311G(d) in Dalton 2011 and with $\Gamma = 0.124$ eV (FWHM) (theoretical spectra calculated with B3LYP/6-311G(d) can be found in SI).

First, we examine the effect of the side-chains onto the TPA and TPCD spectra of His, Phe, and Tyr residues, in their corresponding β-strand, α-helix, and random coil configurations (see Figures 2–8).

In the case of His (Figure 2-4), the straight comparison between the corresponding spectra of a pair of conformers with equal χ_1 uncovers moderately mild differences for TPA, in contrast to considerably big variations in TPCD. The largest difference in TPA is noticed between His11 and His12 in β-strand, which exhibits a noteworthy variation in the peak located at 252 nm. The fact that TPCD is a technique recognized for its sensitivity to small peptide structural distortions such as side-chain conformational angles and bond lengths of residues [23], and Ramachandran dihedral angles [12], can be certified one more time by the occurrence of bands with opposite signs (spectral signatures) among residues of His. For instance, between His11 and His12 at 328 nm and 380 nm, and between His31 and His32 within the spectral region 280-350 nm, there are clear variations in β-strand. In α-helix and random coil, the residues of His look very similar. This makes it challenging to distinguish between a pair of conformers with same χ_1. These TPCD spectra also present some exceptions, specifically in His31 and His32 in α-helix within a wavelength range of 300-340 nm as well as in His11 and His12 in the region between 240 and 260 nm in random coil, where the TPCD signal acquires an opposite sign.

With regard to Phe, in Figure 5 we can see differences in the TPA and TPCD spectra of the three conformers in each configuration. First, in the TPA spectra of all the conformers in β-strand, α-helix, and Random coil one can distinguish the following different spectral features between 280 and 340 nm: i) β-strand: a strong band centered at 330 nm for Phe1 and two strong bands at approximately 300 nm and 340 nm for Phe2, ii) α-helix: a strong band centered at ~ 300 nm for Phe2, and iii) Random coil: a double band with peaks at ca. 300 nm and 315 nm for Phe1.

Second, the TPCD spectra of this amino acid in β-strand reveals evident differences in spectral signatures between the three residues, i.e. alternating sign signals for each conformer within the wavelength range between 300 nm to 380 nm. In α-helix, though, the situation is slightly different, while Phe1 is distinguishable from the rest of the residues through specific negative sign features observed at 250 nm and 280 nm, in Phe2 the differentiation is more obvious at ca. 300 nm where the sign of the characteristic TPCD spectra is opposite for analogous conformers. The same trend can be notorious in random coil where the identification of the Phe1 residue is possible through its opposite sign TPCD signals between 260 nm and 290 nm.

Concerning Tyr (Figures 6-8), it is interesting to highlight the fact that the TPA spectra of all its residues are virtually identical. Likewise, the TPCD spectra of Tyr residues in β-strand and Random coil are very similar between pairs of conformers. However, the TPCD spectra of equivalent conformers in α-helix reveals small but noticeable differences between Tyr11 and Tyr12, and Tyr31 and Tyr32 - distinct peaks between 300 nm and 380 nm.

In summary, one can close this part of the analysis by recognizing that TPCD, contrary to standard ECD [10], allows us to distinguish between pairs of conformers of amino acid residues in different secondary protein configurations, mainly in His. Through the analysis of the OPA and ECD spectra of residues of similar amino acids performed by Kodama and co-authors [10], it has been determined that the aromatic ring present in Phe and Tyr is not anisotropically polarizable as it is in His. Therefore, distinguishing between residues of Phe and Tyr using linear spectroscopy is extremely challenging.

Afterward, we analyzed the effect of the main-chains onto the TPA and TPCD spectra of His, Phe, and Tyr residues in α-helix, β-strand, and random coil conformations. In Figures 9-11, we present, in a comparative fashion, the corresponding linear and nonlinear absorption and CD

spectra of His11, Phe1, and Tyr11 residues calculated with CAM-B3LYP (the comparative plots of all the remaining His$\chi_1\chi_2$, Pheχ_1, and Tyr$\chi_1\chi_2$ residues, calculated with CAM-B3LYP and B3LYP can be found in SI). In order to facilitate the comparison between the linear and nonlinear spectra, OPA wavelength is used for both cases throughout the discussion of Figures 9-11.

Before beginning the discussion in this section, it is worth noting that all three amino acids present more distinguishable and measurable variations in the nonlinear absorption spectra, TPA and TPCD, of the different conformations than in their corresponding linear counterpart, OPA and ECD.

First, in the TPA spectra of His11 (Figure 9) one can precisely identify its presence in all three different conformations through: i) the longer wavelength band at ~ 200 nm in the α-helix, ii) a double peak broad band between 140 nm and 160 nm in β-strand, and iii) the very specific band in the blue side of the spectrum at ~ 130 nm in random coil. On the other hand, in the TPCD spectra of the same amino acid residue, in all three different configurations, one can observe the following spectral fingerprints: i) a strong negative band in the red side of the spectrum at ~ 200 nm in α-helix, ii) a specific negative double band between 160 nm and 180 nm in β-strand, and iii) a clearly strong double peak band between 150 nm and 170 nm in random coil. The presence of these specific bands, that allows the recognition of the different configurations using His residues, reveals the applicability of TPCD as a complementary technique to ECD for the study of complex protein structures. However, the most important point to be highlighted on this part of the analysis is the number of unique and identifiable spectral signatures of His11 in all three configurations down to 110 nm. These features provide additional conformational information to the limited traditional ECD technique. In fact, inspecting the OPA

and ECD spectra of His11 in α-helix, β-strand, and random coil, in the short wavelengths range (< 180 nm) one can see a very large number of convoluted bands overlapping with each other through almost the whole spectrum. Only in the ECD plot, and within the typical spectral range employed for the study of proteins using linear spectroscopy, i.e. above 180 nm, one can observe few spectral signatures for β-strand and random coil useful for the identification of this amino acid in protein secondary structures (shaded area).

Second, in Phe1 (Figure 10) one can distinguish similar characteristics to those observed for His11, i.e. less intricate nonlinear spectra down to 110 nm and the presence of specific bands for each conformation, mostly in the TPCD plot, that allows identification of one from another. Although, in TPA one can only distinguish the β-strand from the other two configurations through a strong peak at ~ 165 nm, in TPCD characteristic signals with opposite sign allow identification of all three conformations independently. For instance, while β-strand presents a unique negative broad TPCD band between 150 nm and 180 nm, the other two configurations present a positive one. To differentiate between α-helix and random coil one should look at longer wavelengths (small positive band at ~ 175 nm) for the former and at very short wavelengths (sharp negative peak at ~ 125 nm) for the latter. Again, the most remarkable aspects of this analysis are, i) the unique and identifiable spectral signatures in all three configurations down to 110 nm, and ii) the accessibility to additional conformational information in the FUV. Even in the event were ECD becomes applicable for the conformational study of proteins in the FUV, the number of convoluted bands, overlapping with each other in that region, would limit its use in that spectral region - ECD is truly functional above ~ 180 nm.

Third, besides the already discussed characteristic nonlinear absorption fingerprints observed in His and Phe, we found that TPA in Tyr (Figure 11) is not reliable for the

differentiation between main-chain "back-bone" conformations. However, TPCD shows obvious differences in certain spectral regions to specifically identify all three conformations: i) α-helix: negative broad band in the red side of the spectrum above 170 nm, ii) β-strand: positive sharp peak at ~ 125 nm, and iii) random coil: positive broad band between 140 nm and 170 nm. One more time, it has been proven that TPCD has all the potential as a tool for the identification between α-helix, β-strand, and random coil utilizing aromatic amino acid residues.

Finally, we present the direct comparison of the TPA and TPCD spectra on a set of two amino acid residues of His, Phe and Tyr (His11-Phe1-Tyr11 and His21-Phe2-Tyr21), in their corresponding β-strand, α-helix, and random coil configurations. As it can be observed in Figures 12-14, they all exhibit clear spectral differences in TPA and TPCD, in each configuration. These signatures could be used, synergistically, for the conformational analysis of peptides and proteins in the FUV.

First, we assess the potential of TPA spectroscopy for the conformational analysis of peptides and proteins examining the specific spectral signatures of aromatic amino acids in different configurations. In β-strand, one can observe, simultaneously, the following complementary spectral fingerprints: a strong peak at ~ 270 nm in Tyr11, a small band at ~ 380 nm in His11, a sharp peak at ~ 280 nm in Tyr21, and a relatively intense band at ~ 290 in Phe2 nm. In α-helix, one can notice the following complementary spectral features: a strong peak at ~ 270 nm in Tyr11, a small band at ~ 390 nm in His11, and a broad band with two strong peaks at ~ 260 nm and 280 nm in Tyr21. In random coil one can distinguish the following spectral signatures: two strong peaks centered at ~ 280 nm and 330 nm in Tyr11, a very broad band with defined spectral structure between 240 nm and 300 nm in Tyr21, and a weak double band above ~ 360 nm in His21. Although, some peaks/bands are similar for different amino acid residues in

different configurations, several others are very specific for each structural arrangement. This outcome shows the limited but still useful potential of TPA for the targeted analysis.

Next, we present the examination of the TPCD spectra of the same amino acid residues, in the same protein secondary structures. As shown subsequently, our results reveal the exceptional spectroscopic ability of TPCD, compared to the traditional ECD, for the analysis of protein structures in a region where the overcrowded ECD spectra make the identification of complex structures using linear spectroscopy very difficult [24].

Using TPA one can attempt to differentiate between protein secondary structures by looking for the following specific fingerprints in Tyr, i.e. a) β-strand: strong band centered at ~ 280 nm with two shoulders on the blue side of the spectrum in Tyr11; b) α-helix: broad band with two sharp peaks at ~ 250 nm and 280 nm in Tyr21; c) Random coil: strong broad band with structure between ~ 240 nm and 300 nm in Tyr21. Other amino acid residues do not present such clear and differentiable signatures in TPA.

Using TPCD one can clearly gather conformational information from all three amino acids. For instance, β-strand can selectively be identified through specific positive or negative signatures at ~ 260 nm in Tyr11, 270 nm in Tyr21, 300 nm in His11 and 290 nm in Phe1; b) α-helix through spectral bands at ~ 360 nm in Tyr11, 240 nm in His21 and in Phe2; c) Random coil by measuring TPCD at ~ 260 nm and 300 nm in Tyr11, 280 nm and 330 nm in His11. TPCD exposes its remarkable capability for the identification of molecules within the same protein structure using specific signatures of different aromatic amino acids, simultaneously.

Conclusions

The potential of TPCD for the study and analysis of complex chemical structures such as peptides and proteins in a region that is impenetrable using linear absorption processes was demonstrated. The specific signatures found in L-histidine, L-phenylalanine, and L-tyrosine residues, in β-strand, α–helix, and random coil configurations, validated the distinguishing ability of TPCD to differentiate between: i) very similar species with equal χ_1, and ii) between equal or dissimilar species in different conformational configuration. The evaluation of the TPCD spectra of the re-optimized structures demonstrated that the main-chain angles are undeniably influenced by the side-chain configuration and *vice versa.* Through the determination of the TPCD signal of multiple amino acids simultaneously one can gather multi-parametric data to certify the existence of a specific conformation and assess a small distortion of the same in the FUV. Finally, we should mention that in some cases the TPA spectra provide additional conformational information that can be used for the study of amino acid residues.

Acknowledgments

This work was partially supported by the National Science Foundation through Grant Number CHE-0840431. The computing time provided by STOKES ARCC is gratefully acknowledged. The authors would like to thank *N.I.K.O.* for his assistance for the completion of this work on time.

References

1. R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2004, 108, 16965-16975.
2. S. M. Kelly, T. J. Jess and N. C. Price, *Biochim. Biophys. Acta: Proteins Proteomics*, 2005, 1751, 119-139.
3. N. J. Greenfield, *TrAC, Trends Anal. Chem.*, 1999, 18, 236-244.
4. S. Y. Venyaminov and J. T. Yang, in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed. G. D. Fasman, Plenum Press, LLC, New York, 1996, ch. 3, pp. 69-108.
5. N. Berova, P. L. Polavarapu, K. Nakanishi and R. W. Woody, *Comprehensive Chiroptical Spectroscopy*, John Wiley & Sons, Hoboken, NJ, 2012.
6. R. W. Woody and K. A. Dunker, in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed. G. D. Fasman, Plenum Press, New York, 1996, pp. 109-158.
7. E. H. Strickland, *CRC Crit. Rev. Biochem.*, 1974, 2, 113-175.
8. P. C. Kahn, *Methods Enzymol.*, 1979, 61, 339-378.
9. A. J. Miles and B. A. Wallace, *Chem. Soc. Rev.*, 2006, 35, 39-51.
10. T. Tanaka, T. S. Kodama, H. E. Morita and T. Ohno, *Chirality*, 2006, 18, 652-661.
11. E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, 1984, 52, 997-1000.
12. G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, *J. Mol. Biol.*, 1963, 7, 95-99.
13. C. Toro, L. De Boni, N. Lin, F. Santoro, A. Rizzo and F. E. Hernandez, *Chem. Eur. J.*, 2010, 16, 3504-3509.
14. I. Tinoco, *J. Chem. Phys.*, 1975, 62, 1006-1009.
15. E. A. Power, *J. Chem. Phys.*, 1975, 63, 1348-1350.
16. D. L. Andrews, *Chem. Phys.*, 1976, 16, 419-424.
17. B. Jansik, A. Rizzo and H. Agren, *Chem. Phys. Lett.*, 2005, 414, 461-467.
18. L. De Boni, C. Toro and F. E. Hernández, *Opt. Lett.*, 2008, 33, 2958-2960.
19. R. W. Boyd, *Nonlinear Optics*, Academic Press, San Diego, CA, 1992.
20. W. Denk, J. Strickler and W. Webb, *Science*, 1990, 248, 73-76.
21. J. M. Squirrell, D. L. Wokosin, J. G. White and B. D. Bavister, *Nat. Biotechnol.*, 1999, 17, 763-767.
22. J. R. Starkey, A. K. Rebane, M. A. Drobizhev, F. Meng, A. Gong, A. Elliott, K. McInnerney and C. W. Spangler, *Clin. Cancer. Res.*, 2008, 14, 6564-6573.
23. B. Jansik, A. Rizzo and H. Agren, *J. Phys. Chem. B*, 2007, 111, 446-460.
24. Y. Vesga, C. Diaz, M. Higgs and F. E. Hernandez, *Chem. Phys. Lett.*, 2014, 601, 6-12.
25. M. Dierksen and S. Grimme, *J. Chem. Phys.*, 2006, 124, 174301.
26. A. D. Becke, *Phys. Rev. A*, 1988, 38, 3098-3100.
27. A. D. Becke, *J. Chem. Phys.*, 1993, 98, 5648-5652.
28. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, 37, 785-789.
29. R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1980, 72, 650-654.
30. A. D. McLean and G. S. Chandler, *J. Chem. Phys.*, 1980, 72, 5639-5648.
31. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, Gaussian, Inc., Wallingford CT, 2009.

32. K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013, DOI: 10.1002/wcms.1172.
33. W. M. McClain, *J. Chem. Phys.*, 1971, 55, 2789-2796.
34. M. G. Vivas, C. Diaz, L. Echevarria, C. R. Mendonca, F. E. Hernández and L. De Boni, *J. Phys. Chem. B*, 2013, 117, 2742-2747.
35. D. L. Silva, N. A. Murugan, J. Kongsted, Z. Rinkevicius, S. Canuto and H. Agren, *J. Phys. Chem. B*, 2012, 116, 8169-8181.
36. M. Guillaume, K. Ruud, A. Rizzo, S. Monti, Z. Lin and X. Xu, *J. Phys. Chem. B*, 2010, 114, 6500-6512.
37. U. J. Meierhenrich, J.-J. Filippi, C. Meinert, J. H. Bredehöft, J.-i. Takahashi, L. Nahon, N. C. Jones and S. V. Hoffmann, *Angew. Chem. Int. Ed.*, 2010, 49, 7799-7802.
38. N. Lin, F. Santoro, X. Zhao, C. Toro, L. De Boni, F. E. Hernández and A. Rizzo, *J. Phys. Chem. B*, 2011, 115, 811-824.
39. K. Kristensen, J. Kauczor, T. Kjærgaard and P. Jørgensen, *J. Chem. Phys.*, 2009, 131, 044112.
40. V. Barone, *Computational Strategies for Spectroscopy: from Small Molecules to Nano Systems*, John Wiley & Sons, Hoboken, NJ, 2012.
41. K. Kristensen, J. Kauczor, A. J. Thorvaldsen, P. Jørgensen, T. Kjærgaard and A. Rizzo, *J. Chem. Phys.*, 2011, 134, 214104.

Figure Caption

Figure 1. Stereochemical structures of L-histidine (left), L-tyrosine (right), and L-phenylalanine (down) models in a random coil configuration. Optimizations were performed with DFT/B3LYP/6-311G(d) in gas phase using Gaussian 09[31].

Figure2. Comparative plots of TPA (left) and TPCD (right) spectra of L-histidine models in β-strand configuration. His11 and His12 (top), His21 and His22 (middle), His31 and His32 (bottom). TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 3. Comparative plots of TPA (left) and TPCD (right) spectra of L-histidine models in α-helix configuration. His11 and His12 (top), His21 and His22 (middle), His31 and His32 (bottom). TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 4. Comparative plots of TPA (left) and TPCD (right) spectra of L-histidine models in random coil configuration. His11 and His12 (top), His21 and His22 (middle), His31 and His32 (bottom). TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 5. Comparative plots of TPA (left) and TPCD (right) spectra of L-phenylalanine models in β-strand (top), α-helix (middle), and random coil (bottom) configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 6. Comparative plots of TPA (left) and TPCD (right) spectra of L-tyrosine models in β-strand configuration. Tyr11 and Tyr12 (top), Tyr21 and Tyr22 (middle), Tyr31 and Tyr32 (bottom). TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 7. Comparative plots of TPA (left) and TPCD (right) spectra of L-tyrosine models in α-helix configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 8. Comparative plots of TPA (left) and TPCD (right) spectra of L-tyrosine models in random coil configuration. Tyr11 and Tyr12 (top), Tyr21 and Tyr22 (middle), Tyr31 and Tyr32 (bottom). TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 9. Comparative plots of TPA (top left), TPCD (top right), OPA (bottom left), and ECD (bottom right) spectra of His11 in random coil (red dotted line), α-helix (black solid line) and β-strand (blue dashed line) configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32]. OPA and ECD for the lowest 80 electronic excited states were computed of all optimized structures at the CAM-B3LYP/6-311G(d) level of theory using Gaussian 09 [31] in gas phase. Shaded area indicates where ECD is truly functional.

Figure10. Comparative plots of TPA (top left), TPCD (top right), OPA (bottom left), and ECD (bottom right) spectra of Phe1 in random coil (red dotted line), α-helix (black solid line) and β-strand (blue dashed line) configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011[32]. OPA and ECD for the lowest 80 electronic excited states were computed of all optimized structures at the CAM-B3LYP/6-311G(d) level of theory using Gaussian 09 [31] in gas phase. Shaded area indicates where ECD is truly functional.

Figure 11. Comparative plots of TPA (top left), TPCD (top right), OPA (bottom left), and ECD (bottom right) spectra of Tyr11 in random coil (red dotted line), α-helix (black solid line) and β-strand (blue dashed line) configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32]. OPA and ECD for the lowest 80 electronic excited states were computed of all optimized structures at the CAM-B3LYP/6-311G(d) level of theory using Gaussian 09 [31] in gas phase. Shaded area indicates where ECD is truly functional.

Figure 12. Comparative plots of TPA (left) and TPCD (right) spectra of Tyr11 (red dotted line), His11 (black solid line) and Phe1 (blue dashed line) (Top), and Tyr21 (red dotted line), His21 (black solid line) and Phe2 (blue dashed line) (Bottom), in β-strand configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 13. Comparative plots of TPA (left) and TPCD (right) spectra of Tyr11 (red dotted line), His11 (black solid line) and Phe1 (blue dashed line) (Top), and Tyr21 (red dotted line), His21 (black solid line) and Phe2 (blue dashed line) (Bottom), in α-helix configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

Figure 14. Comparative plots of TPA (left) and TPCD (right) spectra of Tyr11 (red dotted line), His11 (black solid line) and Phe1 (blue dashed line) (Top), and Tyr21 (red dotted line), His21 (black solid line) and Phe2 (blue dashed line) (Bottom), in random coil configuration. TPA and TPCD response for the lowest 80 electronic excited states of all optimized structures were computed with TD-DFT/ CAM-B3LYP/6-311G(d) in gas phase using Dalton 2011 [32].

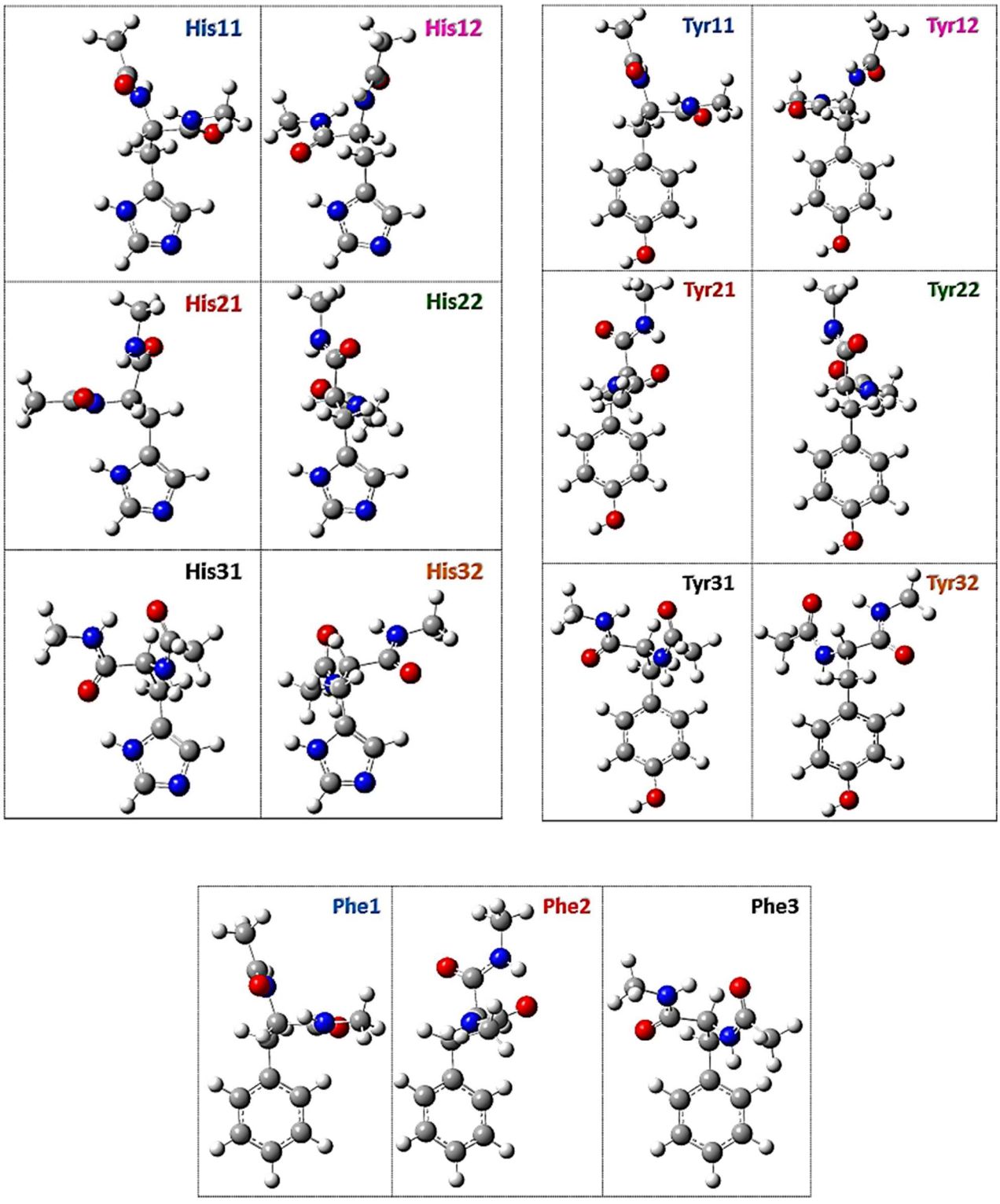

Figure 1.

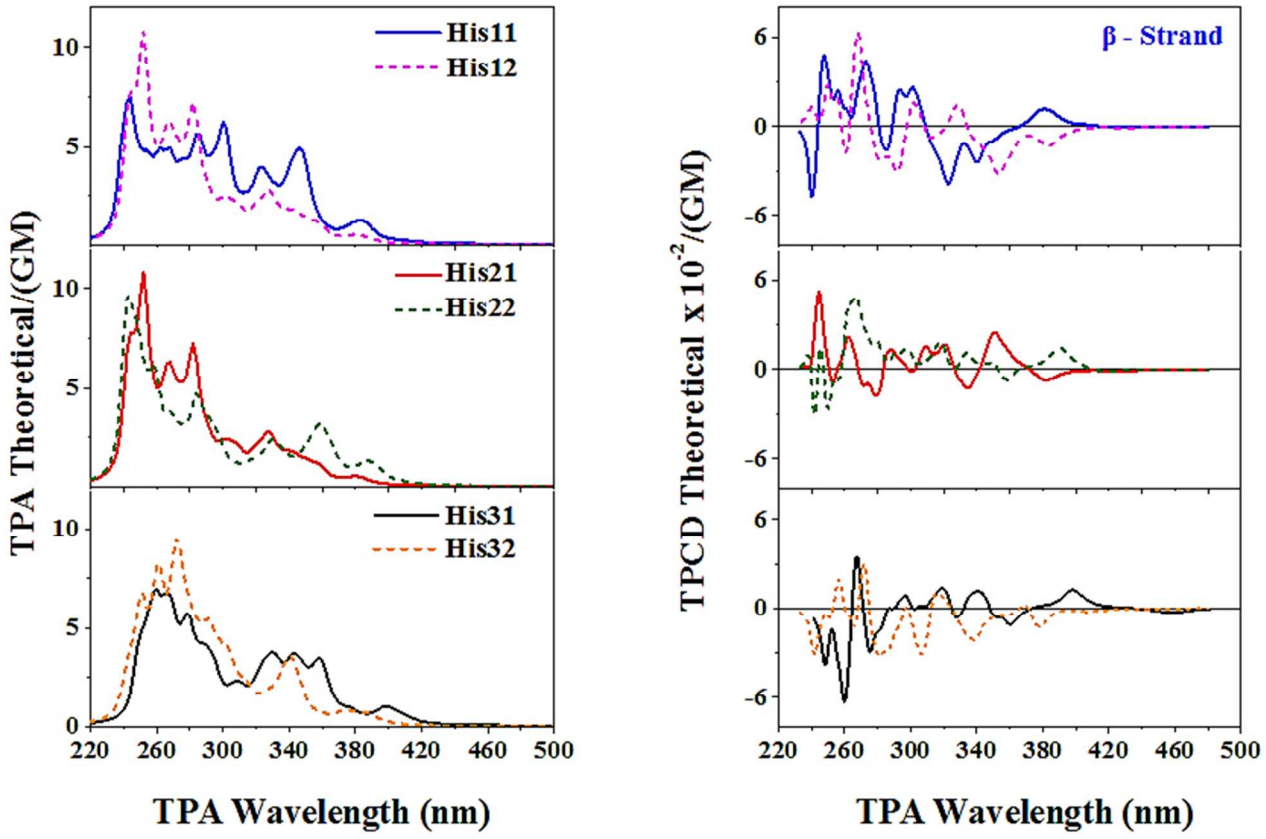

Figure 2.

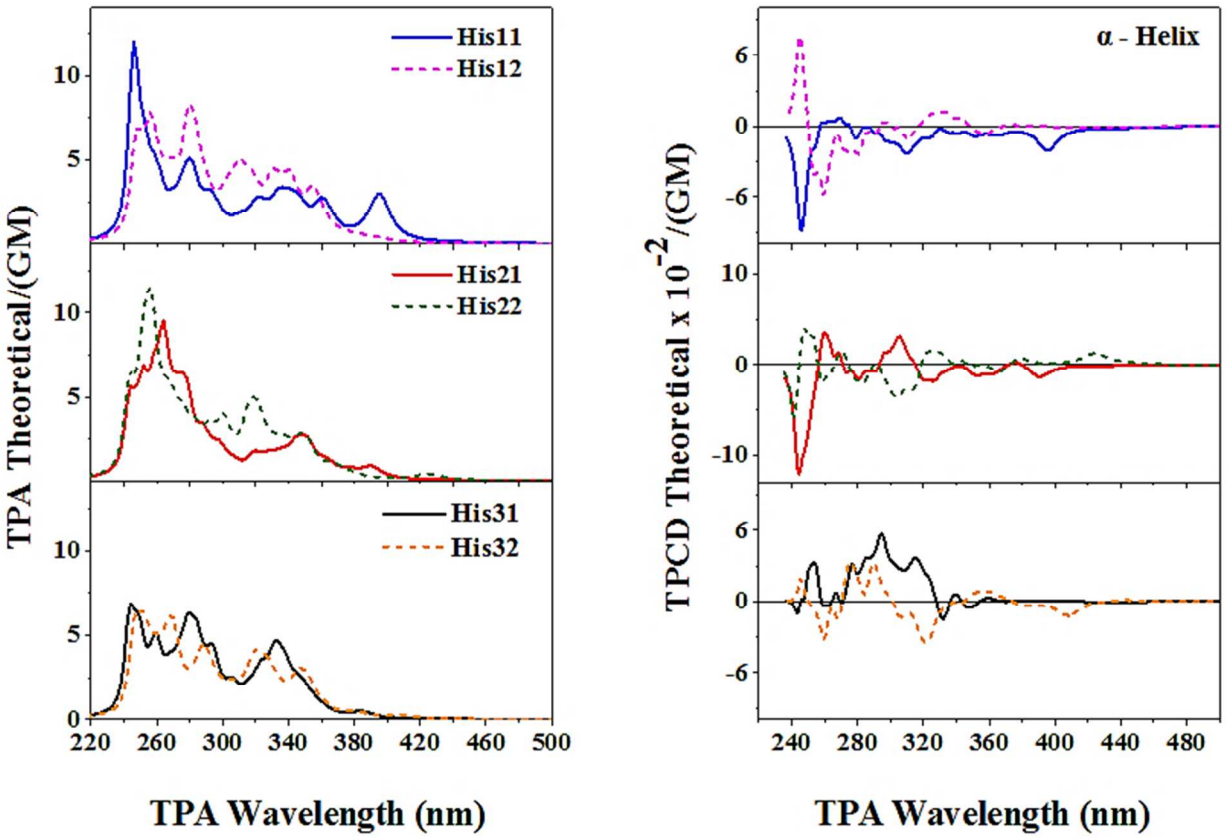

Figure 3.

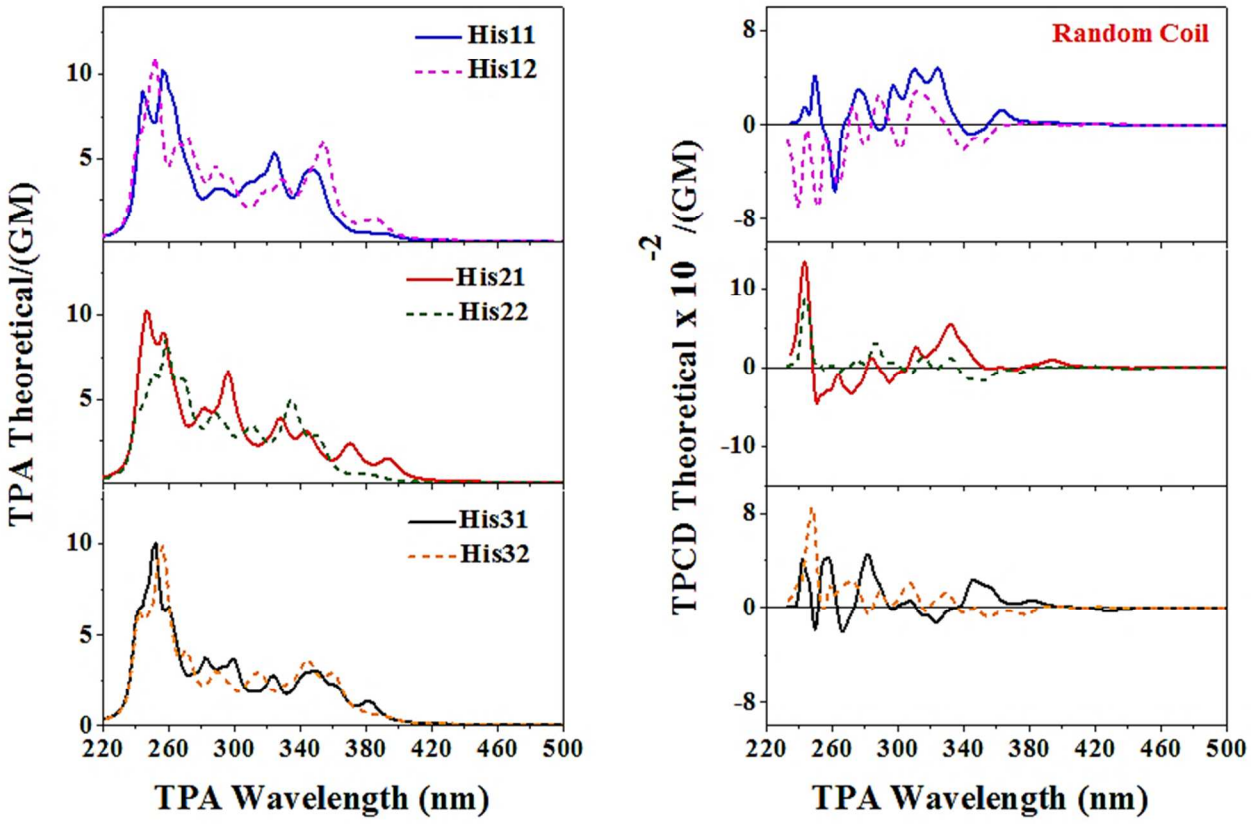

Figure 4.

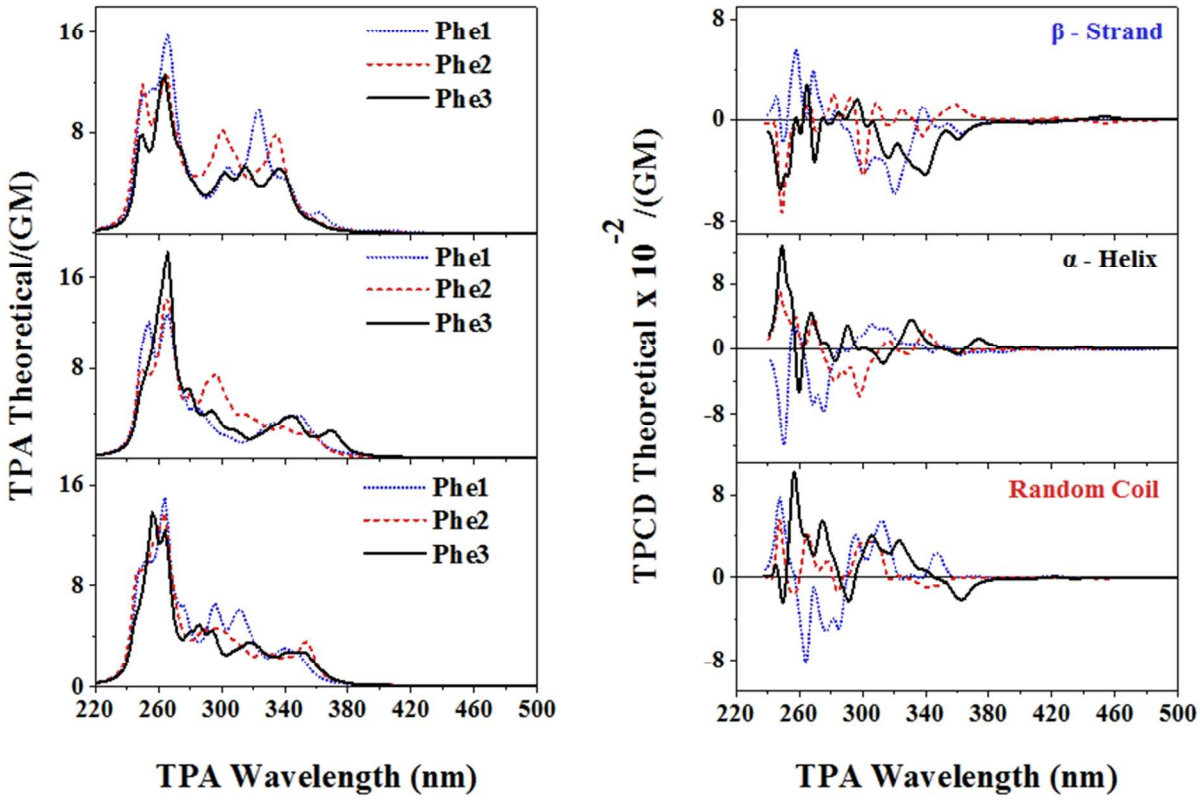

Figure 5.

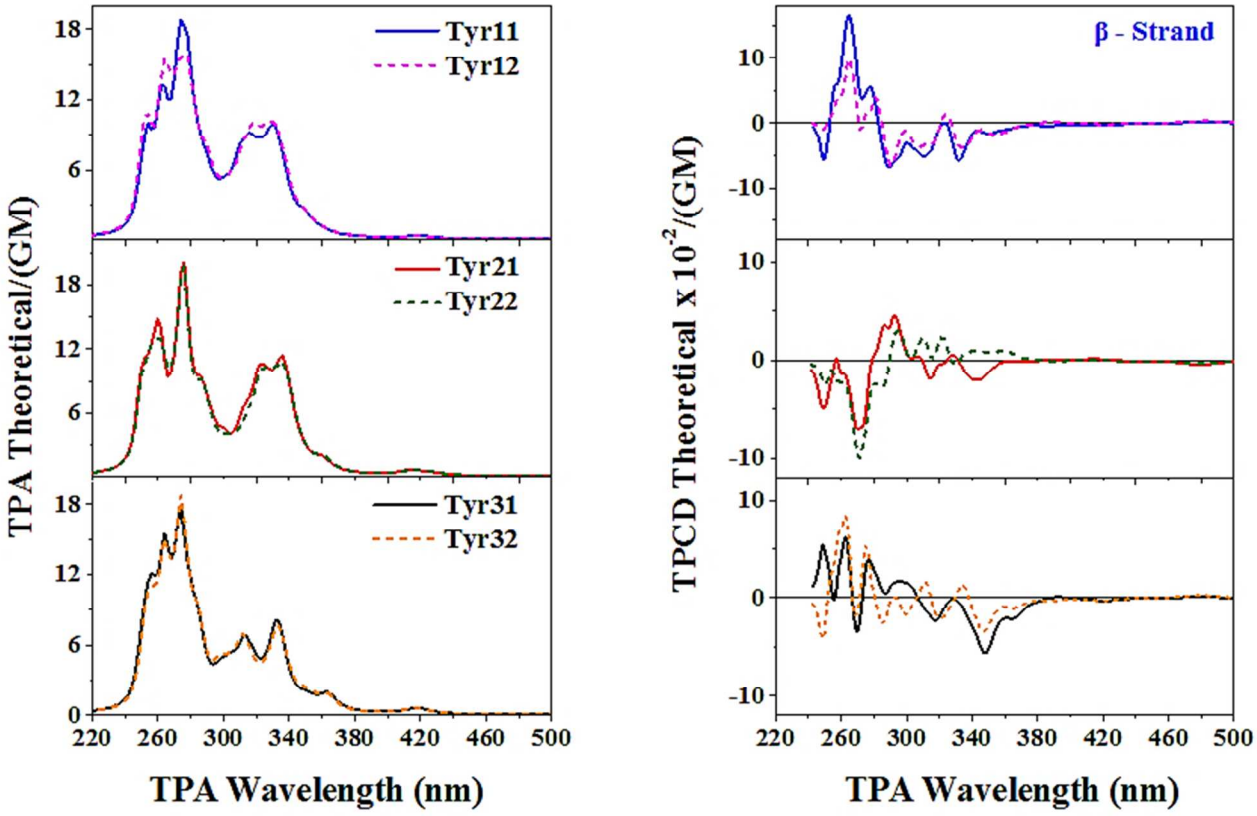

Figure 6.

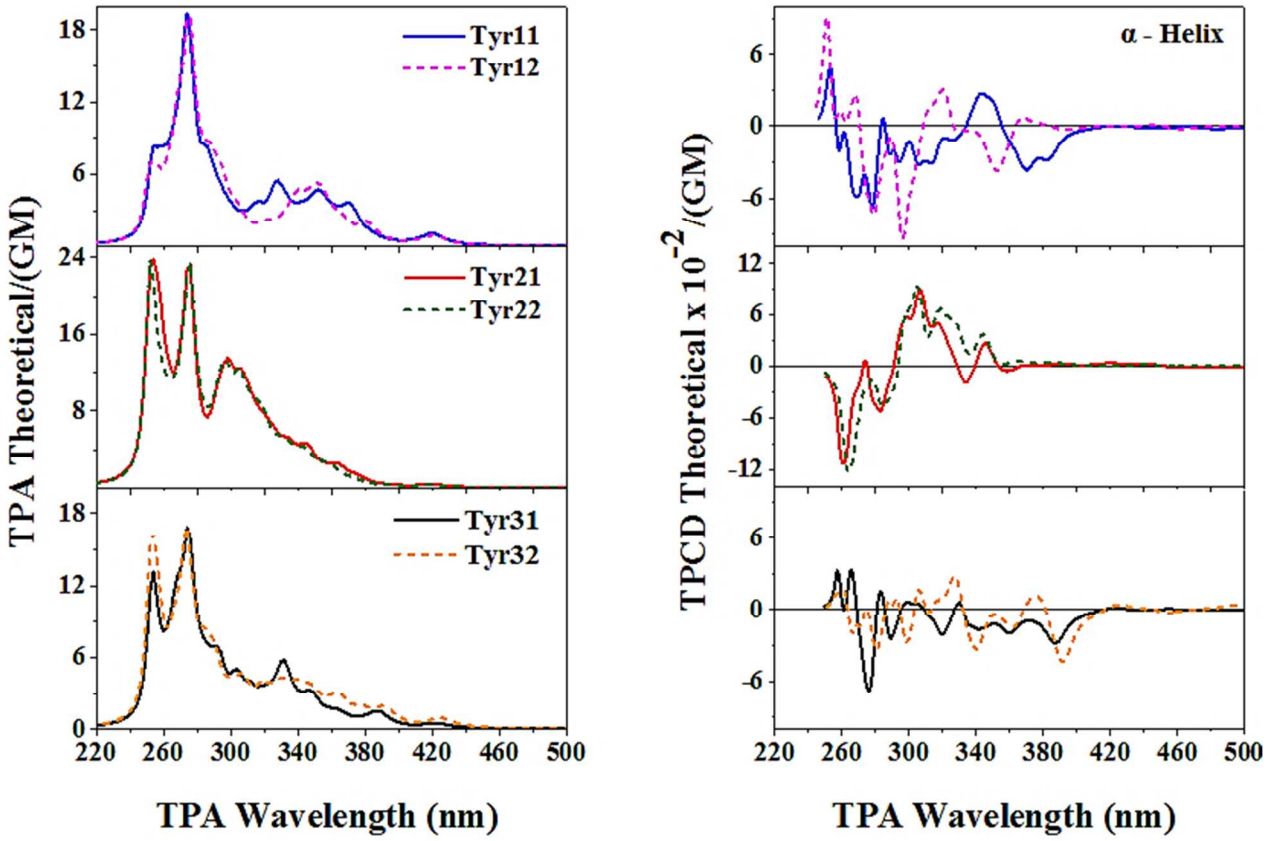

Figure 7.

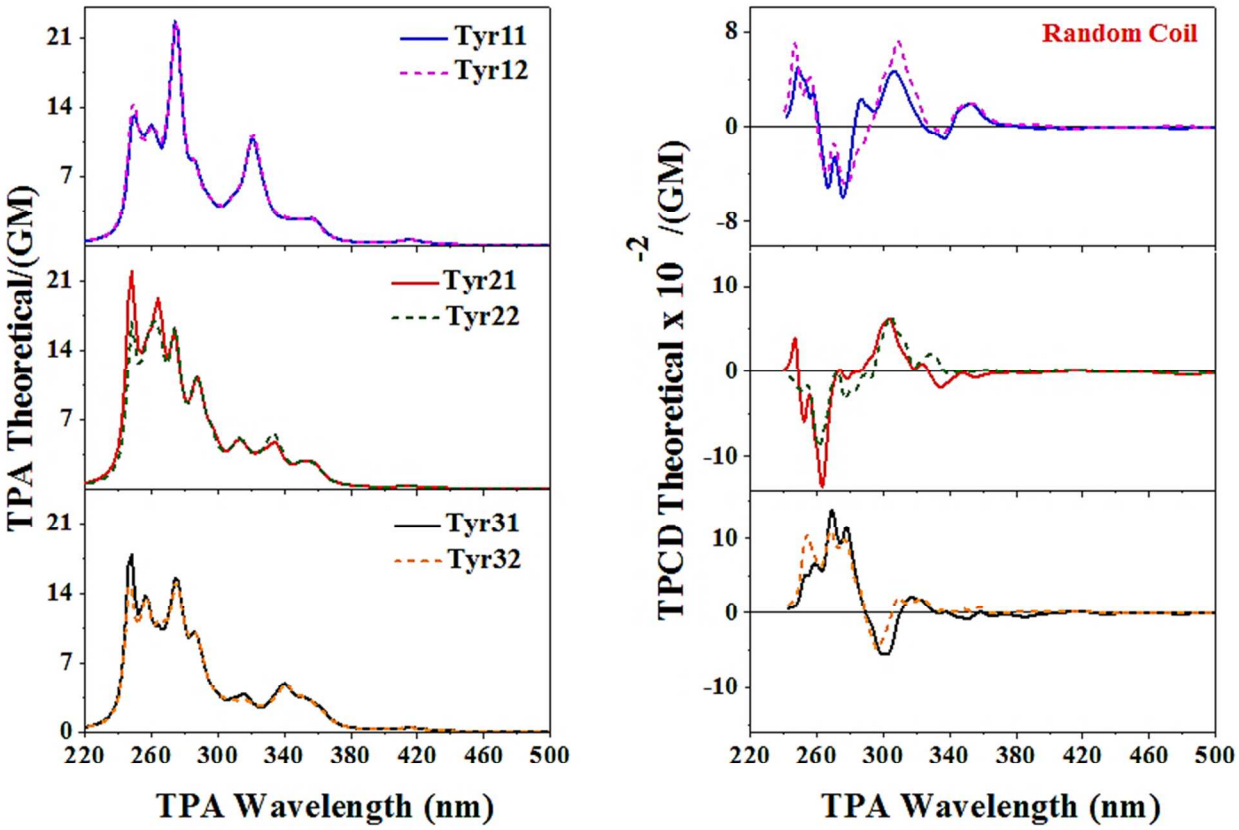

Figure 8.

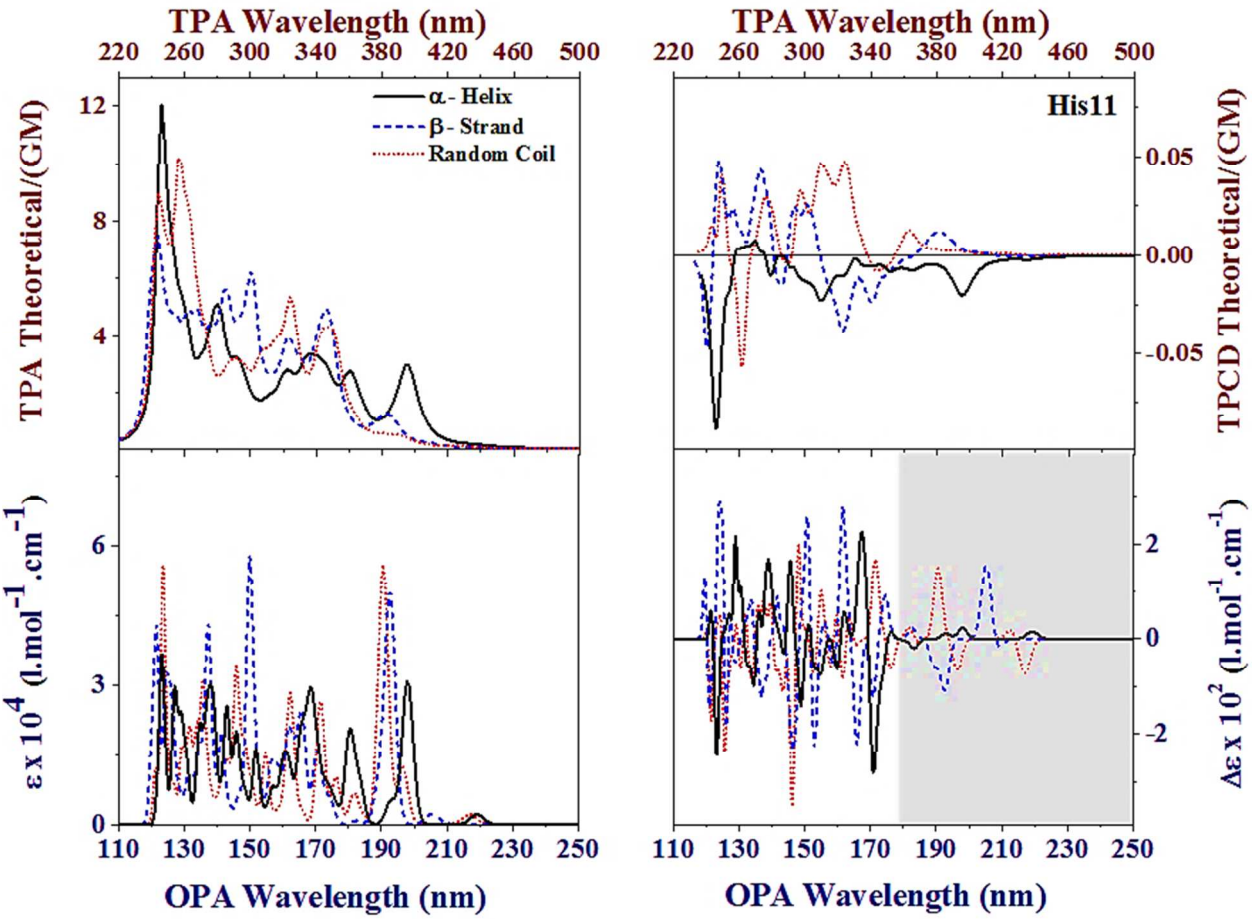

Figure 9.

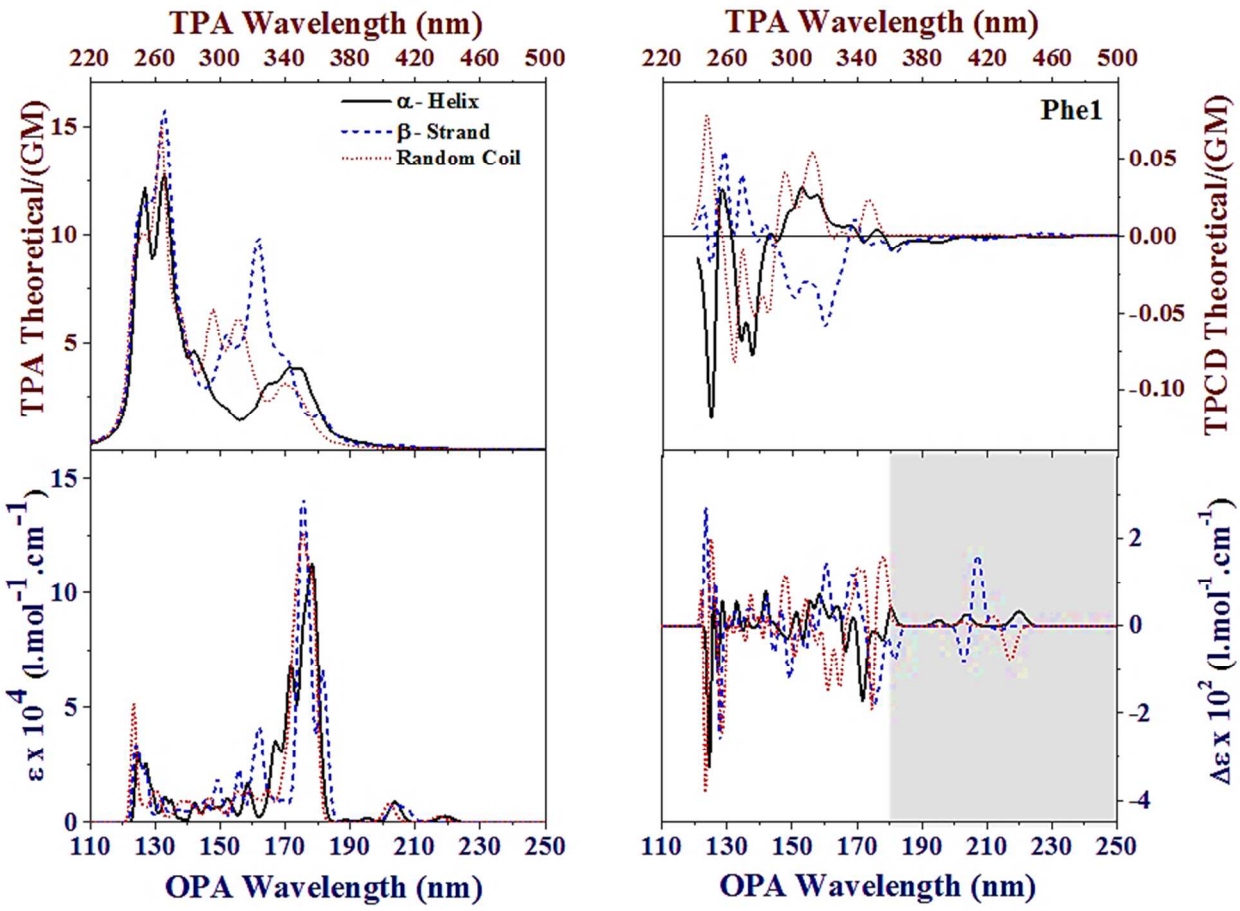

Figure 10.

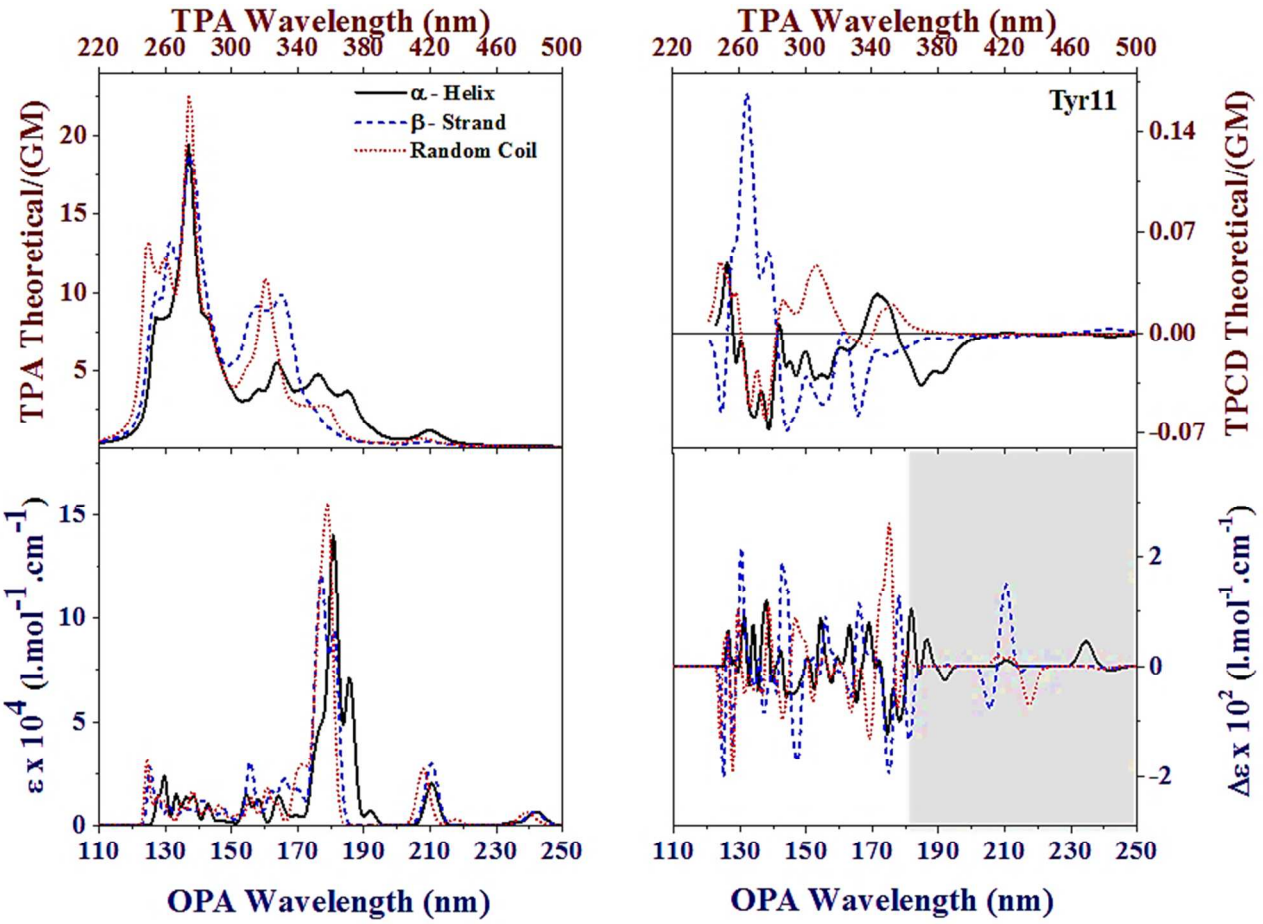

Figure 11.

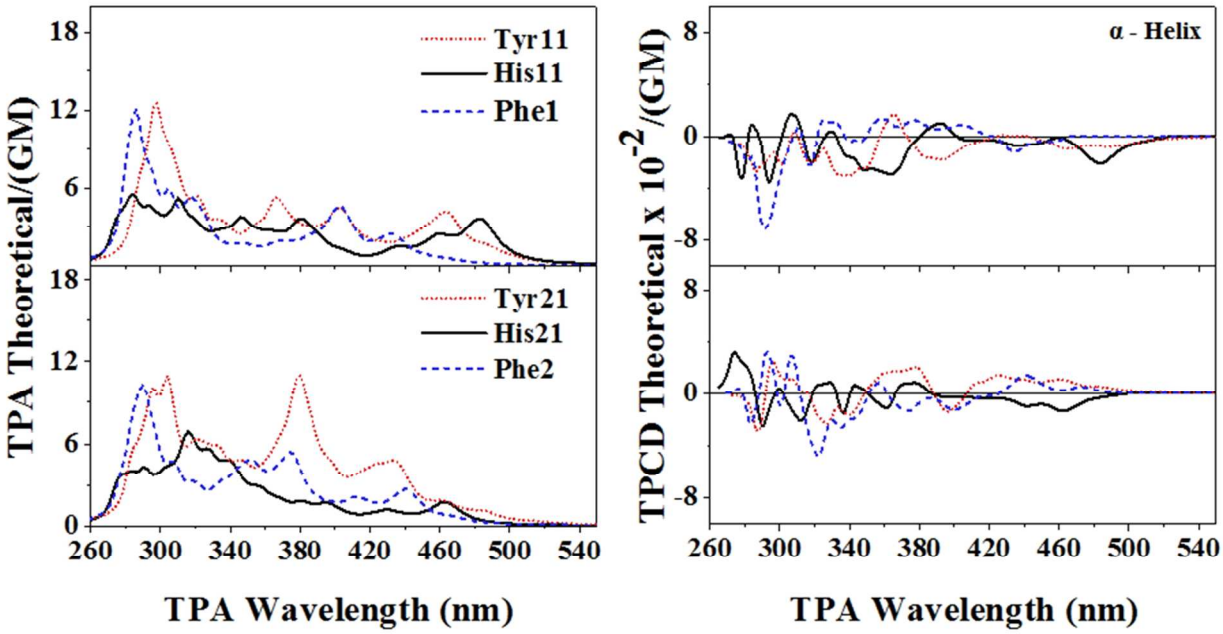

Figure 12.

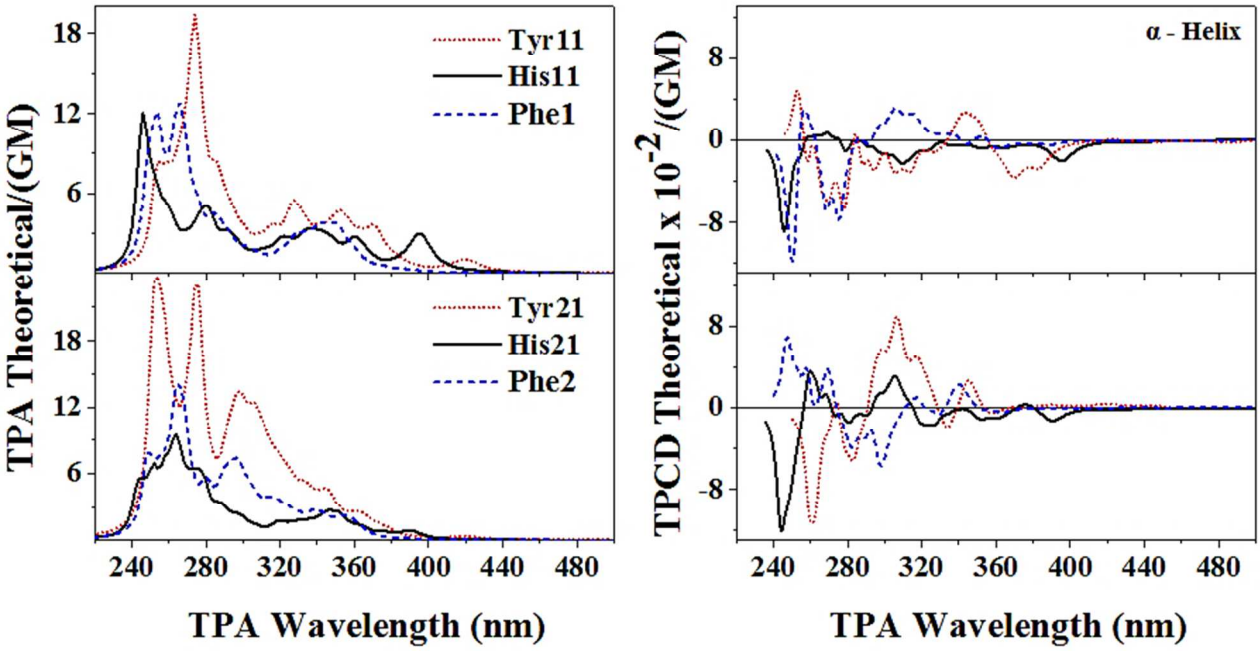

Figure 13.

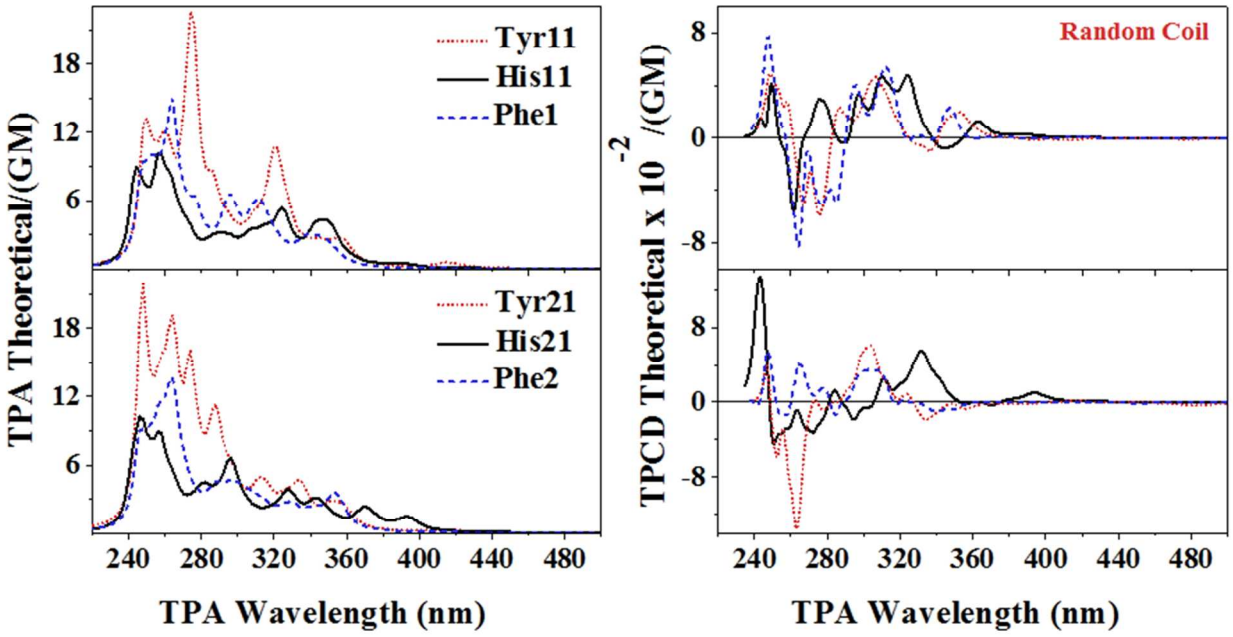

Figure 14.

Table I. Optimized Ramachandran dihedral angles (ψ and ϕ) for the main-chain amide group in L-hisditine and the corresponding angles (χ_1 and χ_2) for the rotational conformers defining the residues (all in degree). Angles are reported for His residues in random coil, α-Helix, β-Strand configuration.

Main-Chain Configuration	Side-Chain Configuration	ϕ	Ψ	χ_1	χ_2
Random-Coil	11	-83.487	80.634	-165.003	85.845
	12	-83.211	70.474	-170.189	110.403
	21	-80.309	77.357	-73.281	51.317
	22	-84.619	72.924	-49.889	113.134
	31	-82.622	68.860	49.524	-111.039
	32	-83.272	65.803	43.188	-117.085
β-Strand	11	-162.708	171.775	-125.454	61.882
	12	-157.729	152.839	-179.823	76.793
	21	-118.667	136.089	-65.477	-14.878
	22	-132.038	153.333	-57.576	-26.088
	31	-149.376	132.726	60.646	-49.944
	32	-150.902	175.458	68.070	-60.460
α-Helix	11	-61.041	-41.509	-165.419	139.644
	12	-61.932	-41.118	-171.243	176.085
	21	-60.080	-41.501	-73.144	-4.945
	22	-63.000	-40.556	-71.054	31.273
	31	-58.168	-46.167	56.581	-110.636
	32	-61.461	-40.596	52.179	-99.841

Table II. Optimized Ramachandran dihedral angles (ψ and ϕ) for the main-chain amide group in L-tyrosine and the corresponding angles (χ_1 and χ_2) for the rotational conformers defining the

residues (all in degree). Angles are reported for Tyr residues in random coil, α-Helix, β-Strand configuration.

Main-Chain Configuration	Side-Chain Configuration	ϕ	Ψ	χ_1	χ_2
Random-Coil	11	-83.172	79.274	-163.680	91.593
	12	-83.164	78.822	-163.729	90.941
	21	-84.599	73.952	-54.497	-67.299
	22	-84.605	73.564	-55.577	109.712
	31	-82.998	60.807	42.379	-101.330
	32	-82.938	60.709	41.995	-102.427
β-Strand	11	-158.147	164.174	-161.169	71.103
	12	-157.355	161.546	-164.719	70.301
	21	-117.278	140.859	-61.060	96.373
	22	-124.004	143.438	-57.805	98.939
	31	-153.901	166.312	59.342	-91.861
	32	-154.404	167.351	59.864	-91.742
α-Helix	11	-56.374	-41.444	-179.701	109.164
	12	-55.709	-41.421	-179.407	89.914
	21	-58.532	-40.313	-48.846	97.869
	22	-58.947	-40.619	-43.136	102.711
	31	-55.253	-41.267	54.833	-75.374
	32	-53.901	-41.385	61.283	-89.946

Table II. Optimized Ramachandran dihedral angles (ψ and ϕ) for the main-chain amide group in L-phenylalanine and the corresponding angles (χ_1 and χ_2) for the rotational conformers defining the residues (all in degree). Angles are reported for Phe residues in random coil, α-Helix, β-Strand configuration.

Main-Chain Configuration	Side-Chain Configuration	ϕ	Ψ	χ_1	χ_2
Random-Coil	1	-83.220	79.360	-163.256	89.235
	2	-84.603	73.399	-55.295	111.016
	3	-83.005	60.863	41.819	-101.829
β-Strand	1	-157.892	161.490	-164.491	70.138
	2	-116.854	139.585	-60.891	96.454
	3	-154.392	166.737	60.085	-90.473
α-Helix	1	-72.924	-41.633	173.254	-109.940
	2	-57.545	-40.589	-69.998	131.189
	3	-61.588	-40.714	52.867	-95.292

TOC