



Enhancement of the chiroptical response of α -amino acids via *N*-substitution for molecular structure determination using vibrational circular dichroism

Prasad L. Polavarapu¹ | Ernesto Santoro¹ | Cody L. Covington^{1,2} | Vijay Raghavan¹

¹Department of Chemistry, Vanderbilt University, Nashville, Tennessee

²Department of Chemistry, Austin Peay State University, Clarksville, Tennessee

Correspondence

Prasad L. Polavarapu, Department of Chemistry, Vanderbilt University, Nashville, TN 37235, USA.

Email: prasad.l.polavarapu@vanderbilt.edu

Funding information

National Science Foundation, Grant/Award Number: CHE1464874; National Science Foundation, Grant/Award Number: CHE-1464874

Abstract

The chiroptical response in the form of vibrational circular dichroism (VCD) in the midinfrared region is found to be enhanced when a hydrogen of amino group of L-tryptophan is substituted with acetyl, acryloyl, or maleyl group. The order of preference for VCD enhancement is found to be acryloyl > acetyl > maleyl group. The resulting experimental VCD spectra are also found to be satisfactorily reproduced by the quantum mechanical (QM) predicted spectra. The QM predicted spectra were simulated using the conformer populations, (a) predicted by Gibbs energies and (b) optimized to maximize the similarity between experimental and predicted VCD spectra. It is found that the conformer populations predicted by Gibbs energies do not yield the maximum possible similarity between experimental and the QM predicted spectra. This work identifies the *N*-substitution of α -amino acids and determining the conformer populations that best reproduce the experimental spectra as two new approaches for molecular structure determination.

KEYWORDS

amino acids, conformer populations, quantum theoretical calculations, tryptophan, VCD

1 | INTRODUCTION

Vibrational circular dichroism (VCD) is being used frequently in recent years for the absolute configuration (AC) determination, along with inferring predominant conformations.¹ These goals are facilitated by the quantum mechanical (QM) predictions of VCD spectra² using density functional theory (DFT) methods and analyzing the experimental spectra with predicted spectra.³ This approach is becoming routine when the sample under consideration has good solubility in solvents

whose vibrational absorption (VA) bands do not interfere in the spectral region of measurement, and measured VCD signals are significantly above the noise level. For cases where the sample of interest suffers from low solubility in suitable solvents and/or exhibits only weak VCD signals, chemical modifications to increase the solubility, and/or enhance the VCD signals, would be required.

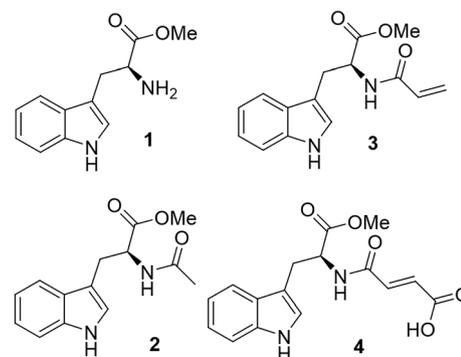
Because of the poor solubility in organic solvents, VCD measurements on simple amino acids are typically limited to the aqueous solutions. Since H₂O solvent has strong infrared absorption at approximately 1650 cm⁻¹, very high concentrations employing narrow (6 micron) pathlength cells had to be used for VCD measurements

Prasad L. Polavarapu, Ernesto Santoro, Cody L. Covington, and Vijay Raghavan contributed equally to this work

on α -amino acids in the midinfrared region.⁴ To avoid these limitations, the general practice has been to make use of the D₂O solvent whose VA bands in the mid-infrared region shift to lower wavenumbers. VCD spectra in D₂O solvent can be measured with 50 micron path length cells in the approximately 1800 to 1300 cm⁻¹ region. Even then, VCD signals for unsubstituted amino acids are often weak in this limited vibrational region. These difficulties led to the measurement of the VCD spectra of α -amino acids in the film^{5a} state. Enhanced VCD signals were obtained when proline, valine, and alanine are coupled to paramagnetic metal ion.^{5b}

The tryptophan-based compounds are important as building blocks of the protein systems.⁶ As a result, determination of the ACs, and of inferring predominant conformations of tryptophan-based compounds, acquires special significance. The VCD spectrum of tryptophan in the approximately 1800 to 1300 cm⁻¹ region could not be measured in aqueous solutions because of its poor solubility but could be measured in the film state^{5b}. Yet the measured signals are quite weak. Moreover, the film-state measurements do not correspond to realistic environments for biochemical applications and are not easy to interpret using theoretical methods. These limitations necessitate the exploration of better approaches to measure the VCD spectra of tryptophan-based compounds in solution phase.

Here, we present a practical method to amplify the VCD response of tryptophan. Tryptophan can be made soluble in organic solvents, that are more suitable for VCD spectroscopy, by esterification with the methyl ester being the simplest for QM predictions. However, esterification does not necessarily enhance the resulting VCD signals. It is known that the presence of delocalized electrons among neighboring C=C and C=O groups enhances the specific rotations.⁷ In the same vein, the introduction of groups with vibrations involving delocalized electrons can be expected to lead to vibrationally modulated delocalized electrons that can be anticipated to amplify the VCD signals. Based on this hypothesis, tryptophan methyl ester, **1**, (see **Scheme 1**) was derivatized via *N*-substitution with acetyl, **2**, acryloyl, **3**, and maleyl, **4**, groups. While acetyl group introduces one C=O group that can couple to the existing C=O group and provide VCD enhancement, acryloyl group introduces an additional mechanism of delocalizing the electrons among neighboring C=O and C=C groups; the maleyl group introduces one more C=O group for extending this delocalization. Therefore, *N*-acetyl, *N*-acryloyl, and *N*-maleyl substitutions might provide insight into how molecular vibrations might influence the circular dichroism associated with vibrational transitions in these groups.



SCHEME 1 Structures of L-tryptophan-Me-ester (**1**), *N*-acetyl-L-tryptophan-Me-ester (**2**), *N*-acryloyl-L-tryptophan-Me-ester (**3**), *N*-maleyl-L-tryptophan-Me-ester (**4**)

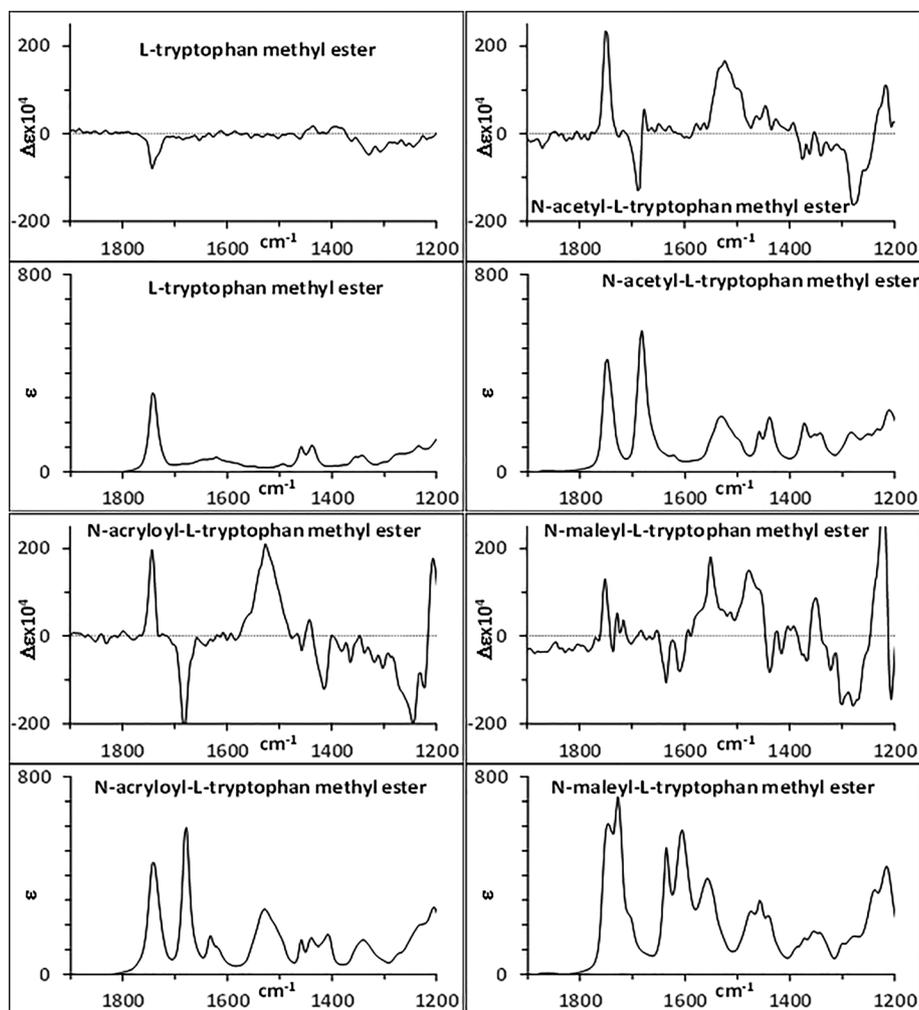
Here, we explore the role of acetyl, acryloyl, and maleyl groups in amplifying the VCD signals associated with tryptophan, as a representative case for α -amino acids, by measuring their experimental VCD spectra. Analysis of these spectra reveals the preferred groups for amplification of VCD signals. The quantitative analysis of similarity between experimental and QM predicted VCD and vibrational dissymmetry factor (VDF) spectra^{3c, 8} is also undertaken for all four molecules. While it is customary to simulate the predicted spectra using conformer populations derived from Gibbs energies, a different approach of determining the conformer populations that maximize the similarity between experimental and simulated VCD spectra is undertaken. This analysis indicates that conformer populations derived from Gibbs energies do not necessarily provide the maximum possible agreement between experimental and QM predicted VCD spectra. Therefore, the determination of the conformer populations that maximize the similarity between experimental and simulated VCD spectra will serve as a useful tool for future research.

2 | MATERIALS AND METHODS

2.1 | Experimental methods

Compounds **1-4** (see Scheme 1) have been synthesized (see ESI for details) starting from L-tryptophan-HCl, through methyl esterification (**1**), followed by substitution of a hydrogen atom of the NH₂ group with acetyl, acryloyl, or maleyl group (**2-4**). The synthetic process does not alter the central chirality, and therefore, the AC of parent compound was retained in **2-4** to be (*S*). The experimental VA and VCD (see Figure 1) spectra were measured in THF-d₈ solvent using a ChiralIR spectrometer, an SL3 cell with BaF₂ windows and 100- μ m path

FIGURE 1 Experimental VCD spectra of L-tryptophan methyl esters



length. The concentrations used are 0.24 (**1**), 0.11 (**2**), 0.11 (**3**), and 0.10 M (**4**).

2.2 | Computational methods

For each of the four molecules studied, conformational analysis was carried out for the (*S*)-enantiomer using CONFLEX program⁹ with built-in MMFF94 force field. In the case of compound **4**, both *cis*- and *trans*-conformers around the double bond of the maleyl substituent have been taken into account. The CONFLEX-generated conformers, within 20-kcal/mol energy window, have been optimized with PM6 method using Gaussian 16 program.¹⁰ The resulting unique conformers within the 10-kcal/mol energy range have been further optimized using B3LYP functional and 6-31G* basis set in gas phase. The resulting conformers, within 5 kcal/mol, have been further optimized at the B3LYP/6-311++g(2d,2p) level with the THF solvent represented by the polarizable continuum model (PCM) using the integral equation formalism (IEFPCM).¹¹ The number of optimized conformers within 2-kcal/mol

energy range is 15 for **1**, 9 for **2**, 25 for **3**, and 15 for **4**. The conformers of **1-4**, with populations greater than 10% are displayed in ESI (Figures S1-S4). These conformers have no imaginary vibrational frequencies, indicating that they are at the potential energy minima. VCD/VA spectra were computed for all of these conformers using B3LYP functional and 6-311++g(2d,2p) basis set with the THF-*d*₈ solvent represented by IEFPCM. Gibbs free energies have been used to derive the populations of conformers at 298.15 K. The simulated spectra were obtained as the sum of population weighted conformer spectra.

The quantification of agreement between experimental and predicted spectra is determined from spectral similarity overlap (SSO) as a function of vibrational frequency scale factor. The *Sim* functions, *SimVA*, *SimVCD*, and *SimVDF* (see ESI for description of these functions) are used to calculate the SSO as implemented in the CDSpecTech program.¹²

It has been a common practice to derive the conformer populations using Gibbs energies and generate population weighted QM predicted spectra for comparison with the experimental spectra. Two recent papers highlighted the uncertainties in Gibbs energies

calculated using QM theories. Nicu et al¹³ stated that the uncertainties in Gibbs energies can be as large as 2 kcal/mol. More recently,¹⁴ it was reported that even larger errors are possible in DFT calculations that use grids smaller than UltraFine grids,¹⁰ because the energies depend on the molecular orientation used for geometry optimization when smaller grids are used. As a result, the reliability of populations of conformers derived from Gibbs energies becomes uncertain. To alleviate these uncertainties, one can independently determine the conformer populations by optimizing them to provide the best possible agreement between population weighted QM predicted spectra and experimental spectra. A comparison of these optimized populations with the populations derived from Gibbs energies will reveal whether or not the latter populations are best suited for reproducing the experimental VCD spectra.

We have taken this new direction in this work by developing in-house a similarity optimization algorithm (see ESI) for optimizing the conformer populations that maximize the *SimVCD*, between experimental and QM predicted VCD spectra for **1-4**. The maximized *SimVCD* values reflect the best possible agreement that one can obtain, at the level of theory used, between experimental spectra and population weighted QM predicted spectra for the (*S*) enantiomers. For situations where VCD spectra are used for determining the unknown AC, one also has to verify if a different combination of conformer populations might also lead to an agreement between experimental spectra and population weighted QM predicted spectra for the opposite enantiomer. For this purpose, the similarity optimization algorithm developed here simultaneously optimizes the populations to minimize the *SimVCD* (ie, find the conformer populations that lead to negative *SimVCD* with maximum magnitude). The minimized *SimVCD* values correspond to conformer populations that give the best possible agreement between the experimental spectra and population weighted QM predicted spectra of opposite enantiomer. As a result, one will be able to assess, from maximum and minimum *SimVCD* values, if different sets of conformer populations might lead to opposite conclusions on AC.

3 | RESULTS AND DISCUSSION

3.1 | Enhancement of experimental VCD signals

The comparison of experimental VCD spectra for the four compounds reveals that **1** possesses the smallest number of VCD peaks, that are also weak. The weak VCD

associated with C=O stretching band at approximately 1743 cm⁻¹ and a weaker broad negative VCD in the 1362 to 1240 cm⁻¹ region are the only two pieces of information for interpreting the spectrum and to extract structural information of **1**. *N*-acetyl substitution results in the generation of bisignate VCD bands, positive at 1751 cm⁻¹ and negative at 1690 cm⁻¹, with significantly larger intensities and associated with well-resolved absorption bands. There are also two prominent additional VCD bands, positive at 1524 cm⁻¹ and negative at 1277 cm⁻¹. Compared with the VCD spectrum of **1**, enhanced VCD information content is clearly apparent in the spectrum of **2**. The acetyl and acryloyl substituents differ only in the presence of a terminal double bond in the latter, which introduces electron delocalization among C=O and C=C bonds. Accordingly, the VCD signals seen for acryloyl substitution are even larger than those seen for *N*-acetyl substitution. *N*-acryloyl substitution results in significantly larger intensities for bisignate VCD bands, positive at 1744 cm⁻¹ and negative at 1686 cm⁻¹, also associated with well-resolved absorption bands. In addition, prominent bands appear at 1524 cm⁻¹ with positive VCD, at 1415 cm⁻¹ with negative VCD, and at 1245 cm⁻¹ with negative VCD. *N*-maleyl substitution results in a “crowding effect” because of the presence of three C=O vibrational transitions in a closely spaced region around 1700 cm⁻¹. This “crowding” results in the disappearance of well-resolved VCD bands, unlike those seen for **2** and **3**, in the 1770 to 1660 cm⁻¹ region. The 1650 to 1250 cm⁻¹ region however appears to be rich in content with additional VCD bands, just as for **2** and **3**.

The substitution of one hydrogen of the amino group generated new bands in compounds **2-4**, especially the one originating from the N–H bending that was not present in the spectrum of **1**. This band, originating from coupling between N–H bending and C*–H bending, is identified through normal mode analysis (vide infra), as the positive broad band at 1524 cm⁻¹. Very interesting observation here is that the sign of N–H bending band is the same in all three *N*-substituted compounds, **2-4**. Since this vibrational chromophore is very close to the chiral center, and is perturbed directly by it, it is possible to use N–H bending VCD as a fingerprint for the AC assignment of the *N*-substituted α -amino acids. It is also interesting to notice that the well-resolved bisignate couplet, with positive VCD at approximately 1740 cm⁻¹ and negative VCD at approximately 1680 cm⁻¹, resulting from *N*-substitution, is reproduced for compounds **2** and **3** but that in the compound **4** is not well resolved because of “crowding” effect.

Thus, from the visual comparison of VCD spectra of **1-4**, it appears that *N*-substitution provides a practical route to enhance the VCD content for structural

investigation of tryptophan-based compounds. In the C=O stretching region, while acetyl and acryloyl groups provided well-resolved VCD bands, the maleyl group appears to be less advantageous because of possible cancelation from oppositely signed VCD band overlap.

Although a clear advantage for *N*-acetyl and *N*-acryloyl substitutions is evident from the experimental VCD spectra in Figure 1, this advantage may not be

useful if the experimental VCD spectra cannot be satisfactorily reproduced using QM predictions. Introduction of *N*-substitution can increase conformational flexibility, and therefore, it is necessary to verify the practical utility of these *N*-substitutions for structural determination by successfully reproducing their experimental spectra using QM predicted VCD spectra. This aspect is discussed next.

FIGURE 2 Comparison between experimental (THF- d_8) and computed VCD/VA spectra of **1** at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory. In the left vertical panel, the predicted wavenumbers are scaled with 0.984 and overlaid on experimental spectra. In the right vertical panel, QC predicted spectra with unscaled wavenumbers are stacked over experimental spectra and band positions labeled

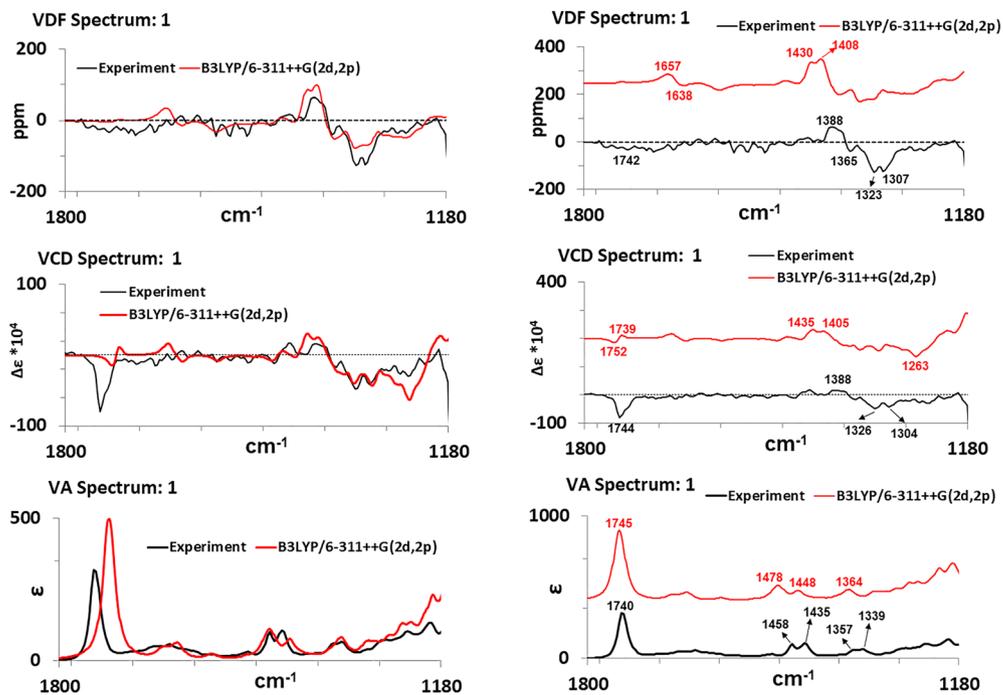
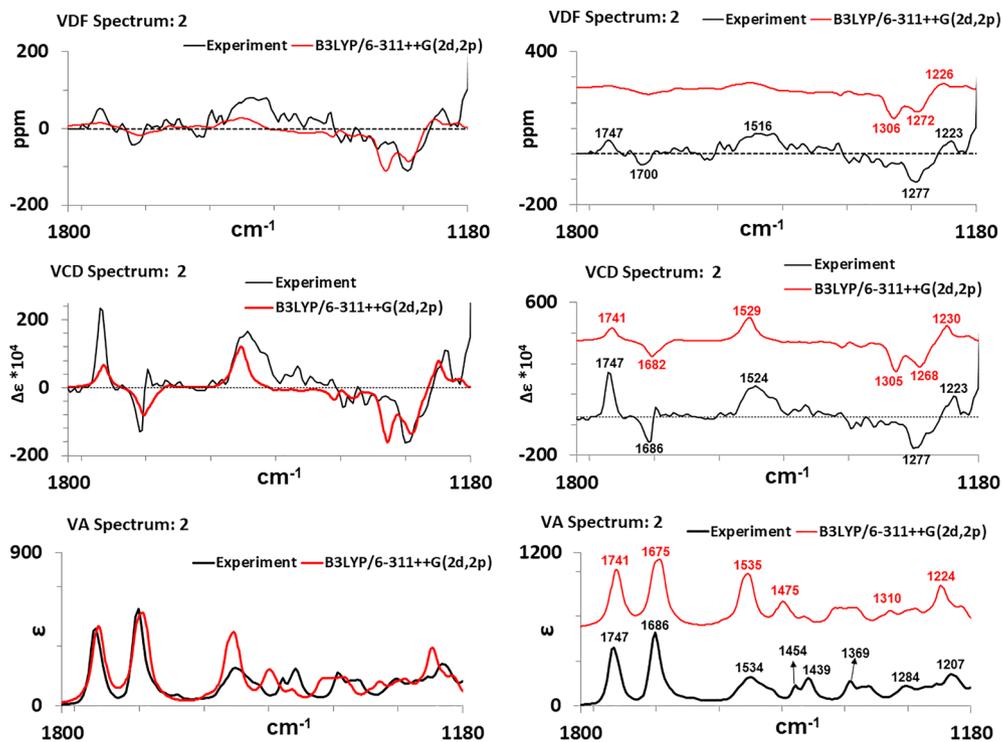


FIGURE 3 Comparison between experimental (THF- d_8) and computed VCD/VA spectra of **2** at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory. In the left vertical panel, the predicted wavenumbers are scaled with 0.997 and overlaid on experimental spectra. In the right vertical panel, QC predicted spectra with unscaled wavenumbers are stacked over experimental spectra and band positions labeled



4 | COMPARISON OF EXPERIMENTAL AND QM PREDICTED VCD SPECTRA

4.1 | Conformer populations determined from Gibbs energies

The dominant conformers of **1-4**, with populations greater than 10% are displayed in ESI. Experimental

VA, VCD, and VDF spectra of **1-4** are compared with corresponding predicted spectra in Figures 2–5, where the predicted spectra used the populations determined from Gibbs energies. In the left vertical panel, the computed spectra are scaled by 0.984 for **1**, 0.997 for **2**, 0.992 for **3**, and 0.983 for **4** and overlaid on the experimental spectra. The scale factors used here correspond to the maximum value of *SimVCD*. In the right vertical panel, the computed

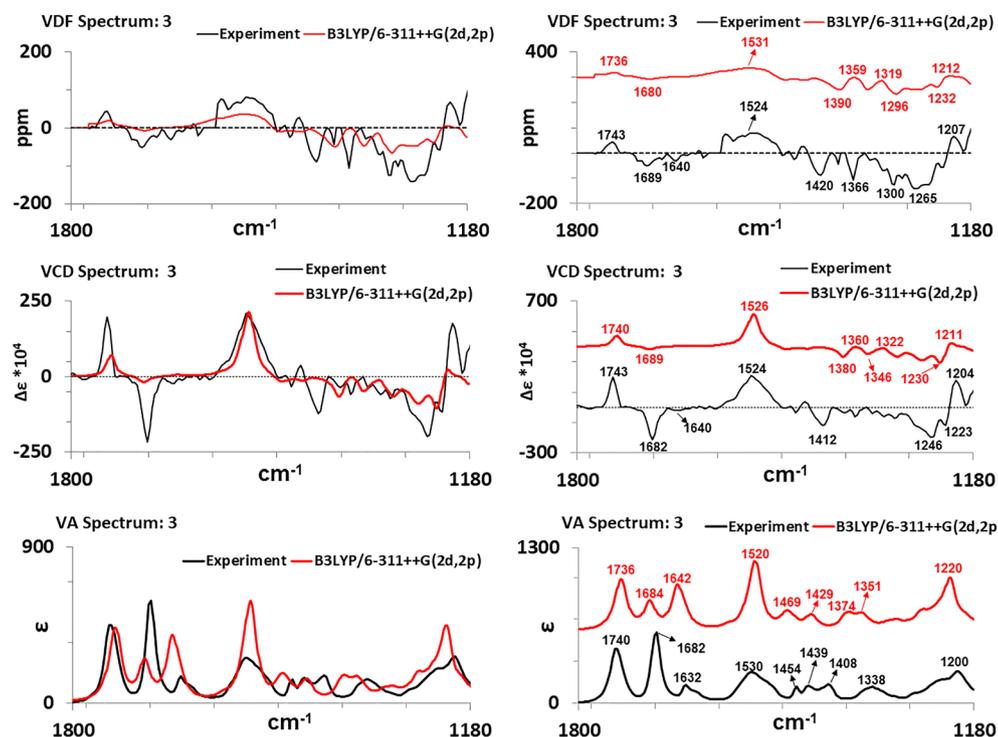


FIGURE 4 Comparison between experimental (THF- d_8) and computed VCD/VA spectra of **3** at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory. In the left vertical panel, the predicted wavenumbers are scaled with 0.992 and overlaid on experimental spectra. In the right vertical panel, QC predicted spectra with unscaled wavenumbers are stacked over experimental spectra and band positions labeled

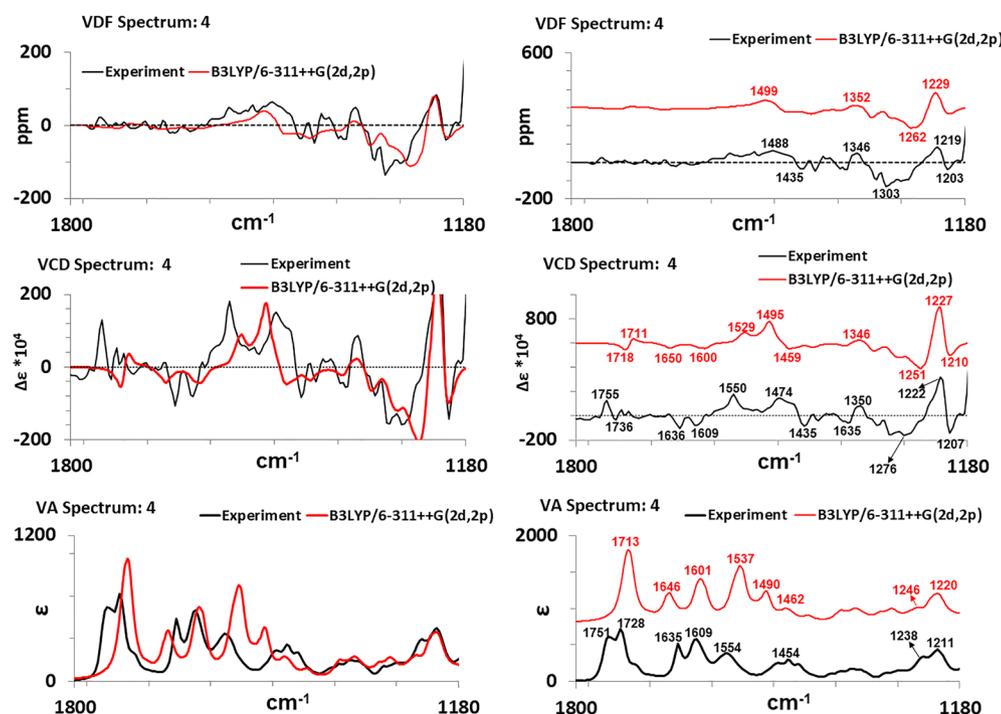


FIGURE 5 Comparison between experimental (THF- d_8) and computed VCD/VA spectra of **4** at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory. In the left vertical panel, the predicted wavenumbers are scaled with 0.983 and overlaid on experimental spectra. In the right vertical panel, QC predicted spectra with unscaled wavenumbers are stacked over experimental spectra and band positions labeled

spectra are stacked over the experimental spectra and not scaled.

The analysis of the normal modes, undertaken using the lowest energy conformer, indicates the following: For **1**, the negative strong VCD band at 1740 cm^{-1} corresponds to the C=O stretching. For **2**, the positive VCD band at 1747 cm^{-1} is due to tryptophan C=O stretching, negative band at 1686 cm^{-1} is due to acetyl C=O stretching, and a positive broad band at 1524 cm^{-1} is due to N–H bending coupled with C*–H bending. The latter signal is a direct result of the *N*-substitution. For **3**, a positive band at 1743 cm^{-1} , because of tryptophan C=O stretching, a negative band at 1682 cm^{-1} , because of acryloyl C=O stretching, and a positive broad band at 1524 cm^{-1} , because of N–H bending coupled with C*–H bending, are analogous to those for **2**. For **4**, the positive VCD band at 1755 cm^{-1} is due to tryptophan C=O stretching; the negative VCD band at 1736 cm^{-1} is due to maleyl terminal C=O stretching coupled to OH maleyl bending; the negative VCD band at 1636 cm^{-1} is due to maleyl C=C stretching; a negative VCD band at 1609 cm^{-1} is due to maleyl C=O stretching close to NH; and a positive broad VCD band at 1550 cm^{-1} is due to N–H bending coupled with C*–H bending.

Quantitative agreement between experimental and predicted spectra is assessed using SSO plots. The maximum *SimVA*, *SimVCD*, and *SimVDF* values are summarized in Table 1, and SSO plots are shown in Figure 6. The experimental region used for SSO calculation is approximately 1850 to 1200 cm^{-1} . From Table 1, it can be seen that the maximum *SimVCD* and *SimVDF* values obtained with populations derived from Gibbs energies are generally satisfactory. We previously recommended¹⁵ *SimVCD* and *SimVDF* values of at least 0.4 for a satisfactory agreement. The *SimVCD* and *SimVDF* values are 0.50 and 0.67 for **1**, 0.54 and 0.44 for **2**, 0.66 and 0.70 for **3**, and 0.51 and 0.55 for **4**. Larger values are obtained for **3**, indicating that *N*-acryloyl substitution is better suited for reproducing the experimental spectra. Since experimental spectra of **3** also exhibit well-resolved and large VCD signals in the 1800 to 1600 cm^{-1} range, it is apparent that *N*-acryloyl substitution is to be preferred among the three *N*-substitutions investigated.

4.2 | Conformer populations optimized to maximize the similarity between experimental and predicted VCD spectra

Figures 2–6 represent the standard practice used for analyzing the experimental spectra for molecular structure determination. However, as mentioned earlier, recent findings^{13,14} indicate that the conformer populations

TABLE 1 SSO obtained using populations derived from Gibbs energies and those optimized to yield maximum and minimum *SimVCD*^a

	Compound 1			Compound 2			Compound 3			Compound 4		
	GE ^b	PopOpt ^a		GE ^b	PopOpt ^a		GE ^b	PopOpt ^a		GE ^b	PopOpt ^a	
	Maximum	Minimum		Maximum	Minimum		Maximum	Minimum		Maximum	Minimum	
<i>SimVA</i>	0.91	0.90	0.87	0.88	0.87	0.76	0.72	0.73	0.62	0.81	0.81	0.56
<i>SimVCD</i>	0.50	0.75	−0.31	0.54	0.68	−0.22	0.66	0.77	−0.09	0.51	0.60	−0.10
<i>SimVDF</i>	0.67	0.67	−0.21	0.44	0.51	−0.05	0.70	0.75	−0.04	0.55	0.69	0.00

^aPopulations were optimized to maximize, as well minimize, the similarity overlap between experimental and predicted VCD spectra.

^bPopulations derived from Gibbs energies.

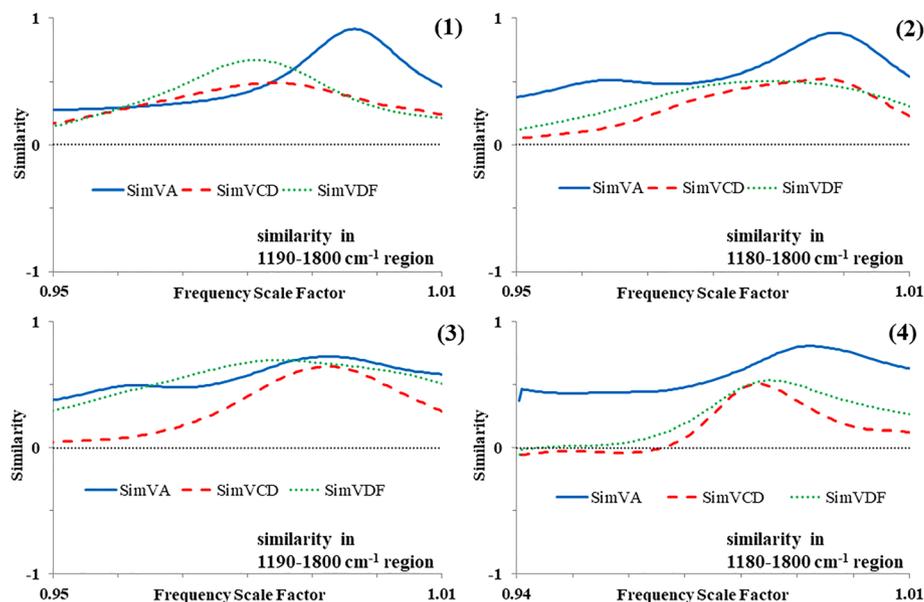


FIGURE 6 SSO plots comparing experimental VA, VCD, and VDF spectra with those predicted for compounds **1-4** using Boltzmann populations derived from Gibbs energies at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory

derived from Gibbs energies can be uncertain. To verify if the Gibbs energies-derived conformer populations are suitable for achieving maximum *SimVCD* and *SimVDF* values, a similarity optimization algorithm, independent of that reported by Nicu et al¹³, has been developed in-house for optimizing the conformer populations to maximize *SimVCD* (see ESI). At the same time, to verify if a different set of conformer populations might also lead to the assignment of opposite AC, conformer populations have also been optimized to minimize the *SimVCD* (which corresponds to negative *SimVCD* with maximum magnitude). The *SimVA*, *SimVCD*, and *SimVDF* values derived in this manner are also included in Table 1. The SSO plots, for experimental and optimized-population weighted QM predicted spectra, are included in ESI (Figures S5-S8). The experimental and optimized-population weighted QM predicted VCD spectra are compared in Figures 7.

From Table 1, it can be seen that the maximum *SimVCD* and *SimVDF* values obtained with optimized populations are uniformly larger than those obtained with Gibbs energies-derived populations. For **1**, the negative strong VCD band at 1740 cm^{-1} , which corresponds to the C=O stretching, is not reproduced well when populations derived from Gibbs energies are used (Figure 2). This deficiency is rectified when populations were optimized to maximize *SimVCD* (see top left panel in Figure 7). For **2**, the intensity of strong positive VCD band at approximately 1747 cm^{-1} is not reproduced well when populations derived from Gibbs energies are used (Figure 3). This situation has improved when populations were optimized to maximize *SimVCD* (see top right panel in Figure 7). The predicted intensity for the negative band at 1682 cm^{-1} of **3** is rather small when populations

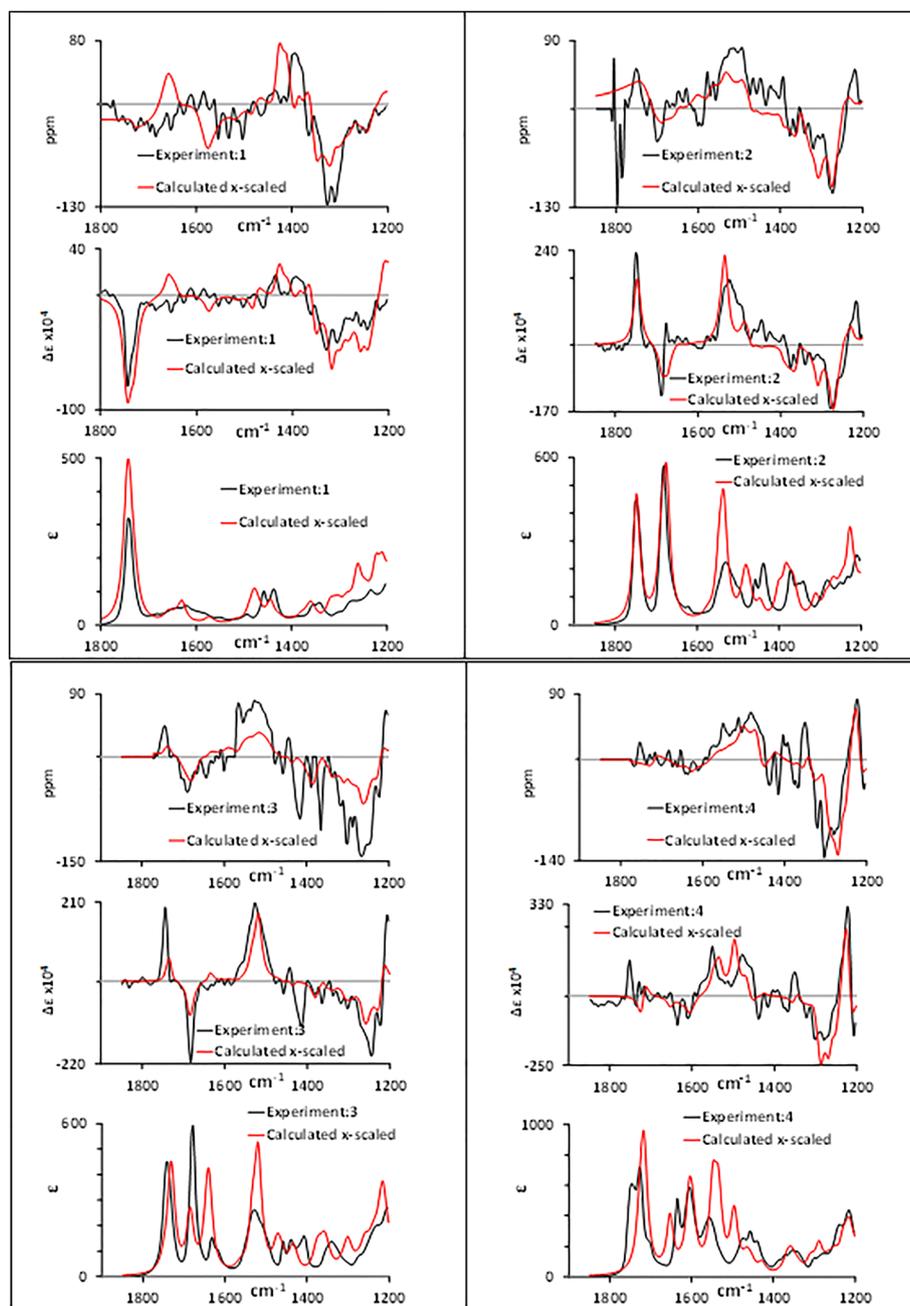
derived from Gibbs energies are used (see Figure 4) but not so when optimized populations are used (see bottom left panel in Figure 7). For **4**, the comparison between experimental and predicted spectra has improved around the approximately 1300 cm^{-1} region, when optimized populations are used (compare Figure 5 and bottom right panel in Figure 7).

When Gibbs energies-derived populations are replaced with conformer populations optimized with similarity optimization algorithm, *SimVCD* value increases from 0.50 to 0.75 for **1**, and *SimVDF* remains about the same; similar increase from 0.54 to 0.68 is seen in *SimVCD* for **2**, 0.66 to 0.77 for **3**, and from 0.51 to 0.60 for **4**; *SimVDF* values have also become larger for **2**, **3**, and **4**. These observations confirm that the Gibbs energies-derived populations do not provide the best possible agreement with experimental VCD spectra. It is apparent that similarity optimization has improved not only *SimVCD* but also *SimVDF* values, a preferred outcome.

Of the four molecules investigated here, the *SimVCD* and *SimVDF* values for **4** are not as high as those for the other three. The addition of maleyl group with more conformational mobility, and multiple vibrations “crowded” in a narrow region, appears to lead to a lower quality comparison between experimental and theoretical spectra. On the other hand, the *SimVCD* and *SimVDF* values are higher for **3** than those of other molecules investigated.

Based on the current observations, the preference for *N*-substitution in tryptophan-based compounds for practical analysis is in the order: *N*-acryloyl > *N*-acetyl > *N*-maleyl. We anticipate the same observations to apply for other α -amino acids. However, it will be prudent to investigate other α -amino acids individually.

FIGURE 7 Comparison between experimental (THF- d_8) and computed VCD/VA spectra of **1-4** at DFT/B3LYP/6-311++G(2d,2p)/IEFPCM (THF) level of theory with populations optimized to maximize *SimVCD*



To assess, what changes in the populations have led to increased *SimVCD* and *SimVDF* values, the populations optimized to maximize the *SimVCD* are compared with those derived from Gibbs energies in Tables 2–5. We will focus only on those conformers that have >10% population (see Figures S1–S4 in ESI). For **1**, the Gibbs energy-derived populations are associated with conformers #12 (24%), #2 (18%), and #15 (14%) while optimized populations are associated with conformers #12 (33%), #15 (26%), #19 (16%), and #5 (14%); two conformers, #12 and #15, accounting for approximately 40% population are common for both cases. For **2**, the former are associated with #30 (40%) and #14 (25%), while the

latter are associated with #10 (52%) and #30 (36%). Here, one conformer, #30, accounting for approximately 40% population is common for both cases. For **3**, the former are associated with #17 (18%), #13 (16%), #9 (12%), and #15 (11%), while the latter are associated with #13 (28%), #44 (27%), #15 (12%), and #6 (11%); here, two conformers, #13 and #15, accounting for approximately 30% population are common for both cases. For **4**, the former are associated with #85 (25%) and #99 (22%), while the latter are associated with #77 (30%), #64 (25%), and #630 (18%), all belonging to the *cis*-maleyl; here, the conformers optimized to yield maximum *SimVCD* are different from those derived from Gibbs energies.

These observations suggest that optimized conformer populations can be significantly different from those derived from Gibbs energies. Since all of the conformers

within 2 kcal/mol energy range are used for VCD calculations, and the uncertainties in Gibbs energies are considered to be at least that much, the redistribution of

TABLE 2 Conformer populations for (*S*)-tryptophan methyl ester, **1**

Sequence Number Loaded Into CDSpectech #	Conformer # ^a	From Gibbs Energies	From Maximizing SSO for VCD	From Minimizing SSO for VCD	Δ (Population)/ Δ (SSO)	
					Max SSO ^b	Min SSO ^c
1	1	0.04	0.02	0.00	-0.08	-0.04
2	2	0.18	0.00	0.65	-0.70	0.53
3	3	0.06	0.00	0.01	-0.22	-0.06
4	4	0.03	0.00	0.28	-0.12	0.28
5	5	0.01	0.14	0.00	0.51	-0.01
6	7	0.01	0.00	0.06	-0.04	0.06
7	8	0.05	0.00	0.00	-0.21	-0.06
8	10	0.05	0.01	0.00	-0.18	-0.06
9	12	0.24	0.33	0.00	0.36	-0.27
10	15	0.14	0.26	0.00	0.46	-0.16
11	19	0.02	0.16	0.00	0.56	-0.02
12	20	0.04	0.02	0.00	-0.08	-0.04
13	21	0.08	0.00	0.00	-0.32	-0.09
14	23	0.05	0.07	0.00	0.06	-0.06
15	32	0.00	0.00	0.00	0.00	0.00

^aThese are numbers associated with Conflex generated conformers for identification purposes.

^bChange in population needed per unit change in maximum SSO for VCD; calculated as (SSOmaxPop-GibbsPop)/change in SSO.

^cChange in population needed per unit change in minimum SSO for VCD; calculated as (SSOminPop-GibbsPop)/Abs (change in SSO).

TABLE 3 Conformer populations for *N*-acetyl-(*S*)-tryptophan methyl ester, **2**

Sequence Number Loaded Into CDSpectech #	Conformer # ^a	From Gibbs Energies	From Maximizing SSO for VCD	From Minimizing SSO for VCD	Δ (Population)/ Δ (SSO)	
					Max SSO ^b	Min SSO ^c
1	1	0.06	0.00	0.12	-0.43	0.08
2	3	0.02	0.00	0.88	-0.11	1.14
3	6	0.09	0.00	0.00	-0.63	-0.12
4	10	0.06	0.52	0.00	3.27	-0.08
5	12	0.07	0.07	0.00	-0.06	-0.10
6	14	0.25	0.00	0.00	-1.78	-0.33
7	20	0.02	0.06	0.00	0.27	-0.03
8	26	0.03	0.00	0.00	-0.21	-0.04
9	30	0.40	0.36	0.00	-0.32	-0.53

^aThese are numbers associated with Conflex generated conformers for identification purposes.

^bChange in population needed per unit change in maximum SSO for VCD; calculated as (SSOmaxPop-GibbsPop)/change in SSO.

^cChange in population needed per unit change in minimum SSO for VCD; calculated as (SSOminPop-GibbsPop)/Abs (change in SSO).

TABLE 4 Conformer populations for *N*-acryloyl-(*S*)-tryptophan methyl ester, 3

Sequence Number Loaded Into CDSpectech #	Conformer # ^a	From Gibbs Energies	From Maximizing SSO for VCD	From Minimizing SSO for VCD	Δ (Population)/ Δ (SSO)	
					Max SSO ^b	Min SSO ^c
1	1	0.08	0.00	0.00	-0.73	-0.11
2	2	0.01	0.00	0.00	-0.07	-0.01
3	3	0.03	0.00	0.00	-0.26	-0.04
4	4	0.01	0.00	0.35	-0.07	0.45
5	6	0.00	0.11	0.00	0.95	0.00
6	7	0.00	0.00	0.65	-0.03	0.86
7	9	0.12	0.00	0.00	-1.06	-0.16
8	10	0.01	0.00	0.00	-0.10	-0.01
9	12	0.01	0.00	0.00	-0.10	-0.02
10	13	0.16	0.28	0.00	1.09	-0.22
11	15	0.11	0.12	0.00	0.03	-0.15
12	17	0.18	0.08	0.00	-0.89	-0.24
13	18	0.01	0.00	0.00	-0.10	-0.01
14	19	0.01	0.00	0.00	-0.13	-0.02
15	20	0.01	0.00	0.00	-0.05	-0.01
16	21	0.01	0.00	0.00	-0.10	-0.01
17	22	0.02	0.00	0.00	-0.16	-0.02
18	23	0.03	0.00	0.00	-0.24	-0.04
19	26	0.05	0.08	0.00	0.23	-0.07
20	29	0.03	0.06	0.00	0.23	-0.04
21	31	0.05	0.00	0.00	-0.48	-0.07
22	32	0.01	0.00	0.00	-0.05	-0.01
23	42	0.00	0.00	0.00	-0.04	-0.01
24	44	0.01	0.27	0.00	2.37	-0.02
25	46	0.03	0.00	0.00	-0.25	-0.04

^aThese are numbers associated with Conflex generated conformers for identification purposes.

^bChange in population needed per unit change in maximum SSO for VCD; calculated as (SSOmaxPop-GibbsPop)/change in SSO.

^cChange in population needed per unit change in minimum SSO for VCD; calculated as (SSOminPop-GibbsPop)/Abs (change in SSO).

populations among these conformers is not surprising. Since the optimized conformer populations provide the best possible agreement with experimental VCD spectra, these populations might be viewed as experimentally determined populations from VCD spectra. An independent verification of the optimized conformer populations in the solution phase using a different experimental technique however remains a challenge at the present time.

It is useful to note that the deduction of correct AC, for the four molecules investigated here, is independent of whether one uses Gibbs energies-derived populations or populations optimized to maximize the *SimVCD*, because both sets of conformer populations yielded high

enough *SimVCD* and *SimVDF* values to confidently assign the AC.

The conformer populations (see Tables 2–5) optimized to yield minimum *SimVCD* value, indicate that the magnitudes of negative *SimVCD* values (see Table 1) are much smaller than the corresponding magnitudes obtained with populations optimized to yield maximum (positive) *SimVCD* values. Thus, there is no possibility for inferring the opposite AC, for any of the four investigated compounds, from the present calculations.

The last two columns in Tables 2–5 summarize the ratio of change in conformer populations to change in *SimVCD*. These ratios provide insight into sensitivity, ie,

TABLE 5 Conformer populations for *N*-maleyl-(*S*)-tryptophan methyl ester, 4

Sequence Number Loaded Into CDSpectech #	Conformer # ^a	From Gibbs Energies	From Maximizing SSO for VCD	From Minimizing SSO for VCD	Δ (Population)/ Δ (SSO)	
					Max SSO ^b	Min SSO ^c
1	<i>cis</i> -11	0.08	0.00	0.00	-0.88	-0.13
2	<i>cis</i> -21	0.07	0.00	0.00	-0.81	-0.12
3	<i>cis</i> -23	0.03	0.07	0.00	0.42	-0.05
4	<i>cis</i> -64	0.07	0.25	0.00	2.07	-0.11
5	<i>cis</i> -77	0.04	0.30	0.00	2.85	-0.07
6	<i>cis</i> -79	0.04	0.00	0.00	-0.40	-0.06
7	<i>cis</i> -85	0.25	0.06	0.00	-2.16	-0.41
8	<i>cis</i> -99	0.22	0.08	0.00	-1.51	-0.36
9	<i>cis</i> -310	0.03	0.06	0.00	0.30	-0.05
10	<i>cis</i> -630	0.05	0.18	0.00	1.39	-0.09
11	<i>trans</i> -5	0.00	0.00	0.00	-0.05	-0.01
12	<i>trans</i> -11	0.01	0.00	0.00	-0.09	-0.01
13	<i>trans</i> -12	0.10	0.00	1.00	-1.07	1.48
14	<i>trans</i> -18	0.00	0.00	0.00	-0.02	0.00
15	<i>trans</i> -53	0.00	0.00	0.00	-0.03	0.00

^aThese are numbers associated with Conflex generated conformers for identification purposes.

^bChange in population needed per unit change in maximum SSO for VCD; calculated as (SSOmaxPop-GibbsPop)/change in SSO.

^cChange in population needed per unit change in minimum SSO for VCD; calculated as (SSOminPop-GibbsPop)/Abs (change in SSO).

how big or small change in populations is needed for unit change in *SimVCD*. A larger ratio would mean less effectiveness in influencing *SimVCD* and vice versa.

Investigation of conformer populations optimized to maximize/minimize *SimVCD* are in the very early stages. It should be remembered that several factors influence the optimized conformer populations. These factors

include (a) the theoretical level used for predicting the vibrational frequencies and intensities; (b) quality of the experimental spectra; and (c) the spectral range used for comparison. For sufficient confidence in the optimized populations, it is important to utilize most reliable theoretical levels and high signal to noise experimental spectra covering a large number of vibrational bands.

**FIGURE 8** Koji starting the Magic show

FIGURE 9 Koji and Prasad Polavarapu during the Magic show



FIGURE 10 Koji and Nobu Harada entertaining the audience during magic show



FIGURE 11 Conference Dinner at the University Club of Nashville, Vanderbilt University. Seated (L to R): Bharathi Polavarapu, Koji Nakanishi, Bob Woody, and Nina Berova; standing (L to R): Jim Cheeseman, Rina Dukor, Prasad Polavarapu, A-Young Woody, and Larry Nafie





FIGURE 12 Cruise Dinner aboard General Jackson river boat on Cumberland River in Nashville. L to R: Jochen Autschbach, Rina Dukor, Koji Nakanishi, a conference attendee, Pat Vaccaro, Yunjie Xu, Jim Cheeseman, Paul Nicu, and Markus Reiher

5 | CONCLUSION

The current investigation indicates that *N*-substitution can provide enhanced VCD signals for tryptophan-based compounds, in particular with acryloyl or acetyl substitutions. In the 1800 to 1500 cm^{-1} region, new VCD features that can be diagnostic of the chiral center are seen. The maleyl substitution is not as attractive because of the “crowding” effect in the aforementioned region resulting in poorly resolved and weaker VCD signals. The QM simulated spectra using conformer populations derived from Gibbs energies do not necessarily provide the best possible agreement with the experimental spectra. The conformer populations can be derived from the experimental VCD spectra by optimizing the populations that maximize the similarity between experimental and QM predicted VCD spectra. The conformer populations derived in this manner match those derived from Gibbs energies to a limited extent for (*S*)-tryptophan methyl ester, *N*-acetyl (*S*)-tryptophan methyl ester, and *N*-acryloyl (*S*)-tryptophan methyl ester. But for *N*-maleyl (*S*)-tryptophan methyl ester, optimized conformer populations are different from those derived from Gibbs energies. The determination of conformer populations that maximize the similarity between experimental and simulated VCD spectra is likely to become a useful aspect of analyzing the future experimental VCD spectra.

ACKNOWLEDGEMENTS

Funding from National Science Foundation (CHE-1464874) is gratefully acknowledged. This work was conducted in part using the resources of the Advanced

Computing Center for Research and Education (ACCRES) at Vanderbilt University.

DEDICATION

We are honored to dedicate this article to the memory of Professor Koji Nakanishi. Koji nurtured a huge number of students, colleagues, and scientists around the world. Koji is well known not only among organic chemists but also among chiroptical spectroscopists. Koji attended chiroptical spectroscopy conferences (formerly known as circular dichroism [CD] conferences) regularly and delighted the participants with his signature magic show. The 14th International Conference on Chiroptical Spectroscopy (CD2013), held at Vanderbilt University, Nashville, TN (June 9-13, 2013) was the last known scientific conference that Koji actively participated. We are pleased to present below some pictures (Figures 8–12) taken at this conference to honor the memory of Koji.

ORCID

Prasad L. Polavarapu  <https://orcid.org/0000-0001-6458-0508>

REFERENCES

- (a) Berova N, Polavarapu PL, Nakanishi K, Woody RW. *Comprehensive Chiroptical Spectroscopy*. John Wiley and Sons; 2012. (b) Nafie LA. *Vibrational Optical Activity*. John Wiley and Sons; 2011. (c) Polavarapu PL. *Chiroptical Spectroscopy: Fundamentals and Applications*. Taylor and Francis; 2017.
- (a) Buckingham AD, Fowler PW, Galwas PA. *Chem. Phys.* 1987;112:1-14. (b) Stephens PJ. *J. Phys. Chem.* 1985;89:748-752.

3. (a) Stephens PJ, Devlin FJ, Pan J. *Chirality*. 20; 2008:643-663. (b) Debie E, De Gussem E, Dukor RK, Herrebout W, Nafie LA, Bultinck P. *ChemPhysChem*. 2011;12:1542-1549. (c) Covington CL, Polavarapu PL. *J. Phys. Chem. A*. 2013;117:3377-3386.
4. (a) Freedman TB, Chernovitz AC, Zuk WM, Paterlini MG, Nafie LA. *J. Am. Chem. Soc.* 1988;110:6970-6974. (b) Diem M. *J. Am. Chem. Soc.* 1988;110:6967-6970. (c) Poopari MR, Dezhahang Z, Yang G, Xu Y. *ChemPhysChem*. 2012;13:2310-2321. (d) Zhu P, Yang G, Poopari MR, Bie Z, Xu Y. *ChemPhysChem*. 2012;13:1272-1281. (e) Poopari MR, Dezhahang Z, Xu Y. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2015;136:131-140. (f) Dobrowolski J, Lipiński PFJ, Rode JE, Sadlej, J. α -Amino Acids In Water: A Review Of VCD and ROA Spectra. In Baranska M, Editor *Optical spectroscopy and computational methods in biology and medicine*, Springer Netherlands, 2014, pp 83-160.
5. (a) Zhang P, Polavarapu PL. *Appl. Spectrosc.* 2006;60:378-385; (b) Domingos SR, Huerta-Viga A, Baij L, et al. *J. Am. Chem. Soc.* 2014;136:3530-3535.
6. Mahalakshmi R, Shanmugam G, Polavarapu PL, Balaram P. *ChemBioChem*. 2005;6:2152-2158.
7. (a) Caricato M, Vaccaro PH, Crawford TD, Wiberg KB, Lahiri P. *J. Phys. Chem. A*. 2014;118:4863-4871. (b) Moore B, Srebro M, Autschbach J, *J. Chem. Theory Comput.* 2012; 8: 4336-4346.
8. Johnson JL, Nair DS, Pillai SM, et al. Dissymmetry Factor Spectral Analysis Can Provide Useful Diastereomer Discrimination: Chiral Molecular Structure of an Analogue of (-)-Crispine A. *ACS Omega*. 2019;4(4):6154-6164.
9. Conflex: High Performance Conformational Analysis. www.conflex.net, November 17, 2019
10. Gaussian 16, Revision B.01, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich AV, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA Jr., Peralta JE, Ogliaro F, Bearpark MJ, Heyd JJ, Brothers EN, Kudin KN, Staroverov VN, Keith TA, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, and Fox DJ, Gaussian, Inc., Wallingford CT, 2016.
11. Mennucci B, Tomasi J, Cammi R, et al. *J. Phys. Chem. A*. 2002; 106:6102-6113.
12. Covington CL, Polavarapu PL. CDSpecTech: A single software suite for multiple chiroptical spectroscopic analyses. *Chirality*. 2017;29(5):178-192.
13. Koenis MAJ, Xia Y, Domingos SR, Visscher L, Buma WJ, Nicu VP. *Chem. Sci.* 2019;10:7661-7678.
14. Bootsma AN, Wheeler S, ChemRxiv 2019, <https://doi.org/10.26434/chemrxiv.8864204.v5>
15. Polavarapu PL, Covington CL, Raghavan V. To Avoid Chasing Incorrect Chemical Structures of Chiral Compounds: Raman Optical Activity and Vibrational Circular Dichroism Spectroscopies. *ChemPhysChem*. 2017;18(18):2459-2465.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Polavarapu PL, Santoro E, Covington CL, Raghavan V. Enhancement of the chiroptical response of α -amino acids via *N*-substitution for molecular structure determination using vibrational circular dichroism. *Chirality*. 2020;1-15. <https://doi.org/10.1002/chir.23205>