JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Statistical Validation of Absolute Configuration Assignment in Vibrational Optical Activity

Jelle Vandenbussche,[†] Patrick Bultinck,*[,†] Anna K. Przybył,[‡] and Wouter A. Herrebout[¶]

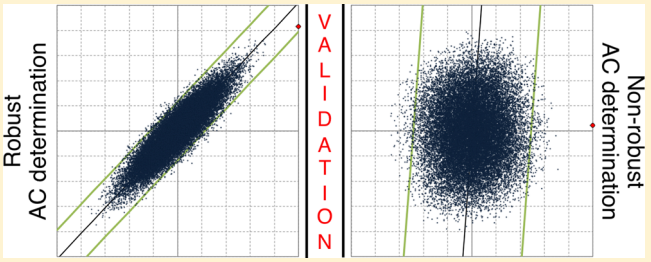
[†]Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S3), 9000 Ghent, Belgium

[‡]Faculty of Chemistry, Adam Mickiewicz University, Grunwaldzka 6, 60-780 Poznań, Poland

[¶]Department of Chemistry, University of Antwerp, Groenenborgerlaan 171 (G.V.023), 2020 Antwerp, Belgium

Supporting Information

ABSTRACT: Chiroptical spectroscopy usually requires theoretically computed spectra to assist in the elucidation of the absolute configuration of samples for which experimental spectra have been recorded. Due to the inherently different nature of these two types of spectra, perfect agreement is quasi impossible. Several methods exist to quantify the degree of similarity between the two spectra, but rather limited work has been done to evaluate the robustness of the similarity between theory and experiment. In this work, a novel method is described to determine the statistical significance of the numerical degree of similarity between experimental and calculated vibrational circular dichroism spectra and to offer valuable support for performing absolute configuration assignments. The approach is successfully applied to a number of quinolizidine alkaloids.



INTRODUCTION

Often experimental data are used to validate new quantum chemical methods through the comparison of computed data with experimental data. Despite usually inherent shortcomings in the computational modeling such as limited inclusion of solvation effects, incomplete basis sets or treatment of electron correlation and many more, good agreement with experiment is seen as evidence of the validity and applicability of the new method. On the other hand, some experimental observations can only be interpreted by comparison with computed data because the experimental data cannot be linked in a straightforward fashion to e.g., a molecular structure. Theory and experiment thus more and more develop into a symbiotic duo.

A good example of such a symbiotic duo is chiroptical spectroscopy which relies on the different response of different absolute configurations (AC) toward polarized light. Such techniques include electronic circular dichroism (ECD),[1] the polarized light analog of UV−vis spectroscopy, vibrational circular dichroism (VCD)[2,3] as analog of IR spectroscopy, and Raman optical activity (ROA)[4] as the polarized light extension of Raman spectroscopy. Whereas for the nonpolarized spectroscopies, the assignment of spectra can be based on numerous data sets describing typical absorptions linked to functional groups, the polarized counterparts are much harder to assign despite that the absorption frequencies are the same. The problem lies in the fact that the polarized counterparts are differential spectroscopies, and thus the absorption bands are characterized by either positive or negative absorption differences and simple rules of thumb are problematic. Typically, quantum chemical simulations of the differential absorption are used to generate theoretical spectra. When these spectra match sufficiently well the experimental ones, one concludes that the experimental sample has the same AC as that used in the simulations. However, a key issue in this is how to define "sufficiently well". The risk of bias toward deciding sufficiently high agreement is often present. The present paper has the aim of checking the robustness of quantitative measures expressing the agreement between theory and experiment. As correlation (similarity) is different from causality, extra tests have to be performed to determine whether a degree of similarity is truly robust.

The following focuses on VCD although the methods developed are in principle also applicable to ROA, ECD, and even dissymmetry factor spectra as in the recent approach of Covington and Polavarapu.[5] VCD has become a common technique for the determination of the AC of chiral natural and pharmaceutical products in the solution phase.[6] It has many advantages compared to alternative techniques[7] yet suffers the drawback that often-used quantum chemical methods have important limitations such as the necessity to omit explicit solute−solute and solvent−solute interactions, to discard anharmonicity effects etc., and to use limited basis sets and approximate treatments of electron correlation. As a result, theoretically obtained spectra differ from experiment both in band positions, intensities, and in the worst cases even in sign. Erroneous matching of VCD bands can partially be avoided by doing so in agreement with the IR absorption bands as these occur at the same frequencies. Nonetheless, the identification of

Received: September 25, 2013
Published: November 6, 2013

   dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

equivalent absorption bands is not straightforward and requires a certain level of experience.

In order to avoid human bias during manual peak-to-peak AC assignments, researchers tend to quantify the agreement between theory and experiment numerically. This was the basis of previously introduced similarity measures.[5,8−14] Even though these similarity measures provide an objective approach to describe the agreement between theory and experiment, they do not necessarily lead to straightforward conclusions with respect to AC assignments. The question may arise, for example, whether a particular VCD similarity of +50% is sufficient to ensure that the AC of a molecule is correctly determined. As a result, there is a need for knowing how good a similarity value is or, in other words, knowing whether the obtained result is statistically significant. In the literature several proposals have been formulated to address this significance issue, although most often not with robust statistical techniques used here. Some studies employ simple empirical rules, stating that the AC is confidently determined when the similarity rises above a particular value.[9] Such an approach, however, does not take into account that an experimental VCD pattern consisting of numerous positive and negative absorption bands is harder to reproduce than a pattern containing only a few bands. A lower similarity for the former can still be statistically more significant than a higher similarity for the latter. There may hence be a risk of putting the agreement in a wrong perspective when applying simple rules of thumb in order to determine the robustness of a similarity value.

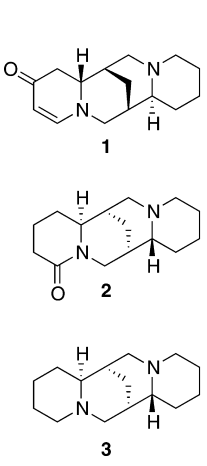
Shen and co-workers[10] recently addressed the robustness problem on a statistical basis. Their study shows that the VCD spectra of the calculated conformations for a particular flexible molecule can be divided into two groups of which one shows high similarity with experiment and the other does not. The similarity values of the latter group of conformers form a normal distribution from which, for example, twice the standard deviation allows to determine the 95% robustness level that a spectrum indeed fits the experiment. The authors merely intended to statistically validate the previously employed empirical robustness levels.[9] As mentioned above, however, these levels are unlikely transferable across different spectral regions or molecular systems, and one would therefore like to address the issue more directly. Unfortunately, the method described by Shen and colleagues is not generally applicable because for molecules having limited conformational flexibility the normal distribution cannot be adequately determined.

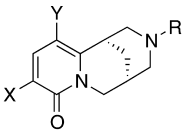
Another method, also known under the commercial trade name CompareVOA,[8] addressed the problem in a different way. It projects computed similarity values in a database of independently validated good and bad VCD-based AC assignments yielding a confidence level (CL). Even though the use of a database is a fundamentally correct approach, in practice it can be less straightforward for several reasons. To function optimally, the database size should be very large in order to represent the enormous variety of chirality in molecules and different experimental conditions (solvents, instrumentation, etc.), and the independent validation of AC determination for molecules in the database may be technically unachievable (for example, when no single crystals can be obtained for XRD analysis). By the projection against the database, a user can assess how good his assignment was with respect to earlier correct assignments. This serves a different purpose than the present paper. Up to now, there is no method that allows for determining the robustness of the similarity between any pair of spectra individually, and it is this gap that the present paper bridges. So, it answers the question whether a given similarity value is robust. It determines whether it is significantly better than the degree of similarity that could be reached with simply a set of random spectra. In this sense, testing the robustness can be a prior step in a CompareVOA analysis, but it also serves a purpose in itself. If, for example, the database is insufficiently large or if with respect to the database the confidence level is not among the best, reporting the robustness still allows to assess whether the similarity is statistically meaningful.

In addition to formulating a solution for the robustness problem, we introduce a visual representation of the agreement between calculated and observed VCD spectra which is reminiscent of a model validation approach employed in quantitative structure−activity relationship (QSAR) studies.

Our method is illustrated for a set of quinolizidine alkaloids with rather limited conformational degrees of freedom. For this reason, they are considered an ideal first test case for our new method. The aim is to address more complicated molecular systems in the near future. The quinolizidine alkaloids that were measured are as follows: (−)-multiflorine (**1**), (−)-lupanine (**2**), (+)-sparteine (**3**), (−)-cytisine (**4**), (−)-3,5-dibromocytisine (**5**), (−)-*N*-acetylcytisine (**6**), (−)-*N*-propionylcytisine (**7**), (−)-*N*-benzylcytisine (**8**), (−)-*N*-benzoylcytisine (**9**), (−)-3-bromo-*N*-acetylcytisine (**10**), (−)-5-bromo-*N*-acetylcytisine (**11**), and (−)-3,5-dibromo-*N*-acetylcytisine (**12**). Their structures are shown in Scheme 1. Except for sparteine,[15] none of these molecules had its AC assigned previously using VCD.

Scheme 1. Quinolizidine Alkaloid Test Set of Molecules Used





	X	Y	R
4	H	H	H
5	Br	Br	H
6	H	H	$COCH_3$
7	H	H	COC_2H_5
8	H	H	$CH_2C_6H_5$
9	H	H	COC_6H_5
10	Br	H	$COCH_3$
11	H	Br	$COCH_3$
12	Br	Br	$COCH_3$

METHODOLOGY

Throughout, the Carbó-similarity index is used to evaluate the similarity between two spectra f and g over a wavenumber range $[\bar{\nu}_1, \bar{\nu}_2]$:[11,12,16,17]

$$R_{fg} = \frac{\int_{\bar{\nu}_1}^{\bar{\nu}_2} f(\bar{\nu}) g(\bar{\nu}) d\bar{\nu}}{\left(\int_{\bar{\nu}_1}^{\bar{\nu}_2} f^2(\bar{\nu}) d\bar{\nu}\right)^{1/2} \left(\int_{\bar{\nu}_1}^{\bar{\nu}_2} g^2(\bar{\nu}) d\bar{\nu}\right)^{1/2}} \quad (1)$$

Values lie in the interval [0,1] for nondifferential spectra like IR. In case of differential spectra like VCD, the interval becomes [−1,1]. A value of 1 indicates a perfect match between the spectra, while −1 represents an exact agreement when one of the spectra is inverted. The latter result implies that sample and

dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

calculated structure are enantiomers. Note that this means that contrary to our previous work, we no longer split the VCD spectra in either regions with positive and negative semidefinite regions. As compared to the previously used enantiomeric similarity index,[8] the evaluation is much simpler and thus meets the note by Simmen and co-workers[18] that the former measure was rather complicated. The Carbó-index has the advantage over the alternative Tanimoto-index[9] that it is independent of linear intensity scaling. Multiplying the intensity of the f spectrum with a constant factor, for example, does not affect the similarity. This is particularly useful when comparing calculated with experimental VCD spectra because the concentration of the sample is usually not precisely determined. In the present paper, both spectra f and g have been neighborhood weighted according to

$$f(\bar{\nu}) \leftarrow \int_{\bar{\nu}-L}^{\bar{\nu}+L} w(\bar{\nu}_0 - \bar{\nu}) f(\bar{\nu}_0) d\bar{\nu}_0 \tag{2}$$

and analogously for $g(\bar{\nu})$. L denotes the one-sided wavenumber neighborhood and was set to 15 cm^{-1}. Here, the triangular weight function was employed:[19]

$$w(\bar{\nu}_0 - \bar{\nu}) = \begin{cases} 1 - \dfrac{|\bar{\nu}_0 - \bar{\nu}|}{L} & \text{for } |\bar{\nu}_0 - \bar{\nu}| < L \\ 0 & \text{for } |\bar{\nu}_0 - \bar{\nu}| \geq L \end{cases} \tag{3}$$

Suppose now that f corresponds to a calculated spectrum. This is a line spectrum broadened using Lorentzian broadening (full width at half-maximum of 10 cm^{-1}) and frequency scaled with a global scale factor σ:[20]

$$f(\bar{\nu}) \leftarrow f(\sigma\bar{\nu}) \tag{4}$$

Because our method requires several similarity evaluations between different pairs of spectra we introduce the following notation. In general, the similarity between an arbitrary spectrum x and the experimental spectrum will be denoted by $R^{(IR)}_{x,exp}$ and $R^{(VCD)}_{x,exp}$ in case of IR and VCD spectra, respectively. The similarity between an arbitrary spectrum x and the computed spectrum, on the other hand, is denoted by $R^{(IR)}_{x,calc}$ and $R^{(VCD)}_{x,calc}$. Analogously, the direct similarity between the experimental and computed spectrum will be indicated by $R^{(IR)}_{calc,exp}$ and $R^{(VCD)}_{calc,exp}$.

Reported values for $R^{(IR)}_{calc,exp}$ and $R^{(VCD)}_{calc,exp}$ in the present paper are obtained using the global frequency scale factor σ (eq 4) that yields a maximum $R^{(IR)}_{calc,exp}$-value except for compounds **2** and **9**. For the latter, IR based scaling was found inconsistent with respect to the manual AC assignment, and therefore the scale factor was used that maximizes $R^{(VCD)}_{calc,exp}$. The search for the optimal σ occurs in a point-wise fashion in the interval [0.960, 1.020] with increment 0.001. By using eq 4, several physical effects and effects originating from artifacts due to the use of limited basis sets, limited treatment of electron correlation, solvation, etc. are taken into account.

One of the key questions is whether the computed degree of similarity ($R^{(VCD)}_{calc,exp}$) is significant (see Figure 1). In order to establish whether it is, we introduce an approach conceptually based on randomization tests.[21] In QSAR one usually has a vector of experimental observations for a set of molecules and a wide array of descriptors characterizing these molecules. QSAR methods then rely on e.g., regression techniques to derive models allowing to establish the formal link between the two to obtain values for the experimental observations for new molecules. Of main importance in QSAR is the validation of such models. The risk of overfitting is always present, and a model can only be regarded significant if it survives numerous tests such as leave many out tests, leverage tests, and also randomization tests.[21] The latter are the basis of our new method. In randomization tests, one randomly shuffles the elements in the vector of experimental data and attempts to construct a new model with the unshuffled molecular descriptors. A 2-dimensional plot is made with on the ordinate the R^2 value between the original vector of experimental data and the shuffled one and on the abscissa the R^2, and possibly Q^2 describing the regression between the experimental data and those predicted using the new models (see for example ref 22). In the case of good models, every significant shuffling of the experimental data vector (low R^2 on the abscissa) should lower strongly the R^2 value on the ordinate. Graphically, the real model should have R^2 = 1.00 on the abscissa and a high R^2 value on the ordinate, and all other models should lie like a cloud near the origin of the plot. Such validation of a QSAR model is thus highly informative on the model quality. We now follow the same line of thinking for vibrational optical activity (VOA) techniques. Taking VCD as an example, we want to construct a similar plot. There are, however, several reasons why one cannot simply proceed as in QSAR. First of all, there is no descriptor space and thus no model to be built. We have only experiment as a beacon. The computed spectrum is not simply a vector of discrete observations but a true spectrum containing a continuous variable. This issue is solved by rather looking at a spectrum in a deconvolutional way: a number of bands is derived from the spectra and for these bands the intensities are used as parameters. Random spectra can now be generated by simply putting the same number of, now random, bands (random in the sense of random absorption frequency and intensity) in the frequency range considered. The only user intervention is the decision of how many bands to consider in the frequency range. The agreement with the experimental spectrum and with the theoretical spectrum/model is then evaluated not by means of the R^2-value but using the similarity measure in eq 1. All required similarity evaluations are summarized in Scheme 1. A 2-D plot, reminiscent of that in QSAR, is then made. Using the experimental spectrum and original computed spectrum as beacons, on the abscissa the similarity ($R^{(VCD)}_{x,calc}$) is plotted between a new random computational spectrum and the original computed spectrum and on the ordinate the similarity ($R^{(VCD)}_{x,exp}$) between the new random spectrum and the experimental spectrum. Whereas in QSAR, one hopes that for the established model the original one is characterized by a high similarity between the experimental data and the predicted using the model and that all other models for shuffled experimental data have poor similarity, the 2-D plot is somewhat less straightforward for VOA spectra. Obviously, the ideal model has R^2 = 1 between theory and experiment which translates to $R^{(VCD)}_{calc,exp}$ = 1 when VCD

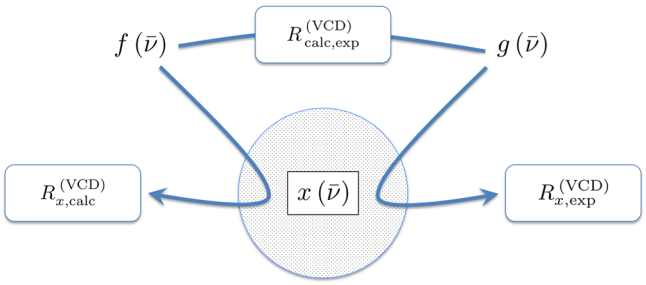


Figure 1. A scheme representing all similarity evaluations within the algorithm. The Carbó-similarity is evaluated between pairs of VCD spectra. f represents the calculated VCD spectrum, and g is the experimental one. x is the collection of computed random VCD spectra.

spectra are considered. This means that both spectra are the same. For every new random spectrum that has a specific $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$-value between the random spectrum and original computed one, the same value is found for $R_{x,\mathrm{calc}}^{(\mathrm{VCD})}$ between the random spectrum and the experimental one. So all points for all random spectra lie on a single straight line with correlation coefficient 1, slope 1, and intercept zero. From the moment that $R_{\mathrm{calc,exp}}^{(\mathrm{VCD})} \neq 1$, there is a chance that some random spectrum may result in higher similarity with experiment, but if the original similarity between theory and experiment is high enough, the chances of this should be fairly small or, alternatively, the raise in similarity should be modest. Not all points in the plot will lie on one line, rendering an inspection of the shape of the data-cloud informative on the quality of the AC assignment. In this respect, the calculation of a linear regression line is considered useful:

$$R_{x,\mathrm{exp}}^{(\mathrm{VCD})} = \alpha + \beta R_{x,\mathrm{calc}}^{(\mathrm{VCD})} \tag{5}$$

Ordinary least-squares has the disadvantage that either the residuals for $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ or $R_{x,\mathrm{calc}}^{(\mathrm{VCD})}$ can be minimized yielding two different (α,β)-solutions and hence two possible regression lines. An alternative regression method is to search for the α and β values for which the sum of squared perpendicular distances of the data points toward the line is minimal. This least-squares fitting using perpendicular offsets yields the following unique solution[23]

$$\alpha = \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle - \beta\langle R_{x,\mathrm{calc}}^{(\mathrm{VCD})}\rangle \tag{6}$$

$$\beta = \frac{(\kappa-\lambda) + \sqrt{(\kappa-\lambda)^2 + 4\mu^2}}{2\mu} \tag{7}$$

where

$$\lambda = \sum_x (R_{x,\mathrm{calc}}^{(\mathrm{VCD})} - \langle R_{x,\mathrm{calc}}^{(\mathrm{VCD})}\rangle)^2 \tag{8}$$

$$\kappa = \sum_x (R_{x,\mathrm{exp}}^{(\mathrm{VCD})} - \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle)^2 \tag{9}$$

$$\mu = \sum_x (R_{x,\mathrm{calc}}^{(\mathrm{VCD})} - \langle R_{x,\mathrm{calc}}^{(\mathrm{VCD})}\rangle)(R_{x,\mathrm{exp}}^{(\mathrm{VCD})} - \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle) \tag{10}$$

Using the standard deviation of the perpendicular distances two additional and parallel lines can be added to the figure to contain 99% of all data points. This is illustrated in Figures 3 and 6 of which the contents will be discussed in the Results and Discussion.

Based on the output of the regression analysis, several quantitative measures can be considered to describe the plot. Examples are the slope and the regression coefficient, but other measures such as the width between the lines containing e.g., 99% of all points can be derived as well. As in QSAR, however, the emphasis lies on the straightforward interpretation such plots have to offer, and hence the development of measures that quantify the agreement based on the appearance of these representations is less essential.

More important is how to obtain the statistical significance of the similarity between a pair of spectra. The robustness can be expressed as the probability that the observed $|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|$-similarity cannot be improved by a random VCD spectrum. The lower this probability of improvement, the more robust is the $R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}$-value. The similarity of a random VCD spectrum x with respect to the experimental one thus represents a stochastic variable $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ from which significance levels can be calculated. Assuming that $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ is normally distributed with mean $\langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle$ and standard deviation $s_{x,\mathrm{exp}}^{(\mathrm{VCD})}$

$$\phi(R_{x,\mathrm{exp}}^{(\mathrm{VCD})}) = \frac{1}{(\sqrt{2\pi})s_{x,\mathrm{exp}}^{(\mathrm{VCD})}} \times \exp\left[-\frac{1}{2}\left(\frac{R_{x,\mathrm{exp}}^{(\mathrm{VCD})} - \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle}{s_{x,\mathrm{exp}}^{(\mathrm{VCD})}}\right)^2\right] \tag{11}$$

The robustness of $|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|$ can be expressed by the probability P that a similarity of at least $|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|$ cannot be obtained at random:

$$\begin{aligned} P &= P[R_{x,\mathrm{exp}}^{(\mathrm{VCD})} < |R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|] \\ &= 1 - [\int_{-\infty}^{-|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|} \phi(R_{x,\mathrm{exp}}^{(\mathrm{VCD})})dR_{x,\mathrm{exp}}^{(\mathrm{VCD})} \\ &\quad + \int_{|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|}^{+\infty} \phi(R_{x,\mathrm{exp}}^{(\mathrm{VCD})})dR_{x,\mathrm{exp}}^{(\mathrm{VCD})}] \end{aligned} \tag{12}$$

This represents the unit surface of the bell shaped distribution from which the tails with absolute similarities higher than $|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|$ are subtracted. The integrals can be expressed by error functions, allowing for rewriting the above equation as

$$P = \frac{1}{2}\left[\mathrm{erf}\left(\frac{|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}| - \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle}{\sqrt{2}\,s_{x,\mathrm{exp}}^{(\mathrm{VCD})}}\right) - \mathrm{erf}\left(\frac{-|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}| - \langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})}\rangle}{\sqrt{2}\,s_{x,\mathrm{exp}}^{(\mathrm{VCD})}}\right)\right] \tag{13}$$

Within the current algorithm, random VCD spectra are generated by randomly positioning K Lorentzian broadened absorption bands (full width at half-maximum of 10 cm^{-1}) with a random sign and intensity within the wavenumber region of interest. K is the number of bands in the spectrum. As the Carbó-index (eq 1) is intensity-scale independent, only relative intensities are important, and the magnitude of the rotational strengths is simply chosen within the interval [0, 1]. The only parameter not chosen at random is the number of bands K which is extracted from the computed spectra after scaling the absorption frequencies.

The higher the number of VCD spectra generated, the more accurately the $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$-distribution can be determined. The normality of the distribution can be verified by constructing a histogram from the obtained $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$-data. Experience has shown that the tails of the distribution are well reproduced for a total of 25000 random spectra, and this number was hence used throughout.

EXPERIMENTAL AND COMPUTATIONAL DETAILS

The isolation, derivatization, and characterization of **1**−**12** have been described by Przybył and co-workers.[24] IR and VCD were recorded on a BioTools dual-PEM ChiralIR-2X spectrometer. The PEMs were optimized for 1400 cm^{-1}, and a resolution of 4 cm^{-1} was used throughout. The path length of the cell equipped with BaF_2 windows is 100 μm. Compounds **1**, **4**, **6**, **7**, **8**, **10**, **11**, and **12** were investigated in solutions of ca. 7.2, 5.5, 9.9, 6.5, 8.3, 8.4, 7.9, and 7.5 mg in 0.1 mL of $CDCl_3$, respectively. The samples of **2**, **3**, and **9** were oils, and their concentration could not be accurately determined. They were diluted until the maximal IR absorbance was below 0.4 units in the spectral region of

dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

Table 1. Summary of Numerical Results and Comparison with CompareVOA[a]

	manual AC	no. conf.	σ	$R^{(IR)}_{calc,exp}$(%)	$R^{(VCD)}_{calc,exp}$ (%)	K	$s_{x,exp}$ (%)	P (%)	CL (%)
1	yes	9	0.972	91.9	83.2	37	20.6	99.99	99
2	yes	13	0.970	81.2	62.3	43	20.8	99.93	99
3	yes	13	0.968	88.1	64.3	46	19.1	99.98	99
4	yes	4	0.964	86.7	56.5	19	19.8	99.62	96
5	no	4	0.998	47.7	4.5	16	20.2	13.79	65
6	yes	4	0.978	91.9	67.2	22	21.3	99.39	99
7	yes	13	0.979	91.0	60.3	26	17.9	99.30	96
8	yes	11	0.974	88.1	70.4	29	18.2	99.94	99
9	yes	8	0.975	77.0	78.6	26	21.3	99.96	99
10	yes	4	0.980	93.1	76.1	21	22.3	99.83	99
11	yes	4	0.980	92.1	53.0	21	19.7	97.28	94
12	yes	4	0.980	91.3	78.4	20	21.5	99.95	99

[a]The CL-value is generated by CompareVOA; all other values are obtained with the present routine. The used wavenumber range was 1500−1000 cm^{-1}, except for **5** where 1600−1100 cm^{-1} was used.

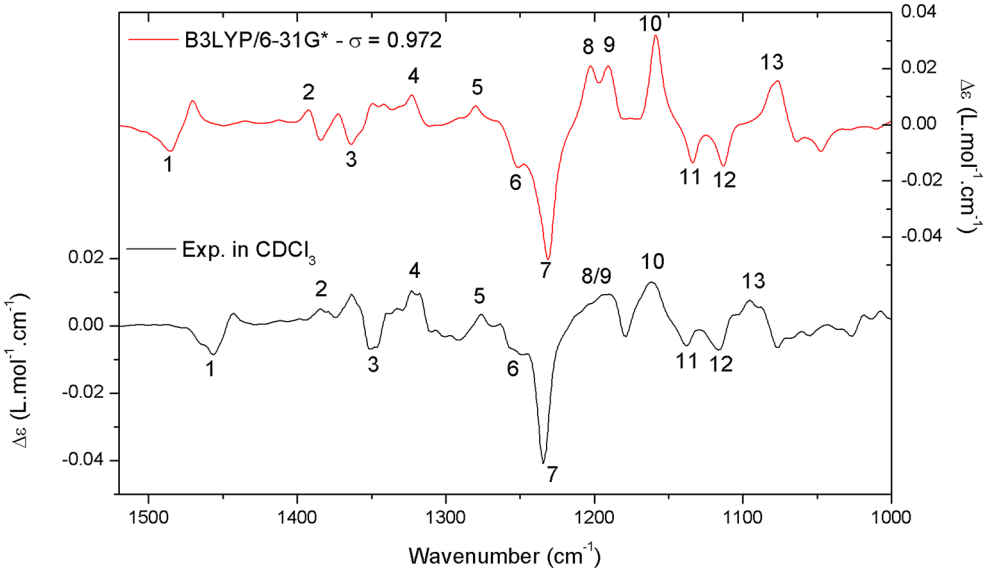


Figure 2. VCD spectra (−)-multiflorine: experimental (below) and calculated (7S,7aS,14S,14aR)-**1** (upper). Bands are numbered in agreement with IR.

interest. Due to the limited solubility in $CDCl_3$ of compound **5**, it was investigated in a solution of 0.7 mg in 0.1 mL of DMSO-d_6. Baseline corrections were accomplished using the spectra of the solvent, except for **2** and **3** for which both enantiomers were measured and a virtual racemate spectrum was used as blanco.

The conformational analysis was performed using the Spartan '08[25] and Conflex[26] worksuites of programs and with the MMFF94[27] and MMFF94S[28] molecular mechanics force fields. The resulting geometries were further optimized, and the Boltzmann weighted IR and VCD spectra were computed at the B3LYP/6-31G* level of theory using Gaussian09.[29]

RESULTS AND DISCUSSION

A manual peak-to-peak AC assignment is performed for each molecule by matching bands between theory and experiment. Since VCD absorption frequencies and IR frequencies are equal, the IR spectra are used as a guideline, and VCD is used as a secondary source of information. Band matching is then performed until consistency is obtained, taking into account that no large shifts in frequencies should be used between theory and experiment. It is the manual assignment of the spectra that is used to establish the AC of the molecule. In the next step, the numerical similarity measures introduced above are calculated.

Proceeding to the similarity evaluation, we focus on two samples (**1** and **5**) to illustrate our method. For **1** the AC can be trivially obtained manually, while for **5** the AC could not be determined. Numerical results for these and the other compounds can be found in Table 1. The AC determination of the other molecules can be found in the Supporting Information.

The experimental and calculated VCD data for **1** are given in Figure 2. The VCD spectrum obtained at the B3LYP/6-31G*-level matches the experimental spectrum quite well and thus allows for establishing the AC of (−)-multiflorine as (7S,7aS,14S,14aR)-**1**. However, the optimal scale factor (0.972) does not allow for aligning all bands perfectly. In particular calculated bands 1, 3, and 13 are shifted with respect to the experiment. The similarity between the VCD spectra is 83.2% (see Table 1). Although this is a quite high value, the magnitude of the value does not by itself establish the robustness of the similarity value. We therefore generated 25000 random spectra by placing 37 (see Table 1) bands with random signs, intensities, and frequencies in the 1500−1000 cm^{-1} wavenumber region. The earlier described randomization plot is generated (Figure 3) showing indeed a quite tight distribution of the spectra. The actual computed similarity where x corresponds to the computed spectrum, denoted by the red dot in Figure 3, is clearly higher

dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

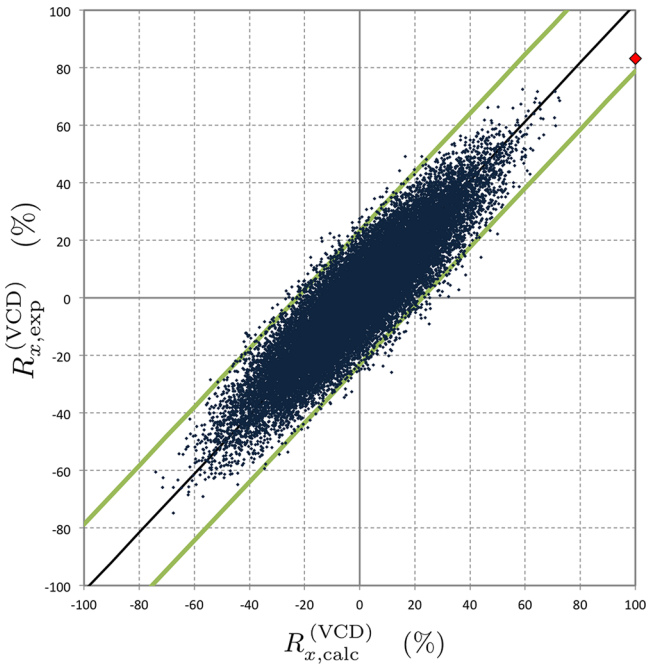


Figure 3. $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ versus $R_{x,\mathrm{calc}}^{(\mathrm{VCD})}$ scatter diagram for (−)-multiflorine. Blue data points represent random spectra. The red data point represents the calculated spectrum which has a 100% similarity with respect to itself and a similarity of $R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}$ with respect to the experiment. The black line is obtained by orthogonal regression. The area between the green lines contains 99% of the data points.

than that of any other point illustrating a robust match between the experimental and computed spectra. The *P*-value (99.99%) shows that, regardless of the residual bad alignment of a few absorption bands, the $|R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}|$-similarity with respect to the experimental VCD spectrum of **1** is practically impossible to obtain at random. The histogram of $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ obtained for **1** (Figure 4) exhibits a clear normal distribution and justifies the use of eqs 11−13.

The experimental and calculated VCD spectra for **5** are depicted in Figure 5. Due to poor solubility of **5** in $CDCl_3$, the VCD spectrum was recorded in DMSO-d_6. This quite often leads

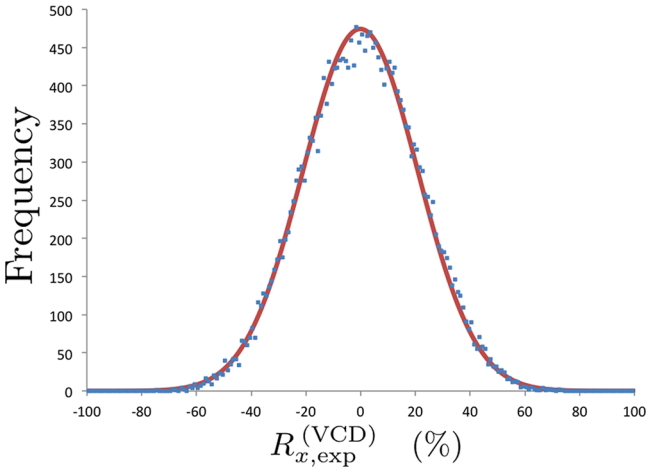


Figure 4. The histogram obtained for $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ (blue points) in the case of (−)-multiflorine. The range is divided in bins of 1%. The normal distribution obtained from the mean and standard deviation is represented in red.

to significant solvent−solute interactions making assignments regularly more difficult. Figure 5 shows the large discrepancy between the theoretical and experimental spectra. Since the noise levels of the experimental spectrum are low, the quality of the VCD measurement can be ruled out as source for the disagreement. Since for the nonbromated analogue (**4**) and the *N*-acetyl substituted derivative (**12**), the B3LYP/6-31G* spectra matched the experimental ones quite well (see Table 1), it is expected that DMSO-d_6 induces specific interactions that are not well described in the theoretical spectra due to either shortcomings in the applied level of theory or the lack in modeling of explicit interactions. On the other hand, the sample concentration is very low, and therefore small enantiomeric excess and/or other contaminations are also reasonable causes for the large disagreement. The spectrum is nonetheless useful to illustrate how our method performs in these circumstances. Figure 5 shows that some absorption bands are entirely absent in the experiment with respect to the calculation. As a consequence, there is insufficient reference signal to allow a coherent band numbering in the spectra and, hence, to conduct a manual AC assignment. From Table 1, we see that $R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}$ is extremely low (4.5%). The significance of this value P = 13.79% indicates, as expected, that this value is entirely nonrobust. This value was obtained using a set of 25000 spectra with 16 bands in the 1600−1100 cm^{-1} wavelength region. For each spectrum obtained by randomly assigning a frequency and a differential absorption to each band, the similarity with respect to the original computed spectrum and the experimental spectrum is computed, and the entire set of these similarities is plotted - see Figure 6. The graph exhibits a rather circular distribution of the data-points with many randomly generated spectra having a higher similarity with respect to experiment. Any assignment based on the computed spectrum is therefore considered highly nonrobust.

Note, that in Figure 7 a small deviation from normality is observed. This is caused by the fact that there appears to be some homogeneous negative absorption spread over the range 1200−1600 cm^{-1} in the experimental spectrum. As a result, there is always some overlap in this region and hence a certain degree of similarity. The latter broadens the distribution of $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ and renders the *P*-value less accurate. Yet, since $R_{\mathrm{calc,exp}}^{(\mathrm{VCD})}$ is very low, the error on *P*, which is caused by the deviation from normality, is considered negligible. Between the straightforward assignment for **1** and the problematic one for **5**, there remains a wide range where intermediate similarities would be best checked for their robustness. The similarity is rather not taken as a single number but attached to it should be a randomization plot. If that testifies that the similarity is insufficiently robust, any further conclusions, or for that matter use of e.g., CompareVOA, become problematic.

For all compounds **1**−**12**, the mean of the normal distributions $\langle R_{x,\mathrm{exp}}^{(\mathrm{VCD})} \rangle$ deviates no more than 0.2% from zero which makes the distribution well centered at the origin. Table 1 includes the standard deviation $s_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ of these distributions. Note that depending on the compound, $s_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ varies over a range of nearly 5%. If $3s_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ is taken as the approximated 99% significance level, then the same robustness is obtained for similarities varying up to 15%. Larger variation is expected depending on the spectral range or other molecules being considered. The latter illustrates that simple rules to obtain significance levels are not transferable and can be misleading. It is therefore important to address the robustness of the similarity for each spectrum individually.

As was stressed already above, it is important to note that an algorithm like CompareVOA serves a different purpose than the

dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

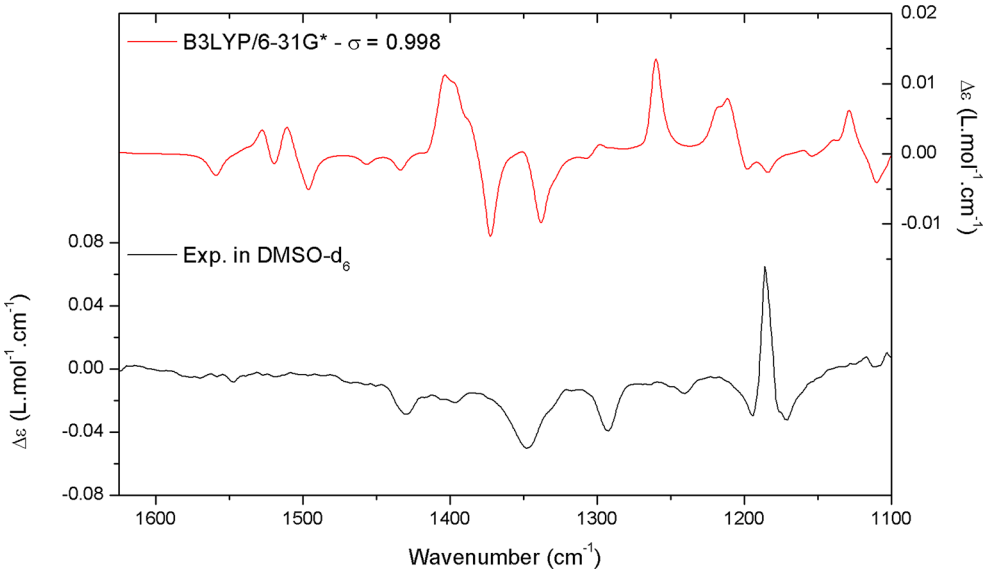


Figure 5. VCD spectra 3,5-dibromocytisine: experimental (below) and calculated (1R,5S)-**5** (upper).

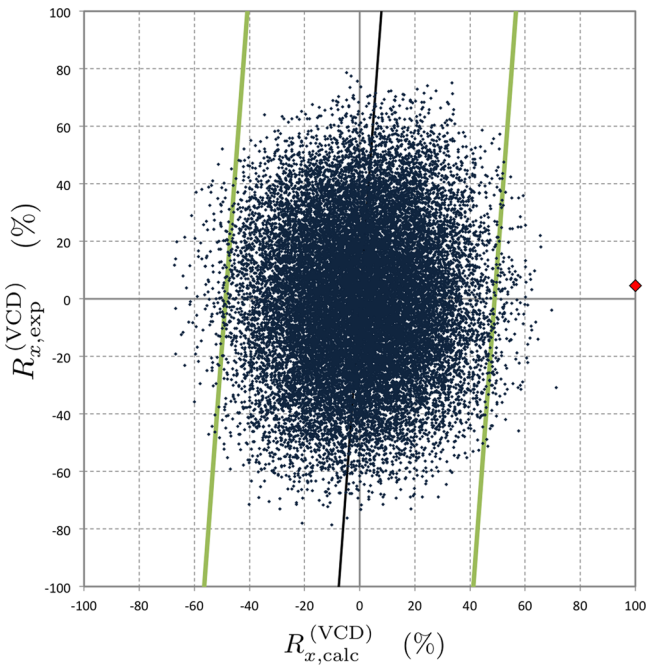


Figure 6. $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ versus $R_{x,\mathrm{calc}}^{(\mathrm{VCD})}$ scatter diagram for 3,5-dibromocytisine. The same formatting was used as in Figure 3.

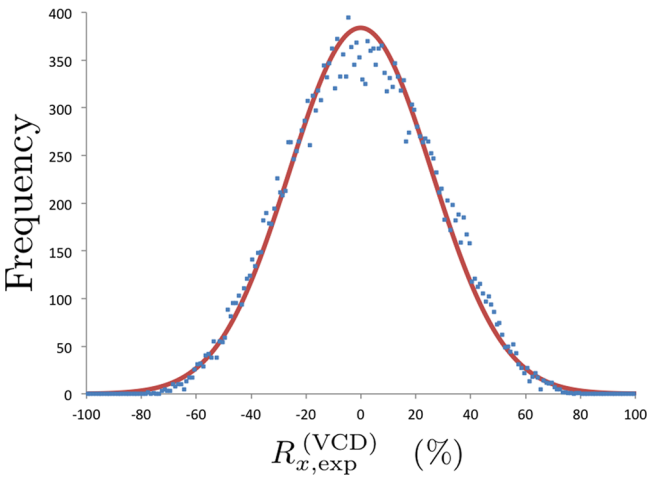


Figure 7. The histogram obtained for $R_{x,\mathrm{exp}}^{(\mathrm{VCD})}$ (blue points) in case of 3,5-dibromocytisine. The range is divided in bins of 1%. The normal distribution obtained from the mean and standard deviation is represented in red.

robustness of the similarity as introduced here. CompareVOA allows users of VCD to establish how good a VCD based assignment is within a database of successful and confirmed correct assignments. To that end it uses similarity measures albeit for both the supposedly correct assignment and the wrong one. The present algorithm establishes whether the similarity is robust. There does, however, appear a qualitative link. If the original computed spectrum clearly has a higher similarity with the experimental spectrum than any of the random spectra, it usually has a higher confidence level. Table 1 shows that the CL-values obtained from CompareVOA and the presently calculated statistical significance *P* are indeed in qualitative agreement. The manual AC assignments presented here show that both methods are consistent. A similar qualitative agreement of CompareVOA and statistical methods was reported by Shen and co-workers[10] as well. Note in Table 1, however, that there are some differences between *P* and CL worth mentioning. For compounds **4** and **7** the CL-value equals 96%, whereas the robustness *P* exceeds 99%. From the latter one concludes that less than 1 on every 100 random spectra would have a higher similarity with respect to the experiment. The CL-value, however, indicates that according to the CompareVOA database, insufficient correct assignments have been performed for the obtained similarity values. Due to the database size dependency of CompareVOA, it is recommended to determine the robustness as well.

CONCLUSION AND OUTLOOK

A new computational algorithm has been proposed to determine the robustness of vibrational circular dichroism (VCD) based absolute configuration (AC) assignments for chiral molecules. Inspired by randomization tests, which are employed for model validation in quantitative structure−activity relationship studies, a 2D-plot is constructed that illustrates how well the similarity between a computed and an experimental VCD spectrum can be

dx.doi.org/10.1021/ct400843e | J. Chem. Theory Comput. 2013, 9, 5504−5512

distinguished from the similarity obtained with random spectra. Based on sound statistics, a measure for the robustness of a similarity value has been proposed. The novel approach was tested on a set of quinolizidine alkaloids and has proven to be supportive with respect to the identification of both correct and incorrect AC assignments. The standard deviation of the normal distributions for these compounds differs across the set of molecules and shows that the minimal similarity required to reach a certain significance level is not transferable among spectra. In order to avoid misinterpretation of similarity values it is thus recommended to address their robustness for each experimental VCD spectrum individually as in the present study.

Our technique has several advantages. It is readily applicable to both rigid and flexible molecules and over any wavenumber range as long as the statistics can be expressed using normal distributions. The method is in principle extendable to other chiroptical techniques such as ECD, ROA, or even dissymmetry factor spectra. The algorithm does not require an experimental database unlike CompareVOA yet serves a different purpose. Whereas CompareVOA verifies whether or not correct AC assignments have been performed for the same similarity values on other molecular systems within the database, the present method addresses the robustness of the similarity value. As a result, each approach can be used separately, but they also complement one another. When a similarity is found to be nonrobust, a further CompareVOA analysis is deemed unreliable. When, on the other hand, a similarity is highly robust, the absolute similarity is usually quite high and should therefore lead to a high confidence level in CompareVOA.

More challenging molecules with higher conformational flexibility will be addressed in the near future in order to further validate and/or perhaps refine the technique.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Computed and measured IR and VCD spectra with band numbering according to the manual AC assignment for compounds **1**−**12** and the associated randomization plots and $R_{x,\mathrm{exp}}$-histograms. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: patrick.bultinck@ugent.be.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

P.B. and W.H. acknowledge continuous support from the Fund for Scientific Research FWO-Vlaanderen, the Bijzonder Onderzoeksfonds (BOF), and the Industrieel Onderzoeksfonds (IOF). The computational resources (STEVIN Supercomputer Infrastructure) and services used in this work were kindly provided by Ghent University, the Flemish Supercomputer Center (VSC), the Hercules Foundation, and the Flemish Government department EWI.

■ REFERENCES

(1) (a) Uray, G.; Verdino, P.; Belaj, F.; Kappe, C.; Fabian, W. *J. Org. Chem.* **2001**, *66*, 6685−6694. (b) Pescitelli, G.; Gabriel, S.; Wang, Y.; Fleischhauer, J.; Woody, R.; Berova, N. *J. Am. Chem. Soc.* **2003**, *125*, 7613−7628.

(2) Stephens, P. J.; Devlin, F. J.; Pan, J.-J. *Chirality* **2008**, *20*, 643−663.

(3) He, Y.; Wang, B.; Dukor, R. K.; Nafie, L. A. *Appl. Spectrosc.* **2011**, *65*, 699−723.

(4) Barron, L.; Buckingham, A. *Annu. Rev. Phys. Chem.* **1975**, *26*, 381−396.

(5) Covington, C. L.; Polavarapu, P. L. *J. Phys. Chem. A* **2013**, *117*, 3377−3386.

(6) Nafie, L. A. *Vibrational Optical Activity: Principles and Applications*, 1st ed.; John Wiley & Sons, Ltd.: Chichester, UK, 2011; pp 1−34.

(7) Kuppens, T.; Bultinck, P.; Langenaeker, W. *Drug Discovery Today: Technol.* **2004**, *1*, 269−275.

(8) Debie, E.; De Gussem, E.; Dukor, R. K.; Herrebout, W.; Nafie, L. A.; Bultinck, P. *ChemPhysChem* **2011**, *12*, 1542−1549.

(9) Shen, J.; Zhu, C.; Reiling, S.; Vaz, R. *Spectrochim. Acta, Part A* **2010**, *76*, 418−422.

(10) Shen, J.; Li, Y.; Vaz, R.; Izumi, H. *J. Chem. Theory Comput.* **2012**, *8*, 2762−2768.

(11) Kuppens, T.; Langenaeker, W.; Tollenaere, J.; Bultinck, P. *J. Phys. Chem. A* **2003**, *107*, 542−553.

(12) Kuppens, T.; Vandyck, K.; Van der Eycken, J.; Herrebout, W.; van der Veken, B.; Bultinck, P. *J. Org. Chem.* **2005**, *70*, 9103−9114.

(13) Kuppens, T. Development of methodology to assign absolute configurations using vibrational circular dichroism. Ph.D. Thesis, Department of Inorganic and Physical Chemistry, Ghent University, 2006.

(14) Kuppens, T.; Vandyck, K.; van der Eycken, J.; Herrebout, W.; van der Veken, B.; Bultinck, P. *Spectrochim. Acta, Part A* **2007**, *67*, 402−411.

(15) (a) Bour, P.; McCann, J.; Wieser, H. *J. Phys. Chem. A* **1997**, *101*, 9783−9790. (b) He, Y.; Cao, X.; Nafie, L.; Freedman, T. *J. Am. Chem. Soc.* **2001**, *123*, 11320−11321.

(16) Carbó, R.; Leyda, L.; Arnau, M. *Int. J. Quantum Chem.* **1980**, *17*, 1185−1189.

(17) Bultinck, P.; Girones, X.; Carbo-Dorca, R. *Rev. Comput. Chem.* **2005**, *21*, 127−207.

(18) Simmen, B.; Weymuth, T.; Reiher, M. *J. Phys. Chem. A* **2012**, *116*, 5410−5419.

(19) De Gelder, R.; Wehrens, R.; Hageman, J. *J. Comput. Chem.* **2001**, *22*, 273−289.

(20) (a) Scott, A.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502−16513. (b) Wong, M. *Chem. Phys. Lett.* **1996**, *256*, 391−399.

(21) (a) Tropsha, A.; Gramatica, P.; Gombar, V. *QSAR Comb. Sci.* **2003**, *22*, 69−77. (b) Gramatica, P. *QSAR Comb. Sci.* **2007**, *26*, 694−701. (c) Van Damme, S.; Bultinck, P. *J. Comput. Chem.* **2007**, *28*, 1924−1928.

(22) Dolezal, R.; Van Damme, S.; Bultinck, P.; Waisser, K. *Eur. J. Med. Chem.* **2009**, *44*, 869−876.

(23) (a) Weisstein, E. W. MathWorld - A Wolfram Web Resource: Least Squares Fitting - Perpendicular Offsets. http://mathworld.wolfram.com/LeastSquaresFittingPerpendicularOffsets.html (accessed August 22, 2013). (b) Adcock, R. J. *Analyst (Cambridge, U. K.)* **1878**, *5*, 53−54.

(24) (a) Przybył, A. K.; Kubicki, M. *J. Mol. Struct.* **2011**, *985*, 157−166. (b) Przybył, A. K.; Nowakowska, Z. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1193−1197. (c) Przybył, A. K.; Kubicki, M. *Tetrahedron* **2011**, *67*, 7787−7793. (d) Przybył, A. K.; Prukała, W.; Kikut-Ligaj, D. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1409−1413. (e) Galasso, V.; Pryzbył, A.; Christov, V.; Kovač, B.; Asaro, F.; Zangrando, E. *Chem. Phys.* **2006**, *325*, 365−377. (f) Galasso, V.; Asaro, F.; Berti, F.; Przybył, A.; Włodarczak, J.; Wysocka, W.; Habuš, I.; Kovač, B. *Chem. Phys.* **2005**, *314*, 25−36.

(25) Spartan '08; Wavefunction, Inc.: Irvine, CA.

(26) (a) Goto, H.; Osawa, E. *J. Am. Chem. Soc.* **1989**, *111*, 8950−8951. (b) Goto, H.; Osawa, E. *J. Chem. Soc., Perkin Trans. 2* **1993**, 187−198. (c) Goto, H.; Ohta, K.; Kamakura, T.; Obata, S.; Nakayama, N.; Matsumoto, T.; Osawa, E. *CONFLEX*; Conex Corp.: Tokyo-Yokohama, Japan, 2004.

(27) Halgren, T. *J. Comput. Chem.* **1996**, *17*, 490−519.

(28) Halgren, T. *J. Comput. Chem.* **1999**, *20*, 720−729.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.;

Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.2*; Gaussian Inc.: Wallingford, CT, 2009.