# A combined Raman optical activity and vibrational circular dichroism study on artemisinin-type products

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

# A combined Raman optical activity and vibrational circular dichroism study on artemisinin-type products†

Jonathan Bogaerts[a], Filip Desmet[a], Roy Aerts[a], Patrick Bultinck[b], Wouter Herrebout[a] and Christian Johannessen*[,a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Artemisinin and two of its derivatives, dihydroartemisinin and artesunate, which are front line drugs against malaria, were investigated using Raman optical activity (ROA) and vibrational circular dichroism (VCD) experiments, both supported by density functional theory (DFT) level calculations. The experimental techniques combined with DFT calculations could show that dihydroartemisinin was present as an epimeric mixture in solution. In addition, an approximation of the epimeric ratio could be extracted which was in agreement with the ratio obtained by $^{1}$H-NMR spectroscopy. The current study also demonstrates that both ROA and VCD are able to assign the correct absolute configuration (AC) of artemisinin and artesunate out of all their possible diastereomers without any explicit knowledge on their correct stereochemistry and accentuates the synergetic effect between ROA and VCD in AC determination.

## Introduction

Raman optical activity (ROA) is measured as the difference in the right-handed ($I_R$) and left-handed ($I_L$) circularly polarized components of Raman light scattered by chiral molecules. Therefore, ROA has an intrinsic potential in the differentiation of enantiomers and the determination of absolute configuration (AC) of chiral molecules. However, ROA historically finds most of its applications in the solution structure elucidation of biomolecules.[1–6] One of its more recent focuses, also explored by our group[7,8], revolves around the resonance enhancement effect to increase the small ROA scattering phenomena ($\sim 10^{-4}$ smaller than the corresponding Raman signal).[9,10] As a consequence, and although ROA has shown its full potential in absolute configuration (AC) assignment in some specific cases[11–14], this possible application of ROA spectroscopy has only scarcely been addressed. In contrast, a plethora of examples can be found in literature in which vibrational circular dichroism (VCD), measured as the differential absorbance of left- and right circularly polarised IR light, is applied to unambiguously assign the AC of a chiral molecule. As it is impossible to cite all VCD studies on AC determination, the authors refer to a review article on natural products by Batista *et al.*[15] and on small organic molecules by Merten *et al.*[16]. The lack of a similar, in depth, review article concerning ROA spectroscopy applied in AC determination is the best illustration of how ROA lags behind VCD in this field of research.

Nevertheless, the workflow to assign the AC based on ROA spectroscopy is identical to the one using VCD spectroscopy: an experimental spectrum needs to be recorded after which a comparison with a predicted spectrum is made. Especially when a solvent such as chloroform ($CDCl_3$) is used, the predictions are mostly straightforward and consist of a thorough conformational analysis together with the calculation of a Boltzmann weighted ROA/VCD spectrum without the need of including explicit solute-solvent interaction which can alter the experimental observed spectra quite drastically.[17,18] In order to bridge the gap between ROA and VCD spectroscopy in AC determination, the current work investigates the possibility of ROA spectroscopy as an alternative for, or as complementary technique to, the more established VCD approach to unambiguously assign the AC of a chiral molecule. To do this, an in depth, side-by-side analysis of ROA and VCD on AC determination is performed. For this case study, we have opted for the sesquiterpene lactone artemisinin and two of its derivatives: dihydroartemisinin and artesunate. The molecular structures of the three compounds are shown in figure 1.

Artemisinin, isolated from dried leaves from the plant *artemisia annua*, is currently the most valuable weapon in the fight against malaria.[19] Furthermore, the easily accessible derivates dihydroartemisinin and artesunate exert a potent antimalarial activity against drug-resistant malaria strains. Dihydroartemisinin and artesunate exhibit the same core structure as artemisinin but are

[a] *Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium. E-mail: christian.johannessen@uantwerpen.be.*
[b] *Department of Chemistry, Ghent University, Krijgslaan 281 (S3), B-9000 Ghent, Belgium.*

† Electronic Supplementary Information (ESI) available: Description of the home-built ROA instrument, additional figures and tables, cartesian coordinates of artesunate. See DOI: 00.0000/00000000.

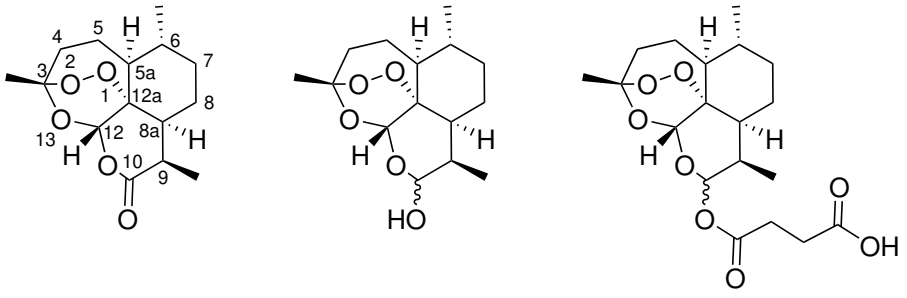


**Fig. 1** Structure of artemisinin, dihydroartemisinin and artesunate with atom numbering. Artemisinin contains seven chiral centres at atom numbers 3, 5a, 6, 8a, 9, 12 and 12a. Dihydroartemisinin and artesunate exhibit an extra chiral center at atom number 10.

differently functionalised at C-10 (figure 1) and therefore display an extra chiral center, compared to artemisinin, which needs to be assigned. Dihydroartemisinin has the lactone of artemisinin replaced by a hemiacetal moiety and artesunate can be formed by a reaction of dihydroartemisinin with succinic anhydride. As dihydroartemisinin and artesunate have a higher bioavailability compared to the parent compound, they have become front-line drugs in the fight against malaria.[20] That these compounds are of high interest is best illustrated by the Noble prize in Physiology or medicine given to Prof. Youyou Tu in 2015, who together with her co-workers discovered and elucidated the structure of artemisinin,[21] and by the efforts made to find (semi)-synthetic pathways to render a scalable approach for the production of artemisinin.[22,23]

Previous work on dihydroartemisinin using VCD spectroscopy by Zhang *et al.*[24] focussed on solvent effects on the VCD spectrum when using DMSO$-d_6$ as solvent. They concluded that the solvent should be taken into account explicitly to achieve a good agreement between experiment and calculations. In the current work, an extensive ROA and VCD study on the three artemisinin-type compounds is presented using $CDCl_3$ as solvent, avoiding the need of this explicit solvation model. We show that a general good agreement can be found between experimental and calculated spectra when the known AC of artemisinin is used, highlighting the potential of ROA spectroscopy in this field of research. Furthermore, we illustrate that ROA and VCD are able to distinguish between the two possible epimers at C-10 of artesunate and that ROA and VCD can be used to give an indication for the epimeric composition associated with the experimental sample of dihydroartemisinin, as was verified by $^1$H-NMR spectroscopy. Throughout the main text, comparison between the Raman/ROA and IR/VCD technique is discussed. Finally, we show that ROA and VCD are able to assign the correct AC of artemisinin and artesunate out of all possible diastereomers. Moreover, when the results of ROA and VCD are combined, the assignment is even more convincing showing the complementarity and the power of combing both chiroptical techniques for AC determination. Therefore this study argues for the further exploration of ROA in the field of AC determination and for combining it with the established VCD technique.

## Experimental methods

**Computational details.** To calculate ROA and VCD (VOA) spectra, all conformers present in solution must be taken into account. Therefore, a conformational search on molecular mechanics level of theory was performed on all the three artemisinin-type compounds. Two software packages Conflex[25] (MMFF94S[26]) and Spartan08[27] (MMFF94[28] and SYBYL[29]) were used for this purpose. The employed force fields are placed between brackets. After elimination of redundant conformations, *ab initio* optimisation at the B3PW91/6-31++G(d,p) level of theory using Gaussian09[30] was carried out. The unique conformers were all checked to be in a local minimum (e.g. exhibiting no imaginary frequencies) and the once having a Boltzmann weight, calculated using $\Delta H^\circ$, above 1% were selected for subsequent VOA calculations at the same level of theory. A global scaling factor on the harmonic frequencies was applied to compensate for *inter alia* the harmonic approximation. Raman and ROA intensities were calculated with an incident light wavelength of 532 nm to match the experimental set-up. Line broadening of Raman and ROA spectra were modelled using a Lorentzian bandshape with a full width at half maximum (FWHM) of 15 $cm^{-1}$ after a Boltzmann factor to correct for temperature (298 K) for each normal mode was applied.[31] The IR and VCD spectra were broadened using a Lorentzian function with a FWHM of 10 $cm^{-1}$. In all cases, the solvent was taken into account by using the integral equation formalism model (IEFPCM) as implemented in Gaussian09 and the dielectric constant for chloroform ($\varepsilon = 4.71$).

**Materials.** Artemisinin, dihydroartemisinin and artesunate were purchased from TCI Europe and used without any further purification. $CDCl_3$ was purchased from Sigma-Aldrich.

**Raman and ROA spectroscopy.** Raman and SCP-ROA spectra were measured on a home-built SCP-ROA spectrometer running at a resolution of $\sim 7$ $cm^{-1}$ and using a laser excitation wavelength of 532 nm. The optical layout and operation of this home-built instrument is based on the high-throughput SCP-ROA spectrometer designed by Hug and Hangartner[32] and incorporates the virtual enantiomers optical offset reduction scheme[33]. A full description of the instrument is given in the supplementary information (see section **??** in ESI†). The Raman intensities are displayed as the sum ($I_R$ + $I_L$) and the ROA intensities as the difference ($I_R$ - $I_L$) in circular intensities with $I_R$ and $I_L$ denoting

for the inelastic Raman scattering with right and left circular polarization, respectively. The laser power was set in the range of 315 - 350 mW at the sample depending on the sample, and total acquisition times were between 3 and 28 h. at ambient temperature. Solvent spectra were subtracted from the Raman spectra after which the baseline correction procedure by Boelens *et al.*[34] was applied. Cosmic ray spikes were removed from the Raman and ROA data by means of a median filter. Subsequently, the ROA spectra were smoothed using a third-order nine-points Savitzky-Golay filter. A concentration of 100 mg.mL$^{-1}$ was used for all three artemisinin-type products.

**IR and VCD spectroscopy.** IR and VCD measurements were carried out on a dual PEM ChiralIR-2X spectrometer (Biotools, Inc.) at room temperature running at a resolution of 4 cm$^{-1}$. A cell with a pathlength of 100 $\mu$m equipped with $BaF_2$ windows and a concentration of 40 mg.mL$^{-1}$ was used for the three samples. A total of 12000 scans (4 $\times$ 3000 scans $\approx$ 4 $\times$ 1h) for each sample was recorded. The final IR and VCD spectra were obtained by subtracting the solvent spectra measured under identical conditions.

## Results and discussion

**Experimental Raman/ROA and IR/VCD spectra of the three artemisinin-type compounds.**

The experimental Raman and ROA spectra of artemisinin, dihydroartemisinin and artesunate in the 950 - 1800 cm$^{-1}$ spectral range are depicted in figure 2. Although some differences in the Raman spectra can be observed, overall the Raman spectra of the three compounds are very similar, apart from the carbonyl stretch present in artemisinin and artesunate observed above 1600 cm$^{-1}$. The ROA spectra of the three compounds are almost identical although some differences are evident, especially below 1200 cm$^{-1}$. The high similarity between the ROA spectra is quite surprising, since one would expect that the addition of a chiral center at the C-10 position in the case of dihydroartemisinin and artesunate would yield larger changes in the chiral sensitive ROA technique.[35]

In figure 3 the experimental IR and VCD spectra of artemisinin, dihydroartemisinin and artesunate in the fingerprint region (950 - 1800 cm$^{-1}$) are reported. Especially in the region below 1200 cm$^{-1}$ differences in the IR spectra are observed, both in band intensity and shape. Throughout the whole shown spectral range of the VCD spectra, differences in band intensities and shape are observed with the most distinct patterns between the three compounds detected below 1200 cm$^{-1}$. This indicates that VCD spectroscopy displays a higher sensitivity towards these derivatisations than ROA spectroscopy as in the latter case only small differences are noted. Next, the comparison between experimental and simulated spectra for all three artemisinin-type compounds is discussed.

In order to avoid any bias towards spectral agreement between experiment and calculations by an expert eye and to quantify the degree of similarity, different approaches exist to perform such an analysis.[36–38] Whatever similarity method is applied, all methods rely on the calculation of a spectral overlap between the experimental and predicted spectra in a specific wavenumber range ($\tilde{\nu}_1, \tilde{\nu}_2$). In the current work, the following overlap integral $S_{fg}$ is used:

$$S_{fg} = \frac{\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} f(\sigma\tilde{\nu}) g(\tilde{\nu}) d\tilde{\nu}}{\sqrt{\int_{\tilde{\nu}_1}^{\tilde{\nu}_2} f(\sigma\tilde{\nu})^2 d\tilde{\nu} \int_{\tilde{\nu}_1}^{\tilde{\nu}_2} g(\tilde{\nu})^2 d\tilde{\nu}}} .100\%$$

In which $f(\sigma\tilde{\nu})$ denotes for scaled calculated spectrum and $g(\tilde{\nu})$ the experimental spectrum. For the comparison between Raman/IR spectra, $S_{fg}$ can take values between 0% and 100%. When ROA/VCD spectra are subjected to the overlap integral analysis, the values range form -100% and 100%, where the former indicates that the two compared spectra are mirror images around the wavenumber-axis (enantiomer spectra) and the latter reflects a perfect overlap with the simulated spectrum. A scaling factor $\sigma$ in the calculated spectrum is used since the wavenumbers in quantum mechanical (QM) calculations are overestimated due to *inter alia* the harmonic approximation. Here, a global scaling factor of 0.982 has been determined by maximizing the overlap of artemisinin taking all spectra (Raman, IR, ROA and VCD) into account within the spectral range 950 - 1550 cm$^{-1}$ (see figure **??** in ESI†). All reported $S_{fg}$ values in the next section are calculated in this wavenumber range.

**Comparison of experimental and calculated Raman/ROA and IR/VCD spectra of artemisinin-type compounds.**

Conformational analysis showed only one dominant (> 99.9%) conformer for artemisinin. In the case of dihydroartemisinin and artesunate, two possible epimeric forms at C-10 need to be considered and therefore, conformational analysis on both epimeric forms were carried out. When the OH group or the succinic anhydride tail at C-10 is placed in the equatorial position it will be referred to the $\alpha$-epimer whereas when placed in the axial position it is further referred to as the $\beta$-epimer. In a previous VCD study on dihydroartemisinin only the $\beta$-epimer was taken into account showing only one dominant conformer.[24] The same conformer was found in this study to account for more than 98%. The $\alpha$-epimer on the other hand, shows two conformers differing from each other in the orientation of the OH group with an almost equal contribution in the Boltzmann weight. $\alpha$- and $\beta$-artesunate exhibit respectively 16 and 18 conformers. Calculating both epimeric forms of dihydroartemisinin and artesunate allows us to verify if Raman/ROA and IR/VCD can distinguish between the $\alpha$- and $\beta$ epimers of dihydroartemisinin and artesunate.

Figure 4 shows the experimental and calculated Raman/ROA (left) and IR/VCD (right) spectra of artemisinin. By using a global scaling factor of 0.982, an almost perfect overlap between the experimental and calculated Raman spectrum is obtained for the artemisinin case. The carbonyl stretch, found experimentally around 1740 cm$^{-1}$ is still slightly overestimated in the calculations. For amide carbonyl stretching modes, it is known that they are more overestimated than the rest of the spectrum and thus need another scaling factor, as explored by Mensch *et al.*[3] This might be the same for artemisinin, taking into account that the carbonyl stretch arises from a lactone moiety instead of an amide

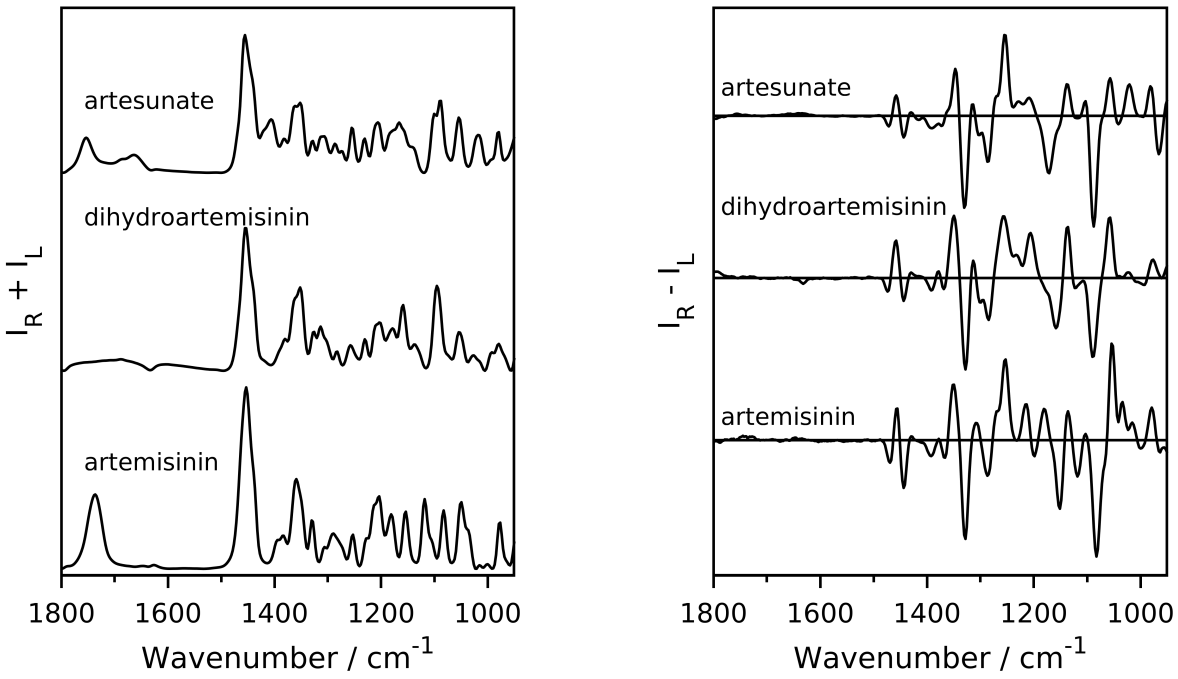


**Fig. 2** Experimental Raman (left) and ROA (right) spectra of artemisinin, dihydroartemisinin and artesunate in $CDCl_3$.

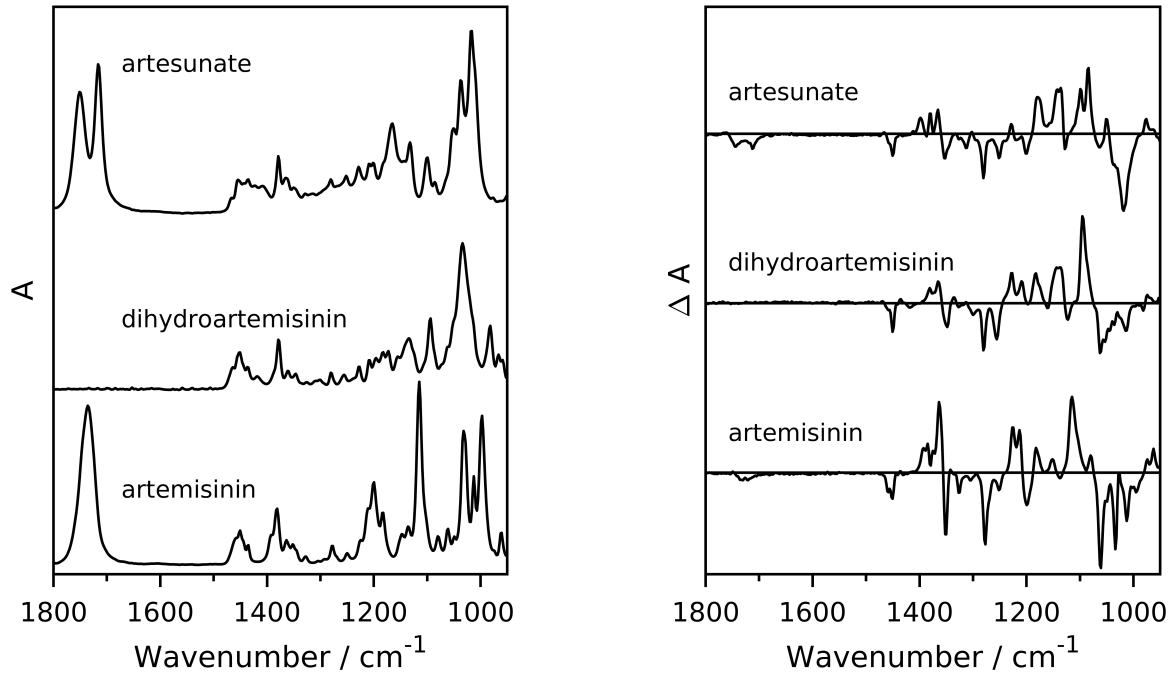


**Fig. 3** Experimental IR (left) and VCD (right) spectra artemisinin, dihydroartemisinin and artesunate in $CDCl_3$.

carbonyl stretch. Nevertheless, using a global scaling factor, a good match is obtained between experiment and calculation in the 950 - 1550 cm$^{-1}$ region of the spectrum. The same can be said for the ROA spectra where all bands observed in the experimental ROA spectrum of artemisinin are well described by the calculations. The $S_{fg}$ values of 88.4% and 83.0% (listed in table 1) for the Raman and ROA spectra of artemisinin respectively quantitatively reflect the high visual agreement between the ex-

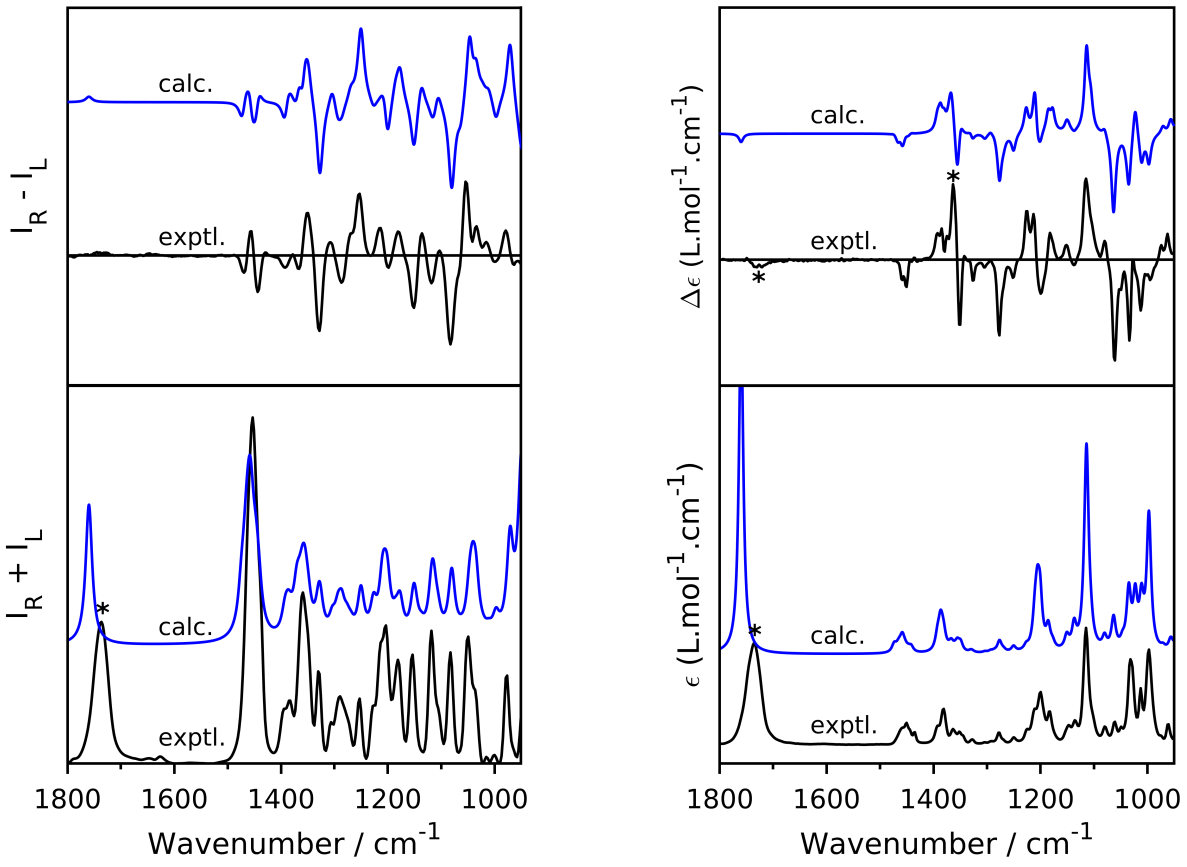


**Fig. 4** Left: Comparison of the experimental (black) and calculated (blue) Raman (bottom) and ROA (top) spectra of artemisinin in $CDCl_3$. Right: Comparison of the experimental (black) and calculated (blue) IR (bottom) and VCD (top) spectra of artemisinin in $CDCl_3$. A global scaling factor ($\sigma$) of 0.982 was used on the calculated wavenumbers. The bands highlighted by an asterisk are those discussed in the main text. Figure **??** shows the experimental and calculated spectra on top of each other (ESI†).

perimental and calculated data.

In line with the calculated Raman and ROA spectra, the IR and VCD spectra are globally scaled using a factor ($\sigma$) of 0.982 (figure 4). Except for the overestimation, of the carbonyl stretch of artemisinin in the IR spectra, both in position and intensity, a very good visual agreement can be found between the experimental and simulated IR spectra. The same is true for the VCD spectra where all experimentally observed bands are well reproduced by the calculations. Even the small positive signal arising from the carbonyl stretch is correctly predicted by the calculations. The largest deviation can be found in the sharp +/- couplet around 1350 cm$^{-1}$, which is less pronounced in the calculated spectrum. This very high degree of similarity both for the IR and VCD spectra is also confirmed by the $S_{fg}$ values being respectively 96.2% and 85.5% (table 1). These values indicate that both VOA techniques are equally accurate in describing the spectral pattern of artemisinin.

In the case of dihydroartemisinin, the two epimers are calculated and are compared to the experimental spectra, which are shown in figure 5. The calculated Raman spectrum of $\alpha$-dihydroartemisinin shows a good agreement with the experiment. However, the calculated Raman spectrum of $\beta$-dihydroartemisinin is highly similar to the $\alpha$-form spectrum making it impossible to distinguish between those two epimers solely based on the Raman spectrum. This is also reflected by the very close $S_{fg}$(Raman) values being 88.3% and 88.6% for $\alpha$- and $\beta$-dihydroartemisinin respectively. When focussing on the calculated IR spectra, more discrepancies in band shape and intensities are observed. Furthermore, the signal around 980 cm$^{-1}$ in the experimental spectrum hints towards the $\beta$-epimer, as this is the only epimer that predicts a band at that position. Despite the quite similar $S_{fg}$(IR) values, 85.3% and 90.9% respectively for $\alpha$- and $\beta$-dihydroartemisinin, these visual observations hint towards the $\beta$-form. Nevertheless, based on the whole shown range, we argue that neither Raman nor IR can unambiguously distinguish between the $\alpha$- and $\beta$-epimer of dihydroartemisinin.

The calculated ROA spectra of the two epimeric forms show, in agreement with their corresponding Raman spectra, again a very similar pattern. Nonetheless, below 1100 cm$^{-1}$ a more distinct pattern between the two calculated spectra is observed, illustrating that ROA indeed is sensitive to the epimeric form of dihydroartemisinin. The sharp positive band from the +/- cou-

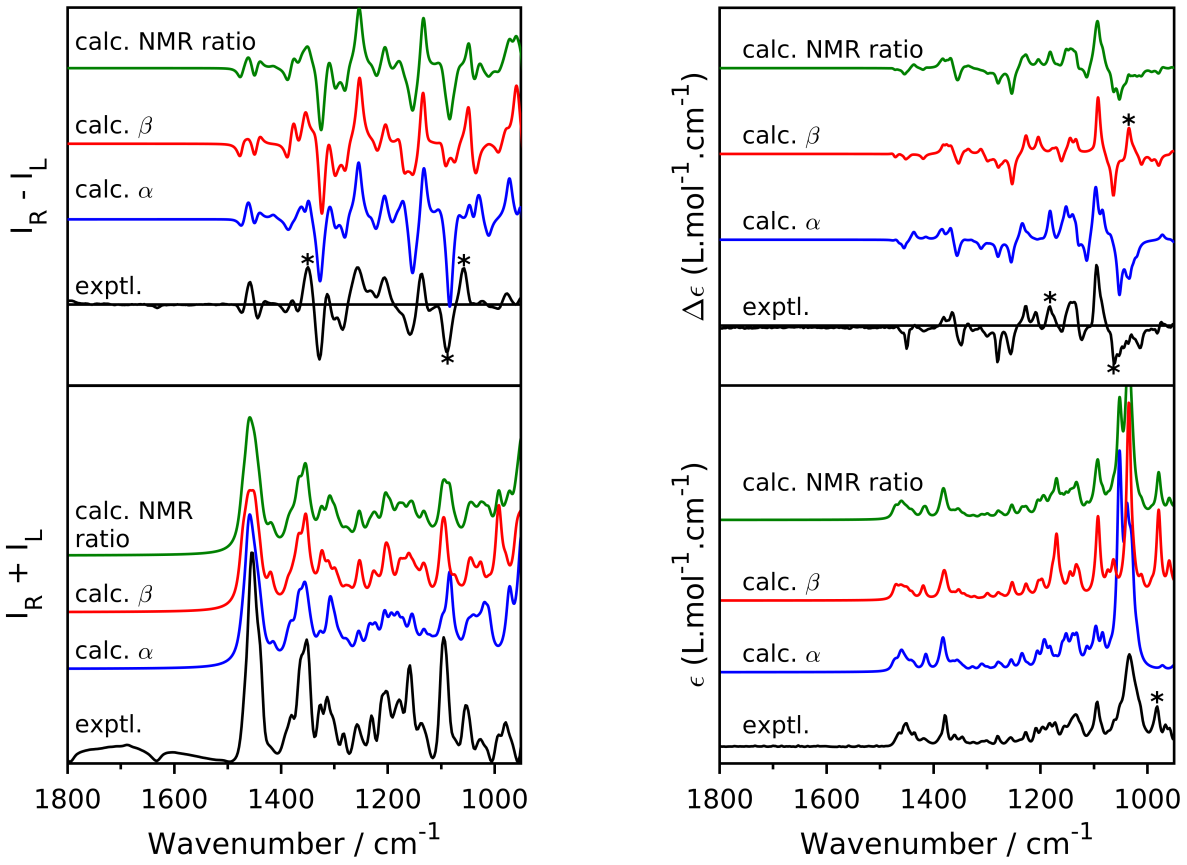


**Fig. 5** Left: Comparison of the experimental (black) and calculated Raman (bottom) and ROA (top) spectra of dihydroartemisinin in $CDCl_3$. Right: Comparison of the experimental (black) and calculated IR (bottom) and VCD (top) spectra of dihydroartemisinin in $CDCl_3$. The calculated $\alpha$- and $\beta$-epimeric forms are shown in blue and red respectively. The green spectra correspond to a $\alpha$ and $\beta$ combined spectrum with $^1$H-NMR determined weights. A global scaling factor ($\sigma$) of 0.982 was used on the calculated wavenumbers. The bands highlighted by an asterisk are those discussed in the main text. Figure **??** shows the experimental and calculated spectra on top of each other (ESI†).

plet around 1100 cm$^{-1}$ observed experimentally is only present as a small shoulder in the predicted $\alpha$-epimer spectrum but the negative part is well described by the predicted spectrum. In contrast, the positive part is well described in the $\beta$-epimer spectrum but at the same time this spectrum is lacking the sharp negative part of the couplet. Another noticeable difference is the sharp positive band of the -/+ couplet around 1350 cm$^{-1}$ which is slightly better described by the predicted $\beta$-epimer spectrum. Since none of the two epimeric spectra comes out as a clear winner after visual inspection, the AC determination of the C-10 chiral center of dihydroartemisinin is impossible solely based on ROA spectroscopy. Unfortunately also the $S_{fg}$(ROA) values (table 1) cannot aid in an unambiguous assignment. When focussing on the calculated VCD spectra, the largest differences between $\alpha$- and $\beta$-dihydroartemisinin are observed below 1100 cm$^{-1}$. The IR spectra already hinted towards the presence of the $\beta$-epimer (*vide supra*), however when the calculated VCD spectrum of $\beta$-dihydroartemisinin is compared with the experimental spectrum, some discrepancies are observed in the low wavenumber part of the shown range. In particular, the sharp positive band calculated at 1050 cm$^{-1}$ is not present in the experimental VCD spectrum, displaying an almost monosignate negative signal in that part of the spectrum. Also the broader positive signal just below 1200 cm$^{-1}$ in the experimental VCD spectrum is not well predicted by the $\beta$-epimer spectrum. These observations are more in agreement with the predicted $\alpha$-epimer spectrum and are also reflected by a larger $S_{fg}$(VCD) value for $\alpha$- than for $\beta$-dihydroartemisinin being 80.2% and 74.2% respectively. These considerations hint towards the $\alpha$-form of dihydroartemisinin present in the experimental spectrum, which is a very peculiar observation as the IR and VCD spectral analysis seemingly contradict each other.

A recent VCD study by Zhang *et al.*[24] on dihydroartemisinin in DMSO$-d_6$ only used the $\beta$-epimer in their calculations, and argued that the mismatches they had observed between experiment and calculations are due to solvent interaction to the OH group of dihydroartemisinin. However, in the present study, $CDCl_3$ is used as a solvent in which such an interaction is very unlikely to occur.[16] A 1D $^1$H-NMR experiment revealed that both the $\alpha$- and $\beta$-epimers are present in $CDCl_3$ (see figure **??** in ESI†). Con-

**Table 1** $S_{fg}$ for Raman/ROA and IR/VCD spectra of the three artemisinin-type compounds in the 950 - 1550 cm$^{-1}$ spectral range

| Compound | $S_{fg}$(Raman) (%) | $S_{fg}$(ROA) (%) | $S_{fg}$(IR) (%) | $S_{fg}$(VCD) (%) |
|---|---|---|---|---|
| Artemisinin | 88.4 | 83.0 | 96.2 | 85.5 |
| $\alpha$-Dihydroartemisinin | 88.3 | 73.0 | 85.3 | 81.5 |
| $\beta$-Dihydroartemisinin | 88.6 | 70.5 | 90.9 | 76.0 |
| $\alpha/\beta$-Dihydroartemisinin | 89.5 | 78.6 | 93.8 | 93.7 |
| $\alpha$-Artesunate | 91.5 | 84.0 | 90.6 | 87.8 |
| $\beta$-Artesunate | 89.7 | 56.5 | 61.2 | 35.6 |

sidering that some parts of the experimental ROA/VCD spectra were better described by one epimeric form and other parts by the other epimer, this did not come as a complete surprise.

When the singlet peak of the proton at C-12 is used for integration, the ratio between $\alpha$- and $\beta$-dihydroartemisinin present in solution is calculated as 50% $\alpha$- and 50% $\beta$-dihydroartemisinin (calculated enthalpies and free energies of these two epimeric forms are listed in table **??**). When this ratio is taken into account in the simulated ROA and VCD spectra of dihydroartemisinin, the visual agreement substantially increases in both cases and all experimental bands observed are well represented by the simulations and all $S_{fg}$ values increase as well (table 1). This evidence supports the above mentioned statement that the $S_{fg}$ values should be used as an additional help to the visual agreement and not as a sole base leading to potentially wrong assignments.

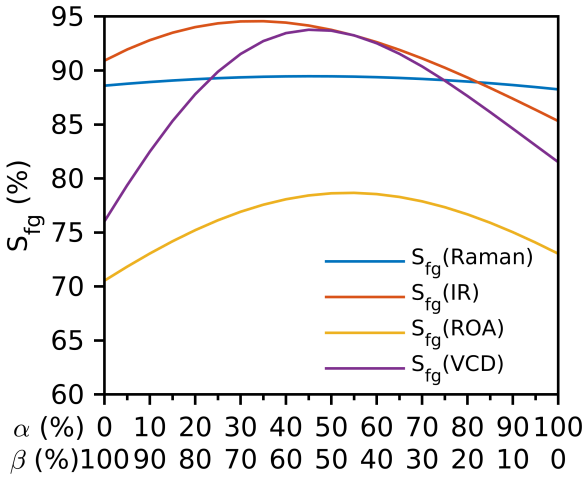


**Fig. 6** $S_{fg}$ values for Raman, IR, ROA and VCD with different $\alpha/\beta$ epimeric ratios of dihydroartemisinin calculated in steps of 5%.

To check if Raman/ROA and IR/VCD can be used to determine the $\alpha/\beta$ ratio, $S_{fg}$(Raman), $S_{fg}$(IR), $S_{fg}$(ROA) and $S_{fg}$(VCD) values were calculated between the experimental obtained spectra and the calculated spectra having different $\alpha/\beta$ ratios, similar to recent work published by Koenis *et al.* in which they use VCD spectroscopy to determine the diastereomeric ratio of dydrogesterone and 6-dehydroprogesterone.[39] The only difference with Koenis *et al.* is that here only computationally the ratio of the two epimers can be changed and that the experimental sample is fixed. The results of the current work are displayed in figure 6 and corresponding numerical $S_{fg}$ values are summarised in table **??** in ESI†.

Figure 6 demonstrates that the agreement between experimental and simulated Raman spectra is almost independent of the epimeric ratio used. This is probably due to the broad signals in which the fine structure is less defined and the small changes in the calculated spectral profile are not reflected by the $S_{fg}$(Raman) value and as stated before, the Raman spectra of $\alpha$- and $\beta$-dihydroartemisinin have very similar $S_{fg}$(Raman) values (*vide supra*). The $S_{fg}$(IR) values give a more clear maximum, indicating that IR spectroscopy is more sensitive than Raman spectroscopy towards the epimeric ratio. A similar curve is found for the $S_{fg}$(ROA) values, albeit with the maximum shifted towards the epimeric ratio determined by $^1$H-NMR. The $S_{fg}$(VCD) values display the sharpest curve accentuating the higher sensitivity of VCD spectroscopy towards the epimeric ratio than ROA spectroscopy. Furthermore, the $S_{fg}$(IR), $S_{fg}$(ROA) and $S_{fg}$(VCD) values emphasises the importance to consider an epimeric mixture to achieve the highest similarity between experiment and calculation as those three displaying a noticeable large overlap when a ratio of the two epimers is used in comparison with the 100% $\alpha$- or 100% $\beta$-epimer of dihydroartemisinin. As such, based on the maxima of the $S_{fg}$(ROA) and $S_{fg}$(VCD) (see table **??** in ESI†) one can predict the epimeric ratio to be approximately 50/50. However, we estimate that the uncertainty on the absolute epimeric ratio is higher here than during the $^1$H-NMR analysis because the interpretation of a difference in $S_{fg}$ values is hard to express in exact uncertainties. Therefore, the authors argue that ROA and VCD in combination with simulations can be used only to derive the approximate epimeric ratio that is present in the experimental spectrum.

As we have highlighted that $S_{fg}$ values should be accompanied with a visual inspection of the experimental and calculated spectra, figure 7 represents the ROA and VCD spectra of the different calculated ratios in comparison with the experimental spectra. The corresponding Raman and IR spectra are shown in supplementary information (see figure **??** in ESI†).

Figure 7 gives further insight on how the spectral pattern changes depending on the different epimeric ratios. Although the ROA spectra are very similar in the 1100 - 1550 cm$^{-1}$ spectral range for all calculated epimeric ratios, a clear variation in the spectral profile is found around 1100 cm$^{-1}$, showing that in order to represent the experimental observed couplet, a ra-

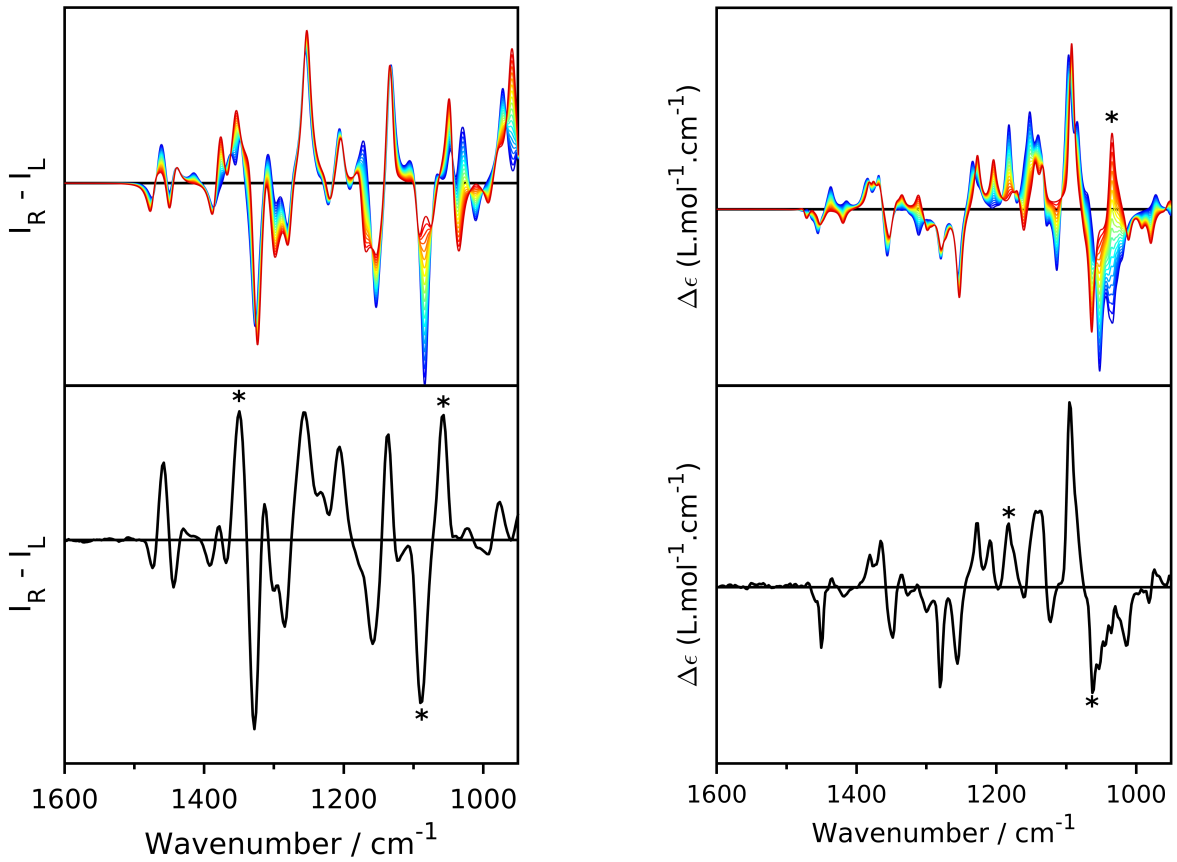


**Fig. 7** Comparison of experimental (bottom) and calculated (top) ROA (left) and VCD spectra (right) of dihydroartemisinin with different ratios of $\alpha$- and $\beta$-dihydroartemisinin. The 100% $\alpha$-dihydroartemisinin spectra are shown in blue which gradually changes to red representing 100% $\beta$-dihydroartemisinin. A global scaling factor ($\sigma$) of 0.982 was used on the calculated wavenumbers. The bands highlighted by an asterisk are those discussed in the main text.

tio of both forms should be taken into account when analysing the experimental spectra. An investigation of the VCD spectra indicates that over the whole wavenumber range, variation in the spectral profile can be observed, being the most prevalent in the low wavenumber (below 1100 $cm^{-1}$) region. This reflects again that, in comparison with the ROA spectra, VCD is more sensitive towards the presence of $\alpha$- and $\beta$-dihydroartemisinin. Furthermore, figure 7 accentuate that indeed a specific ratio between $\alpha$ and $\beta$-dihydroartemisinin needs to be taken into account to simulate the low wavenumber pattern in the experimental observed VCD spectrum.

Finally, we present the comparison between experimental and calculated spectra of artesunate. To parallel the dihydroartemisinin case, both epimeric forms were calculated and are discussed to see if Raman/ROA and IR/VCD can distinguish between the two epimers. As such, the experimental Raman/ROA and IR/VCD spectra together with the calculated spectra of both $\alpha$- and $\beta$-artesunate are shown in figure 8.

The two carbonyl modes, present in the Raman spectrum of artesunate, are slightly overestimated for both the $\alpha$ and $\beta$ epimer of artesunate when a global scaling factor is used. Overall the calculated Raman spectra of the two epimers show a high degree of similarity and on their own have a general good agreement with the experiment. This is also represented by very similar $S_{fg}$(Raman) values, being respectively 91.5% and 89.7% (table 1) for $\alpha$- and $\beta$-artesunate. However, the region below 1100 $cm^{-1}$ is better described by the $\alpha$-epimer spectrum. By investigation of the low wavenumber region in the IR spectra, we can exclude the presence of the $\beta$-epimer. Especially below 1000 $cm^{-1}$, whereas the $\beta$-artesunate show quite strong intensities, no IR bands are observed in the experimental IR spectrum. This is in line with the predicted $\alpha$-artesunate IR spectrum. Also, the strong IR bands experimentally observed in the 1000 - 1100 $cm^{-1}$ range are much better represented by the $\alpha$-epimer. This is supported by the $S_{fg}$(IR) values which are respectively 90.6% and 61.2% for $\alpha$- and $\beta$-artesunate (table 1). Intriguingly, Raman spectroscopy as stand alone tool could not unambiguously differentiate between these two epimers of artesunate.

Regarding the ROA spectra, again a high similarity is observed between the calculated spectra of $\alpha$- and $\beta$-artesunate in the 1100 - 1550 $cm^{-1}$ wavenumber range although some minor difference in relative intensities are observed. Furthermore, in the

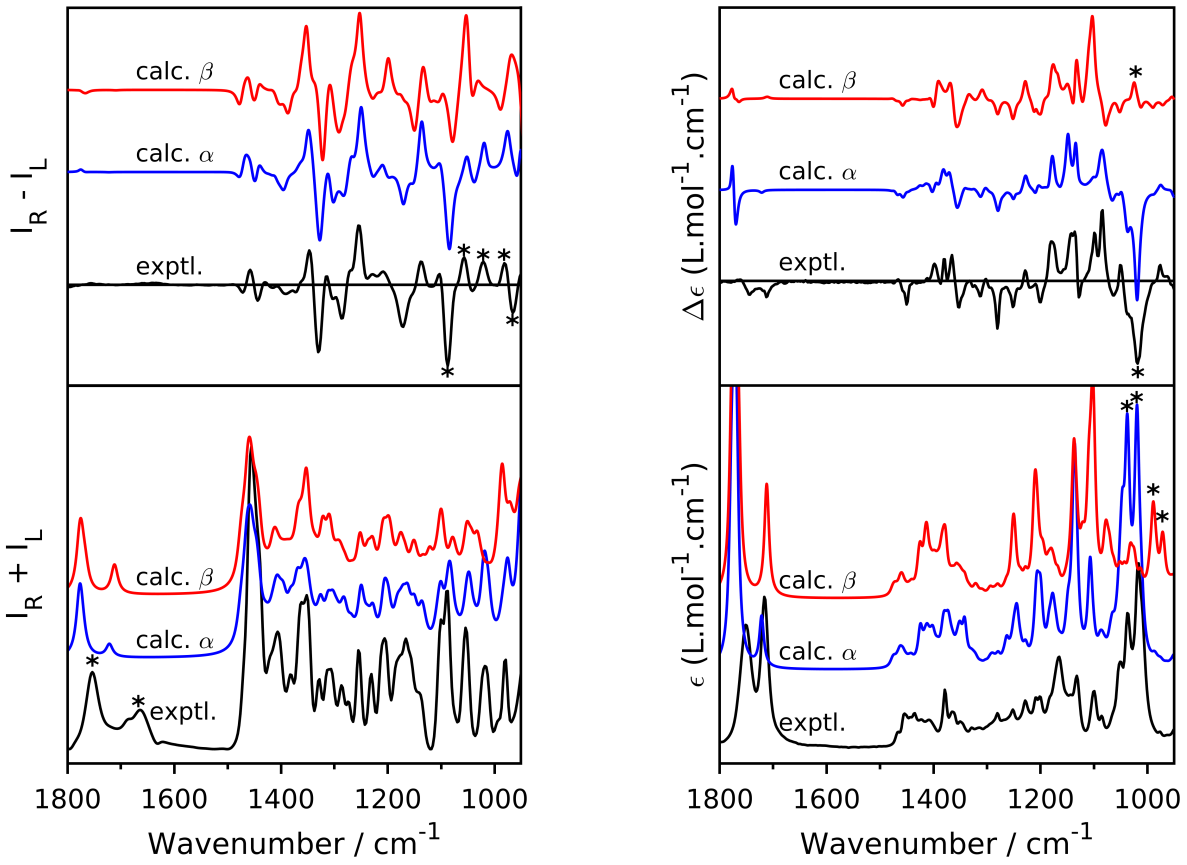


**Fig. 8** Left: Comparison of the experimental (black) and calculated Raman (bottom) and ROA (top) spectra of artesunate in $CDCl_3$. Right: Comparison of the experimental (black) and calculated IR (bottom) and VCD (top) spectra of artesunate in $CDCl_3$. The calculated $\alpha$- and $\beta$-epimeric forms are shown in blue and red respectively. A global scaling factor ($\sigma$) of 0.982 was used on the calculated wavenumbers. The bands highlighted with by asterisk are those discussed in the main text. Figure **??** shows the experimental and calculated spectra on top of each other (ESI†).

low wavenumbers range (950 - 1050 $cm^{-1}$) a distinct pattern is observed between the two epimers. When compared to the experimental ROA spectra of artesunate, the -/+/+/+/- pattern in that low wavenumber region is best represented by the calculated $\alpha$-artesunate ROA spectrum while the rest of the spectral range is almost equally well represented by both epimers. The better visual agreement for the $\alpha$-form is also clearly represented by the $S_{fg}$ values in the 950 - 1550 $cm^{-1}$ wavenumber range, being 84.0% and 56.5%, respectively for $\alpha$- and $\beta$-artesunate. Hence, adding a chiroptical dimension to the Raman scattering phenomenon makes it possible to distinguish between $\alpha$- and $\beta$-artesunate whereas Raman spectroscopy could not.

The VCD spectra reinforce the assignment made by the IR spectra when looking at the same spectral region (below 1100 $cm^{-1}$). The strong negative signal, experimentally observed around 1020 $cm^{-1}$ is solely represented by the $\alpha$-epimer and is even opposite in sign in the predicted $\beta$-epimer VCD spectrum. This is also reflected by a larger difference in the $S_{fg}$(VCD) values, being 87.8% for $\alpha$-artesunate and only 35.6% for $\beta$-artesunate, than was the case for the $S_{fg}$(ROA) values. Notwithstanding, both ROA and VCD assign the AC at C-10 of artesunate as $\alpha$. This is in agreement with the $^1$H-NMR data (see figure **??** in ESI†), that assigned the $\alpha$-form present in solution based on the the large coupling of H-10 at 5.67 ppm ($^3J_{10,9}$ = 9.7 Hz).[40]

All the $S_{fg}$ values indicate that in the case of artemisinin Raman/ROA performs equally well as the more established IR/VCD method in the AC determination. However, IR/VCD performs a slightly better job than Raman/ROA in the case of dihydroartemisinin and artesunate. Nonetheless the results show the potential of ROA to be used as a tool to distinguish between epimers.

So far, we have discussed the possibility to distinguish between the $\alpha$- and $\beta$- epimers of dihydroartemisinin and artesunate via ROA and VCD spectroscopy. However, the ability of ROA and VCD to determine the full AC and to distinguish between all diastereomers has not yet been explored. As dihydroartemisinin appears as an epimeric mixture, one should consider the possibility that this happens for all its diastereomers. The $\alpha/\beta$ epimeric ratio of all other possible diastereomers of dihydroartemisinin cannot be verified experimentally (e.g. by $^1$H-NMR) as they are not naturally or commercially available. This implies that every possible epimeric mixture ratio should be considered as was done in the

previous section. Such an analysis would be severely complicated by false positives (e.g. random band assignment), which would diminish the quality of the analysis and therefore we only evaluate a full AC for artemisinin and artesunate in the presented work.

### Can ROA and VCD distinguish between all diastereomers of artemisinin and artesunate?

Artemisinin has seven chiral centres yielding 128 ($=2^7$) theoretical diastereomers. However, two of the chiral centres are fixed together through the endoperoxide bridge (centres at atom number 3 and 12a in figure 1). Therefore, only 64 ($=2^6$) diastereomers are taken into account. Furthermore, as half of all these possibilities are enantiomers and thus have the same Raman and IR spectra and opposite in sign ROA and VCD spectra, only for 32 diastereomers conformational analysis and subsequent spectra calculations are performed. In comparison, artesunate exhibits one extra chiral centre (at atom number C-10 in figure 1) and therefore conformational analysis on 64 diastereomers and subsequent spectral calculations are carried out.

The scaling factor ($\sigma$) used in the previous section was specifically determined to maximise the overlap for the correct AC of artemisinin. Therefore using the same scaling factor for all other diastereomers might bias the AC assigning of artemisinin or artesunate using ROA and VCD without prior knowledge on the stereochemistry. Hence, for every single diastereomer an optimised scaling factor was used based on all corresponding experimental spectra available (e.g. Raman, IR, ROA and VCD), as was done in the previous part. The final scaling factors are listed in table **??** and in table **??** for respectively artemisinin and artesunate (ESI†).

The $S_{fg}$(ROA) and $S_{fg}$(VCD) values for all possible diastereomers of artemisinin are shown as bar plots in figure 9. The corresponding bar plots for Raman and IR are shown in figure **??** (ESI†). The $S_{fg}$(Raman) values indicate that Raman spectroscopy alone is incapable of assigning the correct AC of artemisinin. IR on the other hands performs slightly better, as the correct AC has the highest $S_{fg}$ value, indicating that IR spectroscopy is more sensitive towards diastereomers than Raman spectroscopy in the case of artemisinin. From figure 9, one clear winner can be extracted which is the same for ROA and VCD and is the one displaying the correct AC of artemisinin. In addition, this figure shows the strength of combining the two techniques to unambiguously assign the AC of artemisinin (RSRSRSR) without any prior knowledge. For example the diastereomer with an AC of SSSSSRS (second to left, green bar in $S_{fg}$(ROA)) has a quite high $S_{fg}$(ROA) value but is countered by a low $S_{fg}$(VCD) value. Likewise, it is found that the high $S_{fg}$(VCD) values of diastereomers RRRSRSR, RSSSRRR and RSSSRSR are countered by the corresponding $S_{fg}$(ROA) values highlighting the complementary of the two techniques. The only diastereomer that exhibits high $S_{fg}$ values for both ROA and VCD next to the correct AC, is RSRSSSR which is not surprisingly an epimer (at C-9) of artemisinin. Nevertheless, the authors argue that the difference in $S_{fg}$ values and the differences in the spectral patterns (see figure **??** in ESI†), in comparison with the correct AC of artemisinin, is large enough to

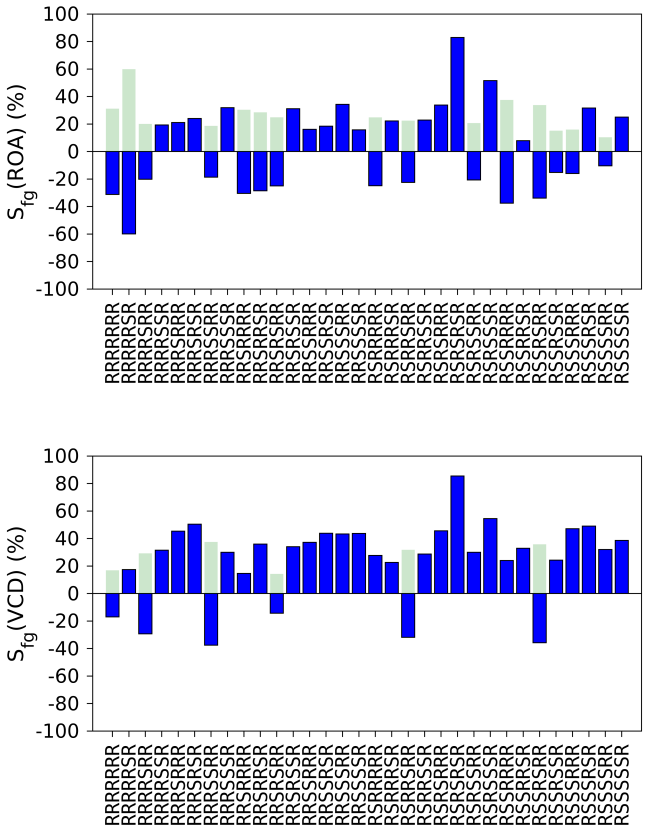


**Fig. 9** $S_{fg}$(ROA) (top) and $S_{fg}$(VCD) (bottom) for all possible diastereomers of artemisinin. The blue bars show the value of the diastereomer as indicated in the labels in which the order is the same as marked in figure 1. For the calculated diastereomers that give a negative $S_{fg}$ value, the green bars indicate the $S_{fg}$ value of the corresponding enantiomer (e.g. absolute value of the $S_{fg}$ value). This to make the comparison with the positive value of the correct AC (RSRSRSR) more clear. Note that those enantiomers are not explicitly calculated.

unambiguously assign the AC of artemisinin based on ROA and VCD.

In the previous section we already showed that ROA and VCD can distinguish between two possible diastereomers of artesunate (e.g. $\alpha$- and $\beta$-artesunate, or here RSRSRSRR and RSRSRRRR). In this section, we extend the analysis to all possible diastereomers. Figure **??** shows that both ROA and VCD assign the highest $S_{fg}$ values to the correct AC of $\alpha$-artesunate out of all 128 ($=2^7$) possible diastereomers. However, some peculiarities are observed in comparison with the artemisinin case. Here, the $S_{fg}$(IR) values are inconclusive as the $S_{fg}$(IR) value for the correct diastereomer ($\alpha$-artesunate, RSRSRSRR) is not unambiguously the highest over all possible diastereomers indicating that the extra dimension of chiroptical spectroscopy is needed to assign the stereochemistry for both Raman and IR spectroscopies. An interesting note to make is that also in the case of artesunate one diastereomer, RSRSSSRR, exhibits in both ROA and VCD high $S_{fg}$ values, which is an epimer of $\alpha$-artesunate at C-9 (figure 1) and is the same epimeric form of the core structure that was found as second best in the artemisinin case. The calculated ROA and VCD

spectra of $\alpha$-artesunate (RSRSRSRR) and RSRSSSRR are shown in figure **??** in ESI†. The differences in the spectral pattern between those two epimers are much more subtle in the VCD than in the ROA spectra, accentuating that in some case ROA performs visually better in distinguishing between epimers than VCD. Besides RSRSSSRR, some $S_{fg}$ values of the VCD spectra of other diastereomers approximate the one of the highest and correct AC of artesunate, which raises doubt in a conclusive AC determination. Here, ROA comes to the rescue by exhibiting small $S_{fg}$ values for the cases e.g. for the diastereomers RSSSRRSR and RSSRSSRR. This highlights once more the potential strength of ROA in AC determination of (natural) products and especially the power of combining these two chrioptical techniques in AC assignments.

## Conclusions

Artemisinin and two of its derivatives, dihydroartemisinin and artesunate were selected as a case study to compare Raman/ROA and IR/VCD spectroscopy in their ability to determine the AC of these compounds and to investigate their complementarity in those assignments. The presented study shows that ROA and VCD are able to distinguish between two specific epimers of artesunate and that they are able to hint towards an epimeric ratio present in the experimental sample in the case of dihydroartemisinin that is in accordance with $^1$H-NMR spectroscopy. The study on dihydroartemisinin also showed that when solely looking at the $S_{fg}$ values, an erroneous conclusion could potentially be drawn and that in the first place the visual agreement between experiment and calculation must be convincing. Furthermore, a full AC determination of artemisinin and artesunate was performed. Although ROA and VCD could independently assign the correct stereochemistry of both natural products out of all their respective possible diastereomers, it was found that when the results of ROA and VCD on all possible diastereomers were combined an even stronger, unambiguous assignment could be made. Therefore, the presented work highlights the necessity to further explore ROA spectroscopy in AC determination and that it should be considered more often as a complementary technique to VCD spectroscopy in the full AC determination toolbox to strengthen the AC assignments, especially when epimers needs to be distinguished.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

J.B. thanks the Research Foundation Flanders (FWO-Vlaanderen) for the appointment of a pre-doctoral scholarship (1198318N) and the Hercules foundation for the purchase of the NMR instrument. The Flemish Supercomputing Centre (VSC) is acknowledged for providing computational resources and support. Financial support allowing the development of the home-built SCP-ROA instrument through BOF, IOF and FWO-Vlaanderen is acknowledged.

## Notes and references

1 J. Hudecová, J. Kapitán, V. Baumruk, R. P. Hammer, T. A. Keiderling and P. Bouř, *Journal of Physical Chemistry A*, 2010, **114**, 7642–7651.

2 L. D. Barron, *Biomedical Spectroscopy and Imaging*, 2015, **4**, 223–253.

3 C. Mensch, L. D. Barron and C. Johannessen, *Physical Chemistry Chemical Physics*, 2016, **18**, 31757–31768.

4 J. R. Cheeseman, M. S. Shaik, P. L. A. Popelier and E. W. Blanch, *Journal of the American Chemical Society*, 2011, **133**, 4991–4997.

5 F. Zielinski, S. T. Mutter, C. Johannessen, E. W. Blanch and P. L. Popelier, *Physical Chemistry Chemical Physics*, 2015, **17**, 21799–21809.

6 R. Pendrill, S. T. Mutter, C. Mensch, L. D. Barron, E. W. Blanch, P. L. Popelier, G. Widmalm and C. Johannessen, *ChemPhysChem*, 2019, 695–705.

7 J. Bogaerts and C. Johannessen, *Journal of Raman Spectroscopy*, 2019, **50**, 641–646.

8 R. Sgammato, W. Herrebout and C. Johannessen, *Journal of Raman Spectroscopy*, 2019, **50**, 1905–1913.

9 L. A. Nafie, *Chemical Physics*, 1996, **205**, 309–322.

10 T. Fujisawa, R. L. Leverenz, M. Nagamine, C. A. Kerfeld and M. Unno, *Journal of the American Chemical Society*, 2017, **139**, 10456–10460.

11 J. Haesler, I. Schindelholz, E. Riguet, C. G. Bochet and W. Hug, *Nature*, 2007, **446**, 526–9.

12 A. F. Monteiro, J. M. Batista, M. A. Machado, R. P. Severino, E. W. Blanch, V. S. Bolzani, P. C. Vieira and V. G. Severino, *Journal of Natural Products*, 2015, **78**, 1451–1455.

13 V. Profant, A. Jegorov, P. Bouř and V. Baumruk, *The Journal of Physical Chemistry B*, 2017, **121**, 1544–1551.

14 E. De Gussem, K. A. Tehrani, W. A. Herrebout, P. Bultinck and C. Johannessen, *ACS Omega*, 2019, **4**, 14133–14139.

15 J. M. Batista Jr., E. W. Blanch and V. d. S. Bolzani, *Natural Product Reports*, 2015, **32**, 1280–1302.

16 C. Merten, T. P. Golub and N. M. Kreienborg, *The Journal of Organic Chemistry*, 2019, **84**, 8797–8814.

17 K. Bünnemann and C. Merten, *Physical Chemistry Chemical Physics*, 2017, **19**, 18948–18956.

18 L. Weirich and C. Merten, *Physical Chemistry Chemical Physics*, 2019, **21**, 13494–13503.

19 Y. Tu, *Angewandte Chemie - International Edition*, 2016, **55**, 10210–10226.

20 M. Ansari, Z. Saify, N. Sultana, I. Ahmad, S. Saeed-Ul-Hassan, I. Tariq and M. Khanum, *Mini-Reviews in Medicinal Chemistry*, 2013, **13**, 1879–1902.

21 *The Nobile Prize, http://www.nobelprize.org/nobel_prizes/medicine/laureates/2015/press.html (accessed December 2019)*.

22 Z. Wang, L. Yang, X. Yang and X. Zhang, *Synthetic Communications*, 2014, **44**, 1987–2003.

23 M. A. Corsello and N. K. Garg, *Natural Product Reports*, 2015, **32**, 359–366.

24 Y. Zhang, M. R. Poopari, X. Cai, A. Savin, Z. Dezhahang, J. Cheramy and Y. Xu, *Journal of Natural Products*, 2016, **79**, 1012–1023.

25 *CONFLEX 6 Program*, Conflex Corporation, Tokyo, 2012.
26 T. A. Halgren, *Journal of Computational Chemistry*, 1999, **20**, 720–729.
27 *Spartan'08*, Wavefunction, Inc. Irvine, CA, 2008.
28 T. A. Halgren, *Journal of Computational Chemistry*, 1996, **17**, 520–552.
29 M. Clark, R. D. Cramer III and N. Van Openbosh, *Journal of Computational Chemistry*, 1989, **10**, 982–1012.
30 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2009.
31 J. R. Cheeseman and M. J. Frisch, *Journal of Chemical Theory and Computation*, 2011, **7**, 3323–3334.
32 W. Hug and G. Hangartner, *Journal of Raman Spectroscopy*, 1999, **30**, 841–852.
33 W. Hug, *Applied Spectroscopy*, 2003, **57**, 1–13.
34 H. F. Boelens, R. J. Dijkstra, P. H. Eilers, F. Fitzpatrick and J. A. Westerhuis, *Journal of Chromatography A*, 2004, **1057**, 21–30.
35 S. Qiu, G. Li, P. Liu, C. Wang, Z. Feng and C. Li, *Physical Chemistry Chemical Physics*, 2010, **12**, 3005–3013.
36 T. Kuppens, W. Langenaeker, J. P. Tollenaere and P. Bultinck, *Journal of Physical Chemistry A*, 2003, **107**, 542–553.
37 J. Shen, C. Zhu, S. Reiling and R. Vaz, *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 2010, **76**, 418–422.
38 C. L. Covington and P. L. Polavarapu, *Chirality*, 2017, **29**, 178–192.
39 M. A. Koenis, E. H. Tiekink, D. M. van Raamsdonk, N. U. Joosten, S. A. Gooijer, V. P. Nicu, L. Visscher and W. J. Buma, *Analytica Chimica Acta*, 2019, **1090**, 100–105.
40 A. Presser, A. Feichtinger and S. Buzzi, *Monatshefte für Chemie - Chemical Monthly*, 2017, **148**, 63–68.