Assessing Cluster Models of Solvation for the Description of Vibrational Circular Dichroism Spectra: Synergy between Static and Dynamic Approaches

Katia Le Barbu-Debus, Jessica Bowles, Sascha Jähnigen, Carine Clavaguéra, Florent Calvo, Rodolphe Vuilleumier, Anne Zehnacker

Assessing Cluster Models of Solvation for the Description of Vibrational Circular Dichroism Spectra: Synergy between Static and Dynamic Approaches

Katia Le Barbu-Debus,[a)] Jessica Bowles,[b)] Sascha Jähnigen,[c)] Carine Clavaguéra,[b)] Florent Calvo,[d)] Rodolphe Vuilleumier,[c)] Anne Zehnacker [a)]

a) *Institut des Sciences Moléculaires d'Orsay (ISMO), CNRS, Université Paris-Saclay, F-91405 Orsay, France*
b) *Université Paris-Saclay, CNRS, Institut de Chimie Physique, UMR8000, 91405 Orsay, France*
c) *PASTEUR, Département de Chimie, Ecole Normale Supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France*
d) *Laboratoire Interdisciplinaire de Physique (LiPhy), CNRS, Université Grenoble Alpes, 140 Avenue de la Physique - BP 87 38402 Saint Martin d'Hères, France*

Abstract

Solvation effects are essential for defining the shape of vibrational circular dichroism (VCD) spectra. Several approaches have been proposed to include them into computational models for calculating VCD signals, in particular those resting on the "cluster-in-a-liquid" model. Here we examine the capabilities of this ansatz on the example of flexible (*1S*,*2S*)-*trans*-1-amino -2-indanol solvated in dimethyl sulfoxide (DMSO). We compare cluster sets obtained from static calculations with results from explicit molecular dynamics (MD) trajectories based on either force field (FF) or first-principles (FP) methods. While the FFMD approach provides a broader sampling of configurational space, FPMD and time-correlation functions of dipole moments account for anharmonicity and entropy effects in the VCD calculation. They provide a means to evaluate the immediate effect of the solvent on the spectrum. This survey singles out several challenges associated with the use of clusters to describe solvation effects in systems showing shallow potential energy surfaces and non-covalent interactions. Static structures of clusters involving a limited number of solvent molecules satisfactorily capture the main effects of solvation in the bulk limit on the VCD spectra, if these structures are correctly weighted. The importance of taking into consideration their fluxionality, *i.e.* different solvent conformations sharing a same hydrogen bond pattern, and the limitations of small clusters for describing the solvent dynamics are discussed.

Introduction

Vibrational circular dichroism (VCD) is an increasingly popular spectroscopic technique that since its discovery in the 1970s broadened its scope of applicability in both academia and industry.[1, 2] Referring to the absorption difference between left- and right-circularly polarised light, VCD is measured in the infrared (IR) region of the electromagnetic spectrum and can therefore be referred to as chiral IR spectroscopy. Hence, being associated to molecular vibrations, it delivers an extremely rich set of stereochemical information, painting the chirality of intra- or intermolecular regions involved in vibrational motion.[3] VCD has become an invaluable tool in the determination of absolute configurations by assigning recorded spectra to structural information obtained from theoretical calculations.[3] It has been applied in the fields of nanoscience,[4] catalysis,[5] solid-state organisation,[6-8] and has also been employed to detect protein fibrils such as amyloids responsible for neurodegenerative disorders.[9] From a more fundamental point of view, VCD is a very sensitive probe of conformational flexibility and molecular interactions.[10-18] At room temperature, flexible molecules can adopt many stable conformations corresponding to different local minima, each of them contributing to the VCD spectrum. Non-covalent interactions have a strong effect on the VCD spectra, even in cases where IR absorption remains unchanged.[15-17, 19] Consequently, effects like supramolecular chirality and chirality transfer can be addressed.[11, 20]

Crucial for a VCD study is the availability of efficient theoretical models needed to interpret and assign the experimental results.[1] In VCD spectroscopy experiments, the molecules of interest are usually solvated and one fundamental issue is to assess the role of the solvent on the molecular conformation and on the spectrum itself. Successful assignment of VCD signals based on computed spectra therefore requires that the solvent is correctly described, and that the spectrum is calculated with sufficient accuracy. Addressing these two tasks simultaneously is possible, in principle, by fully anharmonic calculations based on first-principles descriptions of the electronic structure in which the solute and the solvent are both explicitly accounted for. Such an approach is rigorous but computationally demanding. Among the various approximations that can be introduced to alleviate this shortcoming, continuum solvation models can be particularly appealing, as demonstrated for non-interacting systems such as weakly polar molecules in aprotic solvents.[21] This implicit approach, however, proves insufficient for describing non-covalent interactions like hydrogen bonds, since interactions with the environment influence the line position, the shape, and even the sign of VCD signals.[21, 22]

In the presence of strong intermolecular interactions, including at least some of the solvent molecules explicitly can thus be essential to describe properly measured VCD spectra in solution. In this context, the group of Xu proposed the “clusters-in-a-liquid” model,[23] which encodes a reduced description of the solute-solvent interaction in terms of long-lived clusters of definite size, themselves embedded in a solvation continuum. Being reminiscent of model concepts for liquids, such as the quantum cluster equilibrium (QCE) approach,[24, 25] it has proved to be very efficient for reproducing VCD spectra of small molecules in water[23,

[26-28] or of aggregated alcohols and acids in non-polar solvents.[29, 30] Furthermore, for aqueous solutions this model has led to reports of chirality induction in which vibration modes of solvent molecules become VCD active due to the presence of a chiral solute.[21, 26, 27]

In the cluster-in-a-liquid approach, the IR and VCD spectra of individual members of the cluster set are obtained from a static quantum chemical calculation based on the harmonic approximation, generally considered to be a reasonably accurate description for vibrational spectra of simple molecules.[31] The harmonic calculation involves evaluation of the second energy derivatives and diagonalisation of the Hessian matrix, and for an ergodic system at equilibrium the relative contribution of each local minimum is expected to be proportional to the Boltzmann factor.[20] However, the VCD spectra obtained by Boltzmann-averaging of the different minima contributions do not show optimal agreement with experiment for floppy molecules,[30, 32] suggesting the weights to be obtained instead by adjustment to the experimental VCD spectrum. The discrepancy is in part due to the non-reliability of the Boltzmann factors obtained at the DFT level, which is especially crucial for floppy molecules possessing numerous minima in a narrow energy window. The other reason is inherent to the limitations of a static picture: Vibrational averaging obtained by running linear transit scans along angles associated to low-frequency modes has indeed proved to considerably improve the agreement with experiment.[33, 34] Moreover, introduction of cluster of different sizes may be necessary for good agreement with the experiment. This can be achieved either by adjustment to the experimental VCD spectrum[26] or by using a cluster distribution deduced from the QCE theory.[35] Yet, cluster design on which the description of VCD spectra ultimately relies highly depends on the set of conformations chosen to represent the set, for which chemical intuition alone and even energetics-based strategies might be limited or even misleading.[32, 36] Recent work on *(R)*-2-butanol in CS_2 solution demonstrated that including the very stable ring-tetramer, which would expected from energetic issues, degrade the quality of the spectrum.[35] The clusters-in-a-liquid model focuses on limited numbers of solvent molecules, partly to accommodate with tractable computational resources. This raises the question of the number of solvent molecules needed in practice to describe solvation correctly. The group of Merten recently reported a VCD study of phenyl-containing 1,2-diols,[37] where they showed that in acetonitrile, the intramolecular OH…O hydrogen bond is retained and that merely the consideration of 1:1 solvent-solute complexes is enough to satisfactorily reproduce the experimental VCD spectra. In DMSO, in turn, inter- and intramolecular hydrogen bond formation compete with each other and, depending on the nature of the diol, one or two DMSO molecules are required to reproduce the spectrum. For 1-indanol in DMSO solution, FPMD calculations provided evidence for a stable hydrogen-bonded 1:1 complex similar to those used in the cluster-in-a-liquid model.[38] For *(R)*-2-butanol in CS_2 solutions, QCE approaches including large oligomers improved the agreement with experiment.[35, 39] Whatever the method used, the necessary finite size of the cluster is a limitation that precludes a rigorous comparison with full solvation. In addition to the anharmonicity and temperature issues, this confirms the possible limitations of static

calculations and motivates the use of unbiased approaches based on molecular modeling at approximate (force field) or more rigorous (first-principles) levels of theory.

Atomistic simulations such as molecular dynamics (MD) is a powerful tool to sample the phase space of a molecular system at finite temperature in a realistic way. From the MD trajectories, configurations can be regularly extracted and their energy minimised, their individual spectra being obtained with quantum mechanics (QM) methods, just as in the cluster-in-a-liquid approach.[38, 40] Assuming the MD trajectory to be sufficiently ergodic, averaging over the various individual spectra is expected to provide better comparison with realistic conditions as the molecular system explores regions of the potential energy surface far from minima. Beyond the fully atomistic description, multi-scale approaches separate regions of the chiral centre and the environment to effectively account for the mechanical or polarisation response of the environment towards molecular vibrations and *vice versa*.[41-43] In addition, pre- screening using MD simulations with a force field can be designed to be used in a fully automatic black-box scheme.[44] Though MD generates structures not always inferred by chemical intuition, the reconstruction of the VCD spectrum still relies on minimum-energy structures and the harmonic approximation, which both become computationally very demanding for large systems.[45, 46] Owing to the various binding sites of the solute, solvated compounds often exhibit a particularly rugged energy landscape that produces a large number of inequivalent conformations. At finite temperature, the relative energies of these local minima influence the Boltzmann distribution, hence the spectral features of the VCD spectra through the relative intensities of the various contributing minima. Being exponential functions, they can be particularly sensitive to slight conformational distribution changes and to the details of the underlying QM method.[47] A simpler approach consists in assigning arithmetic weights from snapshots regularly obtained from a MD trajectory, thereby accounting for thermal disorder and especially the various inequivalent orientations of the solvent molecules despite a limited number of hydrogen bond networks. [43, 48]

The IR and VCD spectra can also be obtained without resorting to the harmonic approximation by conducting MD simulations directly and Fourier transforming the appropriate time correlation functions (FT-TCF).[22, 42, 49-53] In spectroscopy, the quantum FT-TCF corresponds to the line shape function;[54, 55] its classical limit is a good approximation to the quantum response.[56] Abbate and co-workers introduced the concept of self- and cross-correlation functions of time-dependent electric and magnetic dipole moments for computation and analysis of circular dichroism spectra.[55] One of the first MD/VCD calculations following this ansatz, based on a quantum mechanics/ molecular mechanics (QM/MM) charge flow model, was presented by Cho and co-workers at the example of (*1S*)-(–)-β-pinene.[50] Since this early study, classical MD simulations based on force fields (FFs) have been also used to compute IR and VCD spectra.[51] However, conventional FFs usually do not reach the required accuracy for spectroscopy, which has led to recent FF development more strongly connected with quantum mechanical ingredients, notably through an

improved treatment of electrostatics.[57] In particular, polarisation and charge transfer effects have been introduced using induced atomic dipoles or fluctuating charges models. [58-60]

In the MD approach to IR and VCD spectroscopy via FT-TCF, first-principles methods are particularly attractive.[38, 61-63] Beyond the natural account for anharmonic and entropic effects, they provide a complete quantum description of the simulation cell, in particular of the dipole moments that are directly computed from the available electronic wave function. Two complementary quantum mechanical formulations of VCD are available, namely magnetic field perturbation theory (MFPT)[64, 65] and nuclear velocity perturbation theory (NVPT).[66] However, only the latter is suitable to obtain FT-TCF from FPMD since the electronic response can be calculated directly from the phase space of the trajectory. Accordingly, NVPT has been implemented into the CPMD code by Scherrer *et al.*,[67, 68] followed by applications in gas, liquid, and solid phases.[38, 61, 67, 69] Recently, the group of Kirchner presented an approximation to the quantum response, extending the charge-flow model towards time-dependent electron densities in first principles (FP) MD simulations of VCD spectra.[62] However, as with any simulation in which the electronic structure is accounted for explicitly, FPMD trajectories are computationally far more demanding than their force field counterparts are, leading to reduced time scales and limited phase space samplings that possibly lead to ergodicity problems.

We present here an IR absorption and VCD study of a flexible bifunctional molecule, (*1S*,*2S*)-*trans*-1-amino-2-indanol (*trans*-AI, Scheme 1) in DMSO solution. Only little attention has been paid so far to the influence of solvation on the VCD spectrum of 1,2-amino-alcohols.[70] The aim of this work is to unravel the interplay between the intramolecular structure and the solvation network, and explore how the solvation dynamics impinge the VCD spectrum of *trans*-AI. We use standard and polarisable FFMD to fully explore the potential energy surface of the monomer as well as that of its complexes with one and two DMSO molecules. Being designed for bulk DMSO, the FF used in this work is expected to reproduce satisfactorily the solvation trends observed in the system. Representative structures are extracted and further optimised at the density-functional theory (DFT) level to compute IR absorption and VCD harmonic spectra, in the frame of the cluster-in-a-liquid model. Special attention is paid to the factors that influence the reconstructed spectra, in particular through Boltzmann weights devoted to each cluster. Furthermore, polarisable FFMD at fixed temperature is used to explore extensively the potential energy surfaces of the *trans*-AI molecule solvated by one, two and five DMSO molecules, and compare it to the bulk limit. A picture of the solvent fluxionality is provided, which aims at understanding full solvation from the interaction of *trans*-AI with a limited number of DMSO molecules.

The most stable *trans*-AI monomer and its 1:1 clusters with one DMSO molecule are used as starting points for FPMD simulations and calculations of the VCD spectrum by means of the NVPT approach. By comparing the results obtained using these various methods with each other and, ultimately, with experimental measurements, we discuss the efficiency of the clusters-in-a-liquid approach and its fruitful connections to molecular dynamics methods.

Experimental and Theoretical Methods

1. Experimental Methods

The vibrational IR absorption and VCD spectra were measured using a FTIR spectrometer Vertex 70 equipped with a VCD module PMA 50 (Bruker), at a spectral resolution of 4 cm^{-1}. The IR radiation, filtered by a low-pass filter cutting at 2000 cm^{-1}, then polarised with a linear polariser, was modulated by a 50 kHz ZnSe photo-elastic modulator (Hinds). The signal was measured by a MCT IR detector with a BaF_2 window, cooled with liquid nitrogen. The output of the MCT detector was demodulated using a lock-in amplifier (Stanford Research Systems SR 830). The spectra were measured using ~1.1 M solutions in an adjustable cell (Harricks) with a path length of 110 µm and 156 µm. The position of the cell was adjusted by rotating it to minimise its linear dichroism. The alignment was then verified by checking the mirror-image relation between the VCD spectra of the two enantiomers of camphor (0.3 M in CCl_4) in the same cell as used here. The acquisition time was 8h. The spectra shown below are the half difference of those of the two enantiomers. They were recorded in two different cells to ensure reproducibility. The DMSO-*d6* solvent and the enantiopure (*1S,2S*)-(+)-*trans*-1-amino-2-indanol and (*1R,2R*)-(-)-*trans*-1-amino-2-indanol were purchased by Aldrich and used without further purification.

2. Theoretical Methods and Computational Details

2.1. Potential Energy Surface Exploration and Static Cluster-in-a-liquid Calculations

The conformational flexibility of *trans*-AI and the fluxionality of the solvent necessitate extensive explorations of the potential energy surfaces. For the *trans*-AI clusters with one or two DSMO-*d6* molecules, explorations were performed using the Monte Carlo Multiple Minimum method implemented in the MacroModel program[71] associated to the OPLS-2005 force field.[72] Furthermore, MD simulations were performed using the AMOEBA polarisable force field [73] with the aim to generate new low-energy structures that could have been missed with OPLS. In recent years, the AMOEBA polarisable force field has shown its capability to reproduce accurately the competition between different intermolecular interactions in solution and the solvent organisation around a solute.[73-75]

A set of multipoles was generated for the *trans*-AI isolated monomer using the distributed multipole analysis of the MP2/cc-pVTZ electron density.[76, 77] The conformations saved every picosecond were fully optimised with AMOEBA and the lowest-energy structures obtained in a 20 kJ/mol window were then reoptimised at the DFT level using the 6-311++G(d,p) basis set. For all conformations, the B3LYP functional was used.[78] This level of theory has been shown to reproduce the VCD spectrum of similar molecules in solution, as well as their structure and vibrational spectrum in the gas phase.[38, 79, 80] It has also been used successfully for calculating the VCD spectrum of aminoacids in aqueous solution.[43] Solvent effects were taken into account by the IEFPM implicit polarisable continuum model.[81] Vibrational

frequencies were computed at the same level of theory and scaled by 0.98 for correcting for anharmonicity and basis set incompleteness. This value is close to that used for similar theoretical methods.[82] [83] The final vibrational spectra were obtained by convoluting the harmonic intensities with a Lorentzian line shape (FWHM 4 cm-1).

All the DFT static calculations were performed with the Gaussian 09 quantum chemistry package.[84] In what follows, the stability of the calculated clusters is given in terms of relative Gibbs energy, the most stable structure being taken as the zero of the scale.

An important factor in the cluster-in-a-liquid description of solvation is the way one defines the relative contribution of each conformation and/or size of complex to the solution properties. Due to the limited size of the cluster, the solvent density is far from that existing in the bulk, which may introduce artefacts, in particular in the treatment of dispersion. Among the different approaches used for this purpose, mentioned in the introduction, we chose to describe the system in terms of contributions of clusters of a single size (0, 1, or 2 DMSO molecules), weighted with the Boltzmann weights obtained from non-dispersion-corrected Gibbs energies. Recent VCD studies of the solvation of an aromatic carboxylic acid by DMSO indeed suggested that weights obtained this way are well adapted to the description of the solute-DMSO solvent interaction within the cluster-in-a-liquid model.[29] Unless specified otherwise, the values given in the text are therefore non-dispersion corrected Gibbs energies. However, alternative approaches are used by other groups, such as calculating Boltzmann weight using $\Delta H°$.[85] Therefore, both relative Gibbs energies obtained at the B3LYP and the B3LYP-D3BJ levels are listed in Table 1. Comparison between electronic energies including or not ZPE corrections at 0K and Gibbs energies at room temperature are given as supplementary material, as well as the B3LYP-D3BJ structures are given as supplementary material.

2.2. FFMD Approaches

Polarisable MD simulations resting on the AMOEBA force field were carried out to model *trans*-AI in interaction with one, two and five DMSO molecules. The temperature was set to 300 K or 150 K and was controlled using the Nosé-Hoover thermostat.[86] The trajectories at 150 K were performed to constrain the dynamics of the system to remain around a selected local potential well of *trans*-AI. At 300 K, the duration of the trajectories was 1 ns and a time step of 0.5 fs was used. For the clusters with 2 and 5 DMSO molecules, a spherical van der Waals potential of 17 Å of diameter was used to avoid evaporation. At 150 K, the duration of the simulations was increased up to 3 ns to improve potential energy surface exploration but a time step of 1 fs was used to reduce the computation time. MD simulations of bulk systems were carried out using a cubic DMSO box with an edge length of 19.20 Å that contains 56 DMSO molecules. This box was pre-equilibrated before soaking the *trans*-AI solute. Simulations were performed at constant volume and 300 K using periodic boundary conditions, the Berendsen thermostat, and a 0.5 fs time step. Particle-mesh Ewald summation was used for the long-range electrostatic interactions. The Ewald real-space

cutoff was 7 Å, the van der Waals cutoff was 7 Å. Several simulations were carried out for a total simulation time of 4 ns. All simulations were performed with the Tinker program.[87]
To characterise the position of the DMSO molecules with respect to *trans*-AI, 2D contour plots were drawn from the distances between the nitrogen or oxygen of *trans*-AI and the oxygen of DMSO, throughout the trajectory using the Plotly library in the R software.[88, 89]

2.3. FPMD Approaches

For isolated *trans*-AI, the 1:1 as well as the 1:2 clusters with DMSO, one, three, and one starting structures were created, respectively, based on the optimised geometries found with the static calculations. FPMD calculations based on DFT were carried out with the Quickstep module of the CP2K software package.[90, 91] The simulations were of Born-Oppenheimer type with a time step of 0.5 fs, using the BLYP exchange-correlation functional,[92, 93] Grimme's dispersion correction (D3),[94] GTH pseudopotentials,[95-97] and the Gaussian and plane wave basis TZVP-MOLOPT-GTH with an energy cutoff of 400 Ry.[98] All simulations were performed in the canonical ensemble using the CSVR thermostat[99] at a slightly elevated temperature of 340 K to counterbalance the underestimation of temperature by the chosen functional.[100] However, the effect of this temperature on the intermolecular sampling is marginal as shown in Figure S1. The FPMD simulations were carried out in a vacuum supercell, that is each setup, either the isolated molecule or the clusters (1:1 or 1:2) with DMSO, was placed in a box of pre-defined size: 16^3 Å^3 and 20^3 Å^3 for isolated molecule and clusters, respectively. Each sample underwent geometry optimisation followed by a 5 ps equilibration performed under massive thermostatting with a coupling constant of 10 - 500 fs. The production trajectory of 30 ps was carried out under global thermostatting with a coupling constant of 500 fs. For VCD sampling of isolated *trans*-AI and the 1:1 clusters,[61] NVPT calculations were carried out with the CPMD software,[67, 101] sampling over the previously created FPMD trajectory, using Troullier-Martins pseudopotentials[102] and the plane wave basis with a cutoff of 70 Ry. The trajectory was sampled at every eighth step to form the correlation function, resulting in a time resolution of 4 fs. To generate IR and VCD spectra, the distributed molecular origin gauge[61] was imposed and FT-TCF post-processing was carried out following the equations

$$A(\omega) \propto \int_{\infty}^{-\infty} dt\, e^{-i\omega t}\, \langle \dot{\mu}\,(0) \cdot \dot{\mu}(t) \rangle \qquad (1)$$

$$\Delta A(\omega) \propto \int_{\infty}^{-\infty} dt\, e^{-i\omega t}\, \langle \dot{\mu}\,(0) \cdot m(t) \rangle \qquad (2)$$

and using the ChirPy python package, as available on GitLab.[103] The local IR and VCD signatures of the solvent were subtracted from the global spectrum before presentation. Further analysis and data visualisation was realised with ChirPy, NumPy,[104] and Matplotlib.[105]

For better comparison with static calculations, another set of MD trajectories was created using the B3LYP functional with dispersion correction (D3), at a temperature of 320K. From preceding BLYP runs, two starting points were chosen for the isolated molecule and the 1:1 cluster, respectively. The response function used for the generation of the IR and VCD spectra was calculated at BLYP level.

2.4. Nomenclature

Three parameters are important for the description of the *trans*-AI geometry. The first parameter is related to the alicyclic ring puckering motion. Due to the stereochemistry of *trans*-AI, the substituents are both either in axial or equatorial positions, resulting in two kinds of geometry: axial (dihedral angle $C_9C_4C_3C_2$ <0) denoted hereafter **ax**, or equatorial denoted **eq** (dihedral angle $C_9C_4C_3C_2$ >0). The two other parameters are the rotation of the OH and the NH_2 groups. The orientation of the OH group (see Scheme 1) will be denoted by g^+, g^- and t when the HC_2OH dihedral angle is close to 60°, -60° and 180°, respectively. For the orientation of the NH_2 group, we will consider the position of the lone pair (*lp*) relative to C_1H; it will be denoted G^+, G^- and T when the HC_1N*lp* dihedral angle is around 60°, -60° and 180°, respectively.

Results and Discussion

1. Molecular Structure and Solvation Network

1.1. Static DFT Calculations

1.2.a. Monomer

The most stable conformers of *trans*-AI in DMSO (continuum model) are shown in Figure 1. Conformer-selective IR spectroscopy of *trans*-AI under jet-cooled conditions found evidence for one conformer only, assigned to **eq**g^+G^+.[106, 107] It is also the most stable conformer within the calculations in implicit DMSO solvent and has the two functional groups on the same side of the alicyclic ring (OH…N distance of 2.96 Å). It should be noted that this proximity is only possible for **eq**g^+G^+. Below 4.3 kJ/mol, there are only **eq**G^+ conformers in DMSO, with various orientations of the OH group, and one **eq**g^+G^- conformer slightly destabilised by 1.9 kJ/mol relative to the most stable geometry. In contrast, the energy of any **ax** conformers is more than 5.6 kJ/mol above that of the **eq**g^+G^+. The structures are the same whether including dispersion correction or not.

1.2.b. 1:1 Complexes

The structures resulting from the exploration of the potential energy surface can be classified into three families, namely, bidentate, monodentate and “non-hydrogen-bonded” complexes. Each family encompasses several geometries of *trans*-AI that will be included in the Boltzmann-averaged contribution to simulate the IR absorption and VCD spectra.

However, for the sake of clarity, we describe here only the most stable structure defined by each possible family/*trans*-AI geometry combination. The most relevant complexes are shown in Figure 2 and their relative energies are given in Table 1. The complete set of low-energy 1:1 complexes are shown in Figure S2 in the electronic supplementary information (ESI).

In the **bidentate** family, the DMSO oxygen atom interacts with both OH and NH_2 groups of *trans*-AI. All the identified bidentate complexes contain the **eq**g^+ geometry. Formation of this type of complex with the **ax** geometry is prevented by steric constraints. Two sub-families can be identified, based either on the **eq**g^+G^- or the **eq**g^+T geometry of *trans*-AI, for which the hydrogen atoms of the protic groups are in a favourable position for interacting with the DMSO. This allows binding of one or the other hydrogen atom of the NH_2 group. These sub-families will be referred to as bi-**eq**g^+G^- and bi-**eq**g^+T. The bidentate complex with **eq**g^+G^- is the most stable of all identified complexes, whereas **eq**g^+G^- is not the most stable monomer conformer, because its solvation energy is favourable. In contrast, the formation of a bidentate complex with the most stable form of *trans*-AI, **eq**g^+G^+ is not possible because none of the hydrogen atoms of the amino group is in the appropriate position.

In the **monodentate** family, the DMSO oxygen interacts with either the OH or the NH_2 group of *trans*-AI. In principle, this kind of structure with a single hydrogen bond to DMSO can be obtained with any geometry of the bare molecule. Not surprisingly, the most stable group is that with the OH acting as a hydrogen bond donor to the DMSO oxygen. In what follows, these complexes will be called $mono_{OH}$ followed by the *trans*-AI geometry, *i.e.* $mono_{OH}$-**eq**g^-G^+, *etc*. Such structures are obtained in particular for the **eq**g^-G^+, **eq**g^-T, **ax**g^+T, and **eq**g^-G geometries of the monomer, in which the hydroxyl hydrogen atom involved in the interaction points outwards. Therefore, although the DMSO molecule is hydrogen bonded, it is located away from the molecule and does not interact with the aromatic ring. For $mono_{OH}$-**eq** tG^- and $mono_{OH}$-**eq**tT, there is also very limited interaction between DMSO and the aromatic ring of *trans*-AI**.** For $mono_{OH}$-**ax**tG^- or $mono_{OH}$-**ax**tT, the OH group points inwards, so that the hydrogen-bonded DMSO also interacts with the aromatic ring. $mono_{OH}$-**eq**g^+G^+ stands out by the fact that this complex is built from the most stable *trans*-AI geometry. However, the formation of an intermolecular hydrogen bond results in an increase of the OH…N distance from 2.96 in the monomer to 3.14 Å in the 1:1 complex.

The second sub-family is based on a single bond from one NH_2 hydrogen atom to the DMSO oxygen atom. The two identified complexes will be named $mono_{NH}$-**eq**g^+G^+ and $mono_{NH}$-**eq**tT, according to the geometries of *trans*-AI they contain. These geometries are adequate for monodentate complex formation because one of the NH_2 hydrogen atoms points outwards. Their relative energies relative to the most stable bidentate complex are larger than more than 10 kJ/mol. There is almost no interaction between DMSO and the *trans*-AI aromatic ring in these complexes.

The last family is called **non-hydrogen bonded** and contains structures lacking any intermolecular hydrogen bond. One such structure is found based on the most stable

conformation of the bare molecule, **eq**g^+G^+ and will be named noHB-**eq**g^+G^+ in what follows. Its very limited solvation energy is counterbalanced by the intramolecular energy and it is destabilised by only 3.4 kJ/mol relative to the most stable bidentate complex. Very surprisingly, it is more stable than mono$_{NH}$-**eq**g^+G^+, which also contains **eq**g^+G^+. However, this type of structure with a loose DMSO molecule is not likely to be realistic and may result from the lack of dispersion correction in the calculation, from misestimating the thermal corrections, or from the predominant role of the intramolecular energy in the 1:1 complex.

Such a possible influence of the theory level is emphasised by noting that only hydrogen-bonded structures are obtained below 15 kJ/mol when dispersion correction is included. For the bidentate structures, the orientation of DMSO is dictated by the hydrogen bond network and the geometry is very similar with and without dispersion correction. For the monodentate structures, the *trans*-AI geometry and the hydrogen bonds are most of the time independent of whether dispersion correction is included or not. However, the DMSO molecule can rearrange to optimise dispersion, keeping the hydrogen bond, which results in changes in energy ordering. No variation of the intramolecular geometry is observed in mono$_{OH}$-**ax**tG^-, with relative Gibbs energy of 2.4 kJ/mol with dispersion corrections *vs.* 13.8 kJ/mol without. This difference is due to rotation of the DMSO methyl groups to optimise the interaction with the aromatic ring (Figure S3a in the ESI). In contrast, the intermolecular geometry of noHB-**eq**g^+G^+ is strongly modified upon inclusion of dispersion, with the DMSO becoming much closer to the aromatic ring, as shown in Figure S3b. Its relative energy raises upon inclusion of dispersion, likely due to concomitant increase of repulsion.

Comparison between the Gibbs energies at room temperature ΔG, the electronic energies E_0, and the ZPE-corrected electronic energies at 0 K E_0 +ZPE, allows assessing the respective role of ZPE and thermal corrections on the Boltzmann factors (see Table S1). Examination of Table S1 confirms the non-physical character of noHB-**eq**g^+G^+, whose relative electronic energy is >20 kJ/mol. noHB-**eq**g^+G^+ is indeed strongly favoured by thermal corrections due to its loose character. For the other 1:1 complexes, Table S1 indicates that ZPE corrections amount to a maximum of 1.8 kJ/mol. The major difference between energies at 0 K and Gibbs energies at room temperature arises from thermal corrections, which are, as expected, more stabilising for looser complexes, like mono$_{NH}$-**eq**g^+G^+. As expected, considering either E_0 or E_0 + ZPE results in a dominant contribution of the most stable rigid bi-**eq**g^+G^- structure, while including thermal corrections increases the contribution of looser complexes. These results underline the role of entropy and the difficulties of assessing the weight of the different complexes.

1.2.c. 1:2 Complexes

The most relevant 1:2 complexes are shown in Figure 3; the other low-energy complexes are shown in Figure S4 in the ESI. Their relative Gibbs energies are listed in Table 1.

The most stable 1:2 complexes have structures similar to those obtained for the 1:1 complexes, in which the second DMSO molecule does not hydrogen bond directly to *trans*-

AI. Instead, it adds to one of the bidentate or monodentate 1:1 complexes evidenced previously, whose geometry is not modified in a substantial manner. The nomenclature for these complexes rests on that used for the 1:1 complexes, preceded by "dmso": dmso-bi-**eq**g^+G^-, dmso-mono$_{OH}$-**eq**g^-G^+, *etc*. Among them, dmso-bi-**eq**g^+G^-, built on the most stable bidentate 1:1 complex, remains the most stable structure; it is the most stable 1:2 complex.

All the 1:2 complexes involving a single OH…O hydrogen bond are built on the mono$_{OH}$ 1:1 complexes found previously. They are higher in energy than dmso-bi-**eq**g^+G^- by 1 to 7 kJ/mol. Only one complex involving a NH…O interaction is found below 20 kJ/mol, namely dmso-mono$_{NH}$-**eq**g^+G^+.

In addition to the structures built on 1:1 complexes, three complexes display two hydrogen bonds, namely two between the OH group and one of the two DMSO molecules, and one between the NH_2 group and the second DMSO molecule. They will be called mono$_{NH}$-mono$_{OH}$-**ax**tT, mono$_{OH}$-mono$_{NH}$-**eq**tG^+, and mono$_{NH}$-mono$_{OH}$-**eq**g^+G^-, respectively. They are higher in energy than dmso-bi-**eq**g^+G^- by 8 to 16 kJ/mol.

Although the hydrogen bond pattern does not change upon inclusion of a second DMSO molecule, the energy ordering is strongly modified compared to the 1:1 complexes. Moreover, unlike in 1:1 complexes, inclusion of dispersion correction strongly modifies the calculated structures. They become much more compact, as shown in Figures S3c and d. For the most stable 1:2 complex, dmso-bi-**eq**g^+G, the position of the second DMSO molecule changes remarkably. A compact solvated structure is obtained when dispersion correction is included, with the two DMSO molecules close to each other and having their electric dipoles oriented in an antiparallel manner. This corresponds to the onset of a second solvation shell.

The non hydrogen-bonded DMSO molecule (Figure S3 d) of dmso-monoOH-**eq**g^-G^- gets much closer to the aromatic ring when including dispersion correction. It is therefore in the first solvation shell of *trans*-AI with which it is in direct interaction. Also the complexes not derived from 1:1 complexes undergo significant modification when including dispersion correction. For example, mono$_{OH}$-mono$_{NH}$-**eq**g^+G^- undergoes strong stabilisation (Gibbs energy going from 8.3 to 2.7 kJ/mol), due to structural modification, see Figure S3e. While the positions of *trans*-AI and the OH-bonded DMSO are identical, the DMSO molecule interacting with the NH_2 group is strongly shifted towards the aromatic ring. The effect of thermal corrections is similar to that described for the 1:1 complexes and we shall not comment them here.

It can be concluded from the static calculations that solvation by one DMSO molecule hardly modifies the structure of *trans*-AI but changes the conformers relative energies. Solvation destabilises the **eq**g^+G^+ conformer in favour of the **eq**g^+G^-, because a stable bidentate complex cannot be formed with **eq**g^+G^+. Adding another DMSO molecule does not perturb the structure of *trans*-AI neither that of the 1:1 complexes; the most stable 1:2 complexes are built on stable 1:1 complexes to which one DMSO molecule is added. Additional structures with two hydrogen-bonded DMSO molecules also appear among the most stable calculated ones. A static picture of solvation accounts for the solute structure and the main

interactions present in solution. The energy ordering between the families corresponding to different hydrogen bond patterns changes upon addition of a second DMSO molecule or by inclusion of dispersion correction, due to a change in the solvent position. It also changes upon inclusion or not of thermal corrections. This observation points to the difficulty of defining the relative contribution of the different structures by their Boltzmann weights and the intrinsic limitation of describing solvation by a finite number of molecules. Moreover, many almost isoenergetic configurations are found for each cluster family described above. This body of results points to the fluxionality of the solvent studied below more in detail using molecular dynamics.

1.2. Dynamical Solvent Organization around *trans*-AI by MD Simulations

MD simulations performed with the AMOEBA force field have two main purposes, namely to connect with the structural set underlying the cluster-in-a-liquid approach, and second to assess the sampling of a few solvent molecules around *trans*-AI achieved with FPMD at 340 K.

1.2.a. FFMD Simulations at 150 K

A local and detailed description of the conformational landscape around a given structure in a potential depth was obtained using 150 K FFMD trajectories. To this end, we resort to the 2D contour plots drawn from the distances between the nitrogen or oxygen atom of *trans*-AI and the oxygen atom of a DMSO molecule. The distance between the nitrogen atom of *trans*-AI and the oxygen atom of the DMSO is represented on the y-axis and the distance between the oxygen atom of *trans*-AI and the oxygen atom of the DMSO is on the x-axis. Then, the z-axis represents the occurrence of the distance observed over the trajectory in a false colour scale. Each graph is divided in four regions: i) a spot at (x≃3 Å,y≃3 Å) corresponds to the bidentate family; ii) a spot around (x≃3 Å,y≃5 Å) represents a monodentate configuration with the DMSO hydrogen-bonded to the OH group of the *trans*-AI; iii) a spot at (x≃4 Å,y≃3 Å) corresponds to a monodentate configuration with the DMSO hydrogen bonded to the NH_2 group of the *trans*-AI; iv) finally, for a spot at (≥ 5 Å,y≃8 Å), the DMSO is in the vicinity of *trans*-AI without being hydrogen bonded.

At 150 K, the temperature is low enough to catch the *trans*-AI molecule in individual potential wells, allowing the stability of its different conformations to be explored. Four trajectories were carried out for *trans*-AI in the equatorial position: one for the 1:1 bidendate conformation that corresponds to the most stable structure in the static sampling, two for the 1:2 complex (one with one bidentate DMSO and the second DMSO in the vicinity, one as mono-OH and mono-NH conformation) and one with five DMSO molecules spread around *trans*-AI with two molecules hydrogen-bonded to the OH and NH_2 groups. With the *trans*-AI in the axial position, six trajectories were simulated, including two with one DMSO (mono-OH and mono-NH, respectively), three with two DMSO (mono-OH and mono-NH, mono-OH and in the vicinity, mono-NH and in the vicinity, respectively) and one

with five DMSO (with similar configuration of the solvent molecules to that of the equatorial system).

The contour plots (Figures S5 and S6 in the ESI) show that the complexes remain near their starting points, keeping the starting *trans-AI* conformation (equatorial or axial) and the hydrogen-bond interactions with one or two DMSO molecules. In Figure 4, the contour plots of the 1:5 equatorial complex show the five DMSO molecules staying mainly around their initial position. One molecule stays hydrogen bonded to the OH group, two are in interaction with the NH_2 group with one closer to the OH, and the remaining two molecules are in the vicinity of the *trans*-AI solute. These results are consistent with the main intermolecular interactions expected in this complex, *i.e.* OH…O and NH…O hydrogen bonds, and solvent-solvent interactions. The dynamics of the third DMSO molecule is more diverse. Lastly, the bidentate conformation is present but in a smaller amount and does not correspond to a well-defined spot in Figure 4. Still, the superimposed contour plots for the five DMSO molecules show that all the expected interaction sites of *trans-AI* (protic groups and aromatic ring) interact with DMSO, which demonstrates that 1:5 complexes give a realistic description of solvation.

One exception is the simulation starting from *trans*-AI in the equatorial position with one DMSO molecule. In this trajectory, the *trans*-AI switches to the axial position after 1.5 ns and remains axial during the rest of the simulation. This result contrasts with the static DFT results, which give only a few stable structures in axial configuration for the 1:1 complex. It could be explained by an artefact of the force field, which may excessively favour the axial configuration.

1.2.b. FFMD Simulations at 300 K

300 K FFMD simulations were performed to assess the sampling of a few solvent molecules around *trans*-AI achieved with FPMD at 340 K. At 300 K, the FFMD trajectories are less impacted by the starting configuration. Four trajectories were produced, namely two with one DMSO, and two with either two or five DMSO molecules, respectively. For the system with one DMSO molecule, the starting points had the solvent hydrogen-bonded to the OH and NH_2 groups, respectively. Yet, the two trajectories give similar contour plots (see Figure 5). The system with one solvent molecule on the OH and one on the NH_2 group was used as a starting point for the 1:2 complex. Lastly, the system with five DMSO had one molecule hydrogen bonded to the oxygen atom, one in interaction with the NH_2 group and the rest in the vicinity of the *trans*-AI solute. The dihedral angle of the 5-membered ring was monitored to characterise the equatorial and axial conformations (see Figure 6). In the FFMD trajectory for the 1:1 complex, the *trans*-AI spends most the time in axial position (93%), however the results are more balanced for the system with two (80%) and even more five DMSO (45%).

The time evolution of the dihedral angles of the hydrogen bonds of the oxygen and the nitrogen atoms of the *trans*-AI are shown in Figure S7 in the ESI. For all three FFMD trajectories, the dihedral angle HC_2OH stays mainly at around 180° or 60° (t and g^+

conformations), which is consistent with the static picture at the DFT level. The residence time of one value of the dihedral angle increases with the number of DMSO molecules as a confirmation of the stability of these t and g^+ conformations in solution.

At 300 K, the DMSO molecules are much more mobile than at 150 K, moving around the oxygen and the nitrogen atoms of the *trans*-AI. Figure S8 in the ESI provides the distances between the OH and NH_2 groups of *trans*-AI and DMSO as a function of time, showing the fluctuations of the two DMSO molecules moving close, then far away from the *trans*-AI. Exchanges between the two molecules happen several times on the NH_2 group with an approximate residence time, *i.e*. lifetime of the hydrogen bond, between 100 and 200 ps, as shown in Figure S8 in the ESI. Nevertheless, the contour plots indicate that there is a tendency of each DMSO molecule to remain in certain spots throughout the simulation. In the 1:1 complex (Figure 5), the DMSO takes mainly three positions, near the OH and NH_2 groups of the *trans*-AI (red spot) and non-hydrogen bonded to the *trans*-AI. Contrary to the static DFT calculations, the bidentate interaction is not favoured because *trans*-AI is mainly in the axial position during the simulation. With two DMSO molecules as explicit solvent, one DMSO stays mainly hydrogen-bonded to the OH or the NH_2 group of the *trans*-AI (Figure 7, right). The second DMSO molecule takes mainly three positions, hydrogen-bonded to the OH or the NH_2 group of the *trans*-AI and non-hydrogen bonded in the OH region (Figure 7, left).

FFMD simulations were also performed with five DMSO molecules as explicit solvent. The solvent moves around the *trans*-AI with an extensive sampling for all the molecules. The contour plots show similar preferred spots as for the 1:1 and 1:2 complexes: near the OH and NH_2 groups of the *trans*-AI and non-hydrogen bonded in the OH region (Figure S9 in the ESI). The bidentate interaction for one DMSO molecule is also found and can be correlated to a larger proportion of the equatorial conformation. Figure S10 in the ESI provides the evolution with time of the O(*trans*-AI)-O(DMSO) and N(*trans*-AI)-O(DSMO) distances, as well as the radial distribution functions. Two main observations can be derived from these graphs: i) at short distances, one DMSO molecule is always hydrogen-bonded to the OH and NH_2 groups, although this interaction is dynamical with frequent exchanges, leading to a residence time of *ca.* 70 ps and 100 ps, respectively; ii) the other molecules (three or four, depending on the observed conformation) solvate the *trans*-AI solute and are also in its vicinity. These observation points to the mobility of the solvent at room temperature.

1.2.c. FPMD Simulations at 320 and 340 K

FPMD trajectories were out of reach for 1:5 complexes, but were obtained at 340 K for three 1:1 and one 1:2 complexes using the BLYP functional and at 320 K for two 1:1 complexes using the B3LYP functional. The starting 1:1 complex structures contained DMSO interacting either in a bidentate manner or via OH…O interaction. FPMD simulations also show a balance between equatorial and axial configurations. While the FPMD trajectories are much shorter and produce results that depend more strongly on the starting structure, their analy-

sis indicates that *trans*-AI spends 28% and 30% in the axial position for the 1:1 and 1:2 systems, respectively. Both series of results highlight an equilibrium between the axial and equatorial configurations of *trans*-AI in the presence of several solvent molecules even if the accurate proportions cannot be directly compared due to the differences in the computational protocols and inherent limitations of both approaches.

The results of the FPMD simulations for the 1:1 and 1:2 complexes parallel those of FFMD. They too highlight the mobility of the DMSO molecules (Figure 8). In the 1:1 complex, both the bidentate and monodentate (OH and NH_2) conformations are explored. The DMSO molecules remain close to the OH group of *trans*-AI, which parallels the conclusion of static calculations that the $mono_{OH}$ or bidentate complexes are the most stable structures. With two DMSO molecules, conformational exploration is very limited, as manifested by the contour plots of Figure 8 (bottom) where the simulation mainly explores one type of interaction with the DMSO. However, even under these shorter time scale the dynamics shows a partial mobility away from the initial conformation of the *trans*-AI (Figure 8, left graph).

The influence of the functional on the trajectories is illustrated by comparing the results described above with MD trajectories obtained with the B3LYP functional. Figure 8 should be modified and B3LYP contour plots should be added. For the isolated molecule, the B3LYP trajectories also evidence a balanced contribution of axial and equatorial configurations. Their main difference compared to BLYP results is an increased contribution of G^- structures. For the 1:1 complex also, a balanced contribution of axial and equatorial configurations is obtained, as well as a larger contribution of G^- structures relative to BLYP trajectories.

1.2.d. Towards Full Solvation with FFMD Simulations

FFMD simulations of fully solvated *trans-AI* were performed and the solvent organisation and the residence time of the DMSO molecules analysed. The radial distribution functions (RDF) between the *trans*-AI and the oxygen atoms of the DMSO molecules are shown in Figure S11 in the ESI. The RDFs between the oxygen and the nitrogen atoms of the *trans*-AI and the DMSO molecules show an intense peak at 2 Å and 3 Å, respectively. The integration of the radial distribution functions beyond the first peak gives an average of 1.0 and 1.7 DMSO around the OH and NH_2 groups, respectively. The second solvation shells are around 6 Å from the oxygen atom and 5.75 Å from the nitrogen atom of the *trans*-AI. The results confirm that in the bulk solvent, a DMSO molecule is always hydrogen-bonded to the OH of *trans*-AI with an approximate residence time of 160 ps. Conversely, less than two DMSO molecules are hydrogen-bonded to the NH_2 group due to the competition with solvent-solvent interactions, with a residence time, *i.e.* lifetime of the hydrogen bond, of 100 ps. As expected, the residence time increases from the 1:1 complex to the bulk.

2. Spectral Assignment by Comparison with the Experiment

2.1. Comparison with Static Calculations

2.1.a. Monomer

The IR absorption spectrum of *trans*-AI in DMSO is given in Figure 9a. It is compared to the weighted average of the monomer conformers according to the Boltzmann distribution at room temperature, given in Figure 9b. The simulated IR absorption spectrum in the fingerprint region shows discrepancies relative to the experiment. First, the band due to the NH_2 scissoring motion at ~1600 cm^{-1} does not reproduce the experimental bandwidth. The major difference appears in the 1210-1260 cm^{-1} region where coupled β(OH) bend and alicyclic β(CH) bend transitions are calculated but are not seen in the experimental spectrum.

The VCD spectrum of *(S,S)-trans*-AI is given in Figure 10a, together with the weighted average of the monomers conformers according to the same Boltzmann distribution at room temperature (Figure 10b). The calculated VCD spectrum satisfactorily reproduces the experiment below 1500 cm^{-1} with the main positive bands obtained at 1458 and 1386 cm^{-1} and the negative bands at 1235 and 1111 cm^{-1}. Nevertheless, it shows some discrepancy, in particular for the high-energy band at 1614 cm^{-1}, which is positive in the experimental spectrum but negative in the calculated spectrum, the breadth of the most intense feature at 1360-1400 cm^{-1} being not reproduced either.

The individual contributions of the conformers included in the average are shown in Figure S12 in the ESI. The NH_2 scissoring motion region at ~1600 cm^{-1} is characterised by an asymmetric bisignate VCD signal with a more intense positive component. Its sign is wrong for all the calculated conformers except for the G^- geometries in which none of the hydrogen atoms of the amino group point towards the hydroxyl group. For the G^- structures, the positive signal is due to pure NH_2 scissoring motion, while the negative signal corresponds to scissoring of the NH_2 group strongly coupled with the e_{2g} (8b) benzene ring motions. The doublet at ~1450 cm^{-1} contains the contribution of aromatic CH bends and is not very sensitive to conformation. This region could be explained in terms of contributions of all the calculated conformers, apart from the **ax** conformer, which could not contribute strongly to the spectrum. The region between 1450 and 1200 cm^{-1} corresponds to strongly delocalised modes, whose description, frequency, VCD sign and intensity, all depend on conformational changes. The VCD band calculated at ~1380 cm^{-1} contains contributions of NH_2 rocking motion, β(OH) and β(CH) bends, with relative amounts depending on the conformation. It is positive for all the conformers except **eq**g^+T. Another mode can be described in terms of contributions from the same displacement, at ~1230 cm^{-1}. It corresponds to a band that is strongly negative for all the **eq**g^+ conformers and slightly negative for the **eq**g^- conformers. Lastly, the $\nu(CC)_{arom}$ appears as a positive band at ~1180 cm^{-1} in all the **eq** conformers and is negative for the **ax** conformer. Regarding those results, it can be concluded that none of the conformers alone displays a spectrum matching very satisfactorily the experiment.

Despite satisfactory agreement in some spectral regions, both band positions and bandwidths in other regions point towards hydrogen bonding. The experimental IR spectrum

in the OH/NH stretching region (see Figure S13 in the ESI) displays a broad and intense band assigned to the hydrogen-bonded ν(OH). Although the weak ν(NH_2) bands do not appear clearly because the spectrum is dominated by the strong ν(OH) transition, one cannot exclude additional interaction between the solvent and the amino group. This observation justifies considering the 1:1 and 1:2 complexes in more details.

2.1.b. 1:1 complexes

The Boltzmann-averaged VCD spectra of the 1:1 complexes are compared to the experimental spectra in Figures 10 a) and c), respectively. The IR absorption spectrum is in much better agreement with the experimental results than that of the monomer, both in terms of band position and intensity. The calculated VCD spectrum is also in good agreement with the experimental spectrum, in particular in the 1350 cm^{-1} range. The calculated positive band at 1624 cm^{-1} corresponds to that at 1614 cm^{-1} in the experimental spectrum. The 1462 and 1443 cm^{-1} bands and the intense positive band at 1380 cm^{-1} with the shoulder at 1355 cm^{-1} are well reproduced. The negative features at 1217 and 1123 cm^{-1} have also their counterpart in the calculated spectrum. However, the band calculated at 1281 cm^{-1} has no equivalent in the experimental spectrum.

The individual spectra of the most relevant 1:1 complexes are shown in Figure 11 and that of all complexes in Figure S14 in the ESI. At first sight, inspection of Figure 11 indicates that the experimental spectrum can be accounted for by the contribution of a limited numbers of structures. The best match is obtained for $mono_{OH}$-**eq**g^-G^-, which is the second stable complex. All features except the positive band at ~1180 cm^{-1} find their counterpart in the experimental spectrum. The weak positive doublet calculated at 1447 / 1431 cm^{-1} corresponds to that observed at 1462 /1443 cm^{-1}. An intense band is predicted at 1384 cm^{-1}, which meets the position of the intense experimental feature. The negative band calculated at 1331 cm^{-1} may correspond to that observed at 1318 cm^{-1}. The last two calculated negative bands at 1216 and 1126 cm^{-1} are in particularly good agreement with those observed at 1217 and 1123 cm^{-1}. Other complexes show partial overlap with the experimental spectrum and can contribute to it. The VCD spectrum of the most stable bi-**eq**g^+G^- complex displays, like the **eq**g^+G^- monomer, the bisignate transition centred at 1616 cm^{-1}. Two positive features calculated at 1387 and 1345 cm^{-1} are close to the intense and broad experimental band at 1380 cm^{-1} and its shoulder at 1355 cm^{-1}. The negative signal at 1320 cm^{-1} is in good agreement with that observed at 1318 cm^{-1}. However, the negative features observed at 1228 and 1125 cm^{-1} are not predicted for the bidentate complex. They are due to coupled motions of the NH_2 and OH groups, which are strongly modified by the intermolecular hydrogen bond network and loose intensity in the complex. The $mono_{OH}$-**ax** complexes too have satisfactory overlap with the experimental spectrum; in particular, $mono_{OH}$-**ax**tG^- shows the bisignate signature of the G^- forms and it becomes one of the most stable complexes when dispersion correction is considered. The axial complexes all have negative signal in the

region of ~1200 cm^{-1}, as experimentally observed. While the experimentally observed features find at least partial counterpart in the spectra of the $mono_{OH}$ or the bidentate families, this is not the case for $mono_{NH}$, which lacks in particular the strong positive band between 1340 and 1390 cm^{-1}. This poor match, together with very high relative energy, allows discarding the $mono_{NH}$. bi-**eq**g^{+}T cannot contribute to the spectrum either due to the wrong signs of the bisignate at 1620 cm^{-1} and of all the bands between 1100 and 1250 cm^{-1}. Last, $mono_{OH}$ **eq**tG^{-} exhibits a wrong sign for the doublet at ~1450 cm^{-1}.

In summary for the 1:1 complexes, an important role of the solvent is to stabilise the structures selectively, therefore changing the contributions of the different conformers in the spectra. In this respect, it is worth noting that axial forms are stabilised in DMSO complexes, in particular $mono_{OH}$-**ax**tG^{-} can contribute to the final spectrum. Although almost no intramolecular structural modification happens upon solvation, hydrogen bonding induces some band shifting and changes in the intensity or sign of the VCD spectrum in the corresponding regions. This is especially significant in the region of the β(NH) and β(OH) modes, below 1400 cm^{-1}. Not all interaction types perturb the spectrum to the same extent. Non-hydrogen bonded complexes show the same spectrum as the corresponding monomer. bi-**eq**$g^{+}G^{-}$ shows strong modification of the VCD spectrum compared to **eq**$g^{+}G^{-}$ in the 1200 cm^{-1} region, as mentioned above. The interaction with NH_2 perturbs the spectrum much less than the interaction with OH does. The VCD spectrum of $mono_{NH}$-**eq**$g^{+}G^{+}$ is almost identical to that of the corresponding monomer, while that of $mono_{OH}$ complexes is more affected by solvation.

2.1.c. 1:2 complexes

The Boltzmann-averaged VCD spectra for the 1:2 complexes are compared to the experimental spectra in Figures 10a and d. They are in good agreement with the experimental data, as was the case for the 1:1 complexes. The individual contributions are given in Figure S15. The spectrum of the complexes built by adding one DMSO molecule to a 1:1 complex is identical to that of the 1:1 complexes they are associated with. Therefore, similar conclusions can be drawn for the 1:1 and 1:2 complexes: the 1:2 complexes involving an NH…O interaction with the solvent do not match the experimental VCD spectrum. Several $mono_{OH}$ complexes are stabilised in the 1:2 complexes relative to the 1:1 complex, like dmso-$mono_{OH}$-**eq**tT, the spectrum of which displays an intense positive feature located at 1260 cm^{-1} in the region where only small negative transitions are expected. The spectrum of the $mono_{OH}$-$mono_{NH}$ complexes does not fit the experiment either. Despite being built on a G^{-} conformer, $mono_{OH}$-$mono_{NH}$-**eq**$g^{+}G^{-}$ does not account for the bidentate signal but shows a single intense band located at 1664 cm^{-1}, too high in frequency compared with the observed band. Moreover, it displays a strong negative band at 1435 cm^{-1}, where the small positive doublet is observed.

To summarise, the 1:2 complexes not derived from 1:1 complexes do not match the experimental spectra. Those built on a 1:1 complex, where the second DMSO molecule

simply sticks to the preformed 1:1 complex without additional hydrogen bond, show a VCD spectrum identical to that of the parent 1:1 complex. Addition of a second DMSO molecule only modifies the energy ordering. This result suggests that considering 1:1 complexes is a good approximation for explaining the experimental results from a static point of view. A more refined picture of solvation demands addition of more solvent molecules but considering the 1:2 complex is not enough to reproduce the whole solvation process.

2.2. Comparison with VCD spectra from FPMD simulations

2.2.a. Sampling and spectral assignment

The individual results of the FPMD-NVPT samples are shown in Figure 12. To allow comparison with the cluster set from the static calculations, each MD trajectory is endowed with a sampling clock that characterises the phase space visited by the system, according to the nomenclature introduced in section 2.4. The spectra obtained for isolated *trans*-AI in the gas phase (no solvent) are also shown for comparison. We shall first discuss the results resting on BLYP trajectories, then shortly mention the differences when using B3LYP trajectories. Differences between FPMD spectra at the BLYP and B3LYP levels, and their static counterparts, may result from structural effect (different relative contributions of different geometries) or differences in frequencies, due to the differing underlying functionals (B3LYP for static calculations, BLYP or B3LYP in the case of FPMD).

Monomer

The FPMD VCD spectra have in common with their static counterpart and the experiment the main positive signal at ~1400 cm^{-1}, and a weaker positive band at ~ 1440 cm^{-1}. Yet, the region below 1300 cm^{-1} and the NH_2 scissor vibration at ~1570 cm^{-1} (1600 cm^{-1} in the experimental spectrum) differs from the static calculations and shows more variability between individuals trajectories (see Figure S16).

The averaged BLYP FPMD trajectory for the *trans*-AI monomer mainly visits two conformers, **eq**g^+G^+ and **eq**tG^-, whereas only the first clearly agrees with the static results, while the latter does not seem energetically favourable in the static picture. Moreover, **eq**$g^+G^-$and **eq**g^-G^+ - ranked second and third in their static energies - play a minor part in the FPMD trajectory. Results based on B3LYP trajectories also show important contribution of **eq**g^+G^+, the most stable form in static calculations. However, they show much larger contributions of **eq**g^-G^- and **eq**g^+G^- than those based on BLYP. The fact that they favours G^- conformers over G^+ already for the isolated molecule underlines the importance of entropy: even in the absence of the solvent, the energetic preference for **eq**g^+G^+ loses its importance at room temperature. This is in line with the static calculations showing that **eq**g^+G^- is stabilised upon inclusion of thermal corrections.

With respect to experiment, the FPMD spectrum based on BLYP trajectories exhibits similar shortcomings as its static counterpart, showing the additionally predicted IR and VCD peak at 1250 cm^{-1}, while missing the negative VCD signal at 1125 cm^{-1}. However, it contains the main signal at ~1350 cm^{-1} together with its bandwidth. The agreement between experiment and the FPMD spectrum based on B3LYP trajectories is by far better. This improvement is largely due to the increased contribution of G^- conformers, which are responsible for the bisignate doublet experimentally observed at ~1600 cm^{-1}. The improvement is also due to better frequencies calculations, in particular the NH_2 rocking motion at 1380 cm^{-1}.

1:1 complex

As explained below in more details, the position of DMSO is important for conformer stability, but does not actively take part in the local VCD generation, which is why Figure 12 focuses on the conformer space of *trans*-AI alone (the individual contributions of the FPMD trajectories is shown in Figure S16 in the ESI). The correlation between DMSO position and the solute conformers was analysed in section 1.3, where it was found that in the FPMD trajectories of the 1:1 cluster the DMSO molecule remains most of the time with the OH group of *trans*-AI. The phase space sampling of the solute (Figure 12; right) indicates that the OH conformation (*i.e.*, g^+, g^-, t) shows a broader diversity compared to that of NH_2 (i.e., G^+, G^-, T) owing to the varying position of the solvent molecule, for BLYP and B3LYP results alike. Generally, there is a strong bias for the G^- conformer in the FPMD results; the amino group remains oriented towards the main occupation zone of the single DMSO molecule. As indicated in section 1.3, FPMD finds the ratio of axial conformation much higher than the Boltzmann weight of the static cluster predicts, which might contribute to the two negative VCD bands near 1200 cm^{-1}.

Apparent is the absence of **eq**g^-G^+, which seems very unfavourable in the FPMD simulation for both BLYP and B3LYP trajectories. This may be the underlying reason for the underestimation of the region 1100–1250 cm^{-1} in the VCD with respect to experiment, in view of the individual contribution of that **eq**g^-G^+ cluster (see Figure S14). The remaining parts of the measured VCD spectrum appear well reproduced by FPMD. The IR spectrum shows some discrepancies in the intensities for the NH_2 rocking/β(OH)/β(CH) region, which could be an artefact of the missing implicit or explicit solvation of the overall system.

2.2. Locality of vibrational absorption and VCD and the role of the solvent

Figure 13 shows the spatially resolved IR and VCD spectra derived from FPMD-NVPT calculations based on the B3LYP trajectories, using FT-TCF and a radial cutoff function as it has been used and introduced in earlier studies.[38, 61, 69] The results are qualitatively identical for those based on the BLYP trajectories. Therein, the electric and magnetic dipole moments entering the TCF, Equation [2], are evaluated according to their originating position in space. In Figure 13a, the radius *r* corresponds to the distance from the centre of mass of *trans*-AI as it is scanned starting from 0, where only local contributions to the IR/VCD spectrum appear, *i.e.*

those contributions that stem from trans-AI itself. By moving farther from the chiral solute, non-local signatures can be captured; for *trans*-AI this corresponds to contributions stemming from DMSO. Hence, if a non-local signal can be found, the polarisation of the centre (trans-AI) is coupled with the environment (DMSO), which can be due to either coupled oscillation, or induced polarisation.[108] It becomes evident that the spatial region above 2 Å, denoted as solvent effect, hardly returns any signal that may account for intermolecular IR/VCD; the main contributions are indicated as being of local origin, that is, stemming from *trans*-AI itself. Consequently, DMSO as a solvent does not significantly contributes to neither the IR nor the VCD spectrum of *trans*-AI in solution. There are negligible traces at frequencies at about 1360 cm^{-1} and 1400 cm^{-1} in the IR spectrum and, additionally, at 1425 cm^{-1} in the VCD spectrum, where the large dipole moment of the solvent slightly polarises the solute. This does not come as a surprise since these frequencies correspond to modes localised at the OH and NH_2 groups, respectively, which bind the molecule to the solvent. Although in these regions the non-local contributions sum up to 20%, they do not change qualitatively or quantitatively the overall shape of the spectra (Figure 13b). Consequently, what matters is the conformational sampling of *trans*-AI and this is where the solvent effect is discernible.
Such a conclusion is important as it means that in the case of DMSO electrostatics do barely influence the VCD response of *trans*-AI directly. This is at least true for the case of a 1:1 cluster of *trans*-AI and DMSO, on which the FPMD results rely. Yet, as the solvation study in part 2.1 suggests, it can be assumed that this small-scale image generalises to full solvation. The number of solvent molecules surrounding the solute may ultimately lead to a rise of the small contributions shown in Figure 13, but since the nature of the interaction does not change significantly with the number of DMSO molecules, it is unlikely to alter the locality of the VCD spectrum. Consequently, only the electronic wave function of the chiral solute is needed, which encourages theoretical models like QM/MM approaches for IR and VCD determination.

3. Discussion

3.1. Molecular Structure and Solvation Network

Generally, static calculations of the 1:1 and 1:2 complexes do not yield *trans*-AI conformers very different from the minima found on the potential energy surface of the monomer. However, DMSO stabilises higher-energy structures of *trans*-AI. The low-energy structures obtained by static calculations can be first compared to the distribution obtained with FFMD or FPMD trajectories. A static view of the thermal distribution between axial and equatorial conformers, based on Boltzmann populations, would yield a vanishing contribution of the axial conformers to the spectra. In contrast, FPMD calculations indicate that axial conformations are visited slightly for the bare molecule and more for the 1:1 complex. The axial/equatorial ratio is overestimated by the FFMD simulations of the monomer. However, the values obtained for larger clusters match the FPMD results and point towards a balanced

contribution of axial and equatorial configurations. Lastly, the FPMD trajectories, despite being only 30 ps long, explore both axial and equatorial conformations, which point towards the flexibility of the molecule along the puckering coordinate.

In terms of OH and NH_2 orientations, static calculations and FPMD trajectories indicate a lesser contribution of g^+G^+ in the 1:1 complex relative to the monomer. The larger weight of G^- in the 1:1 complex is corroborated by an increased contribution of G^- in the FFMD calculations with increasing cluster size. This observation indicates a good complementarity of FPMD at small sizes with FFMD, on the way towards describing full solvation. Temperature manifests itself in the greater contribution of t conformers. FPMD trajectories give a noticeable fraction of tG^- structures in the monomer, as well as in the 1:1 complex. This *trans* orientation of the OH group is also found in FFMD calculations, for all the clusters sizes. It appears among the most stable structures in the static calculations for the 1:1 complex. It seems therefore that, in static calculations, a solvent molecule is required to stabilise a geometry frequently visited in the FPMD trajectories already for the monomer.

The intramolecular structure and the solvation network are strongly interconnected. In terms of OH orientation, g^+ is favoured by all the methods used because it allows concomitant interactions of DMSO with OH and NH_2 groups. This interaction manifests itself by the presence of a very stable bidentate complex in the static calculations and by spots located near NH or OH in the contour plots obtained from FFMD trajectories. Still, the frozen view of a bidentate complex is replaced by a fluxional description in MD simulations, where the DMSO alternatively interacts with the two substituents. This explains the smaller contribution of g^+G^- conformers in the MD simulations than would be expected based on static calculations, despite these conformers contributing importantly. Structures involving g^- conformations are also among the most stable ones. They involve a less stiff hydrogen bond network, with a single floppy OH…DMSO interaction. However, t and g^+ are favoured in both MD simulations over g^-, which does not even appear in FFMD. Both FPMD and FFMD results also show a decreasing contribution of t conformers with increasing cluster size. This can be interpreted in terms of better solvation of g^+ forms because of the short distance between the solvent and the molecule, which allows DMSO to alternatively interact with any of the substituents.

All the methods used are consistent with one another in terms of solvation trends: the major interaction site is the hydroxyl group, but interaction with the amino group also occurs, either concomitantly, or alternatively. Simultaneous interaction with OH and NH results in the most stable 1:1 complex in which DMSO interacts with both OH and NH in a bidentate fashion. Such a finding has a counterpart in the large occurrence of the g^+G^- 1:1 complex in the FPMD calculations. This also shows that FFMD is well adapted for describing these solvation trends in larger clusters sizes, as the g^+ and G^- contributions both increase with cluster size.

3.2. Solvation and Spectroscopy

The structural description of solvation obtained with the various computational models can now be challenged by the comparison between the simulated and experimental VCD spectra. DMSO does not contribute to the VCD response of *trans*-AI but modifies its conformer distribution and its vibrational modes localized on the amino and hydroxyl groups. Static calculations show that the interaction with OH leads to greater spectral variations in the region of β(OH) et β(NH) than the interaction with NH_2 does. FFMD results on *trans*-AI in bulk DMSO yield the same results, with the average residence time of DMSO involved in an OH…O bond being longer than that involved in an NH…O interaction (160 ps vs. 100 ps). Examination of individual spectra from static calculations shows that G^- structures are necessary to reproduce the spectrum in the 1600 cm^{-1} range, which agrees well with the frequent occurrence of G^- in FPMD trajectories. The poor agreement in the β(OH) et β(NH) region obtained with the bidentate cluster, despite being the most stable energetically, indicates that its contribution as a frozen conformation is not an appropriate vision. Although bidentate clusters are frequently encountered in MD simulations, DMSO molecules are found wobbling around the OH group, making the inclusion of $mono_{OH}$ necessary to produce a more realistic spectrum. Lastly, examination of the individual spectra of the 1:1 complexes from static calculations shows that the **eq**T complexes cannot contribute to the spectrum. This result must be put in perspective with the MD results that indicate no (for FFMD) or very limited (for FPMD) occurrence of the **eq**T complexes. The cluster-in-a-liquid approach does not energetically favour **ax**t complexes for the small sizes, although they could contribute to the experiment in the β(OH) region, especially $mono_{OH}$-**ax**tT. Temperature effects are important in terms of NH_2 orientation and puckering angle and can already be appreciated for small cluster sizes, as the FPMD trajectory samples **ax** and t conformations. This is an entropy effect that disfavours closed structures but biases the system towards more open and more flexible orientations. Lastly, the cluster model has limitations for the description of the amino group orientation that is poorly described by the FPMD simulations containing only one DMSO molecule. Hence, although major parts of the experimental spectrum apparently rely on a bidentate 1:1 cluster, including its spontaneous fluctuations, solvation with more DMSO molecules is important to turn the NH_2 group in the appropriate position. Such conformations can only be achieved by an explicit solvation model.

3.3. Synergy among Computational Approaches

The *trans*-AI molecule in DMSO is a flexible system in which weak intra and intermolecular interactions are intimately related. Beyond its intrinsic experimental interest, deciphering its structure through vibrational spectroscopy proves to be challenging for theoretical chemistry.

The solvent is found to have a major role in stabilising multiple local minima of *trans*-AI, including high-energy conformers. Whatever the method, various structures were identified within an energy range as small as 10 kJ/mol. The number of structures is found to increase

with the number of explicit DMSO molecules. In addition to the intrinsic complexity of their energy landscape, the relative energies of these minima are very sensitive to the details of the calculation, such as the quality of the force field in classical MD, the choice of the functional and the basis set in DFT, the inclusion or neglect of dispersion forces, *etc*. Especially relevant in theoretical chemistry is the role of dispersion effects in the DFT calculations. In the static calculations, dispersion interactions have no significant effect on the structure of *trans*-AI but they induce a change in the solvent organisation toward more compact structures. This is the case for non-hydrogen-bonded 1:1 complexes and for some 1:2 complexes (see Figure S3). These structural changes affect relative and binding energies (see Table 1). Therefore, the Boltzmann weights used to obtain the average IR and VCD spectra exhibit some differences as well. In order to disentangle the effects due to structure and energy, the Boltzmann weights obtained with dispersion-corrected energies were assigned to the structures calculated without dispersion correction. This leads to an increased contribution of bi-**eq**g^+G^-, which results in a less satisfactory agreement with the experimental spectrum in the 1400 cm^{-1} region.

The cluster-in-a-liquid approach may overestimate interactions that would be averaged in the bulk, like dipole-dipole interactions between two DMSO molecules in a 1:2 cluster. This explains that the most stable 1:2 clusters found in the static calculations are solvated 1:1 complexes. The interaction between the solvent and the two adjacent protic substituents influences their conformation in a way that might differ from full solvation conditions. The cluster-in-a-liquid model thus considers solvation as based on two mechanisms: either a single molecule interacts with the system in a bidentate fashion, or interaction with DMSO happens *via* a single substituent. It results in monodentate 1:1 complexes with DMSO located either on NH_2 or OH, or 1:2 complexes with one DMSO molecule on each of these sites. The strong preference for the bidentate 1:1 cluster in the static calculations is related to the intrinsic limitation of the cluster-in-a-liquid description of solvation: it is the only 1:1 cluster able to simultaneously reproduce solvation of both OH and NH_2 groups by DMSO, while a full solvation model would always show a DMSO molecule in the vicinity of the protic groups.

As with most liquids, external factors such as temperature and entropy also play a role. Solvent fluxionality indeed influences the thermal corrections that strongly vary among the different complexes (Table S1). Using thermally corrected energies favours floppy complexes and reduces the contribution of more rigid clusters involving stronger interactions. Since static calculations rely on finding these representative minima, the obtained set of structures may overestimate the importance of certain clusters located in particularly shallow energy valleys. Such a situation was found here for the T conformation, which is stabilised in the static calculations but visited only occasionally in MD simulations. Similarly, the equatorial/axial proportion differs between the methods because of the shallow potential energy surface along the puckering coordinate and the low energy barriers, which are easy to cross at room temperature. The importance of thermal effects is also illustrated by the

comparison between static and FPMD calculations. Examination of the spectra of the individual conformers shows that the contribution of G^- conformers is mandatory for a good match with the experiment. However, G^- conformers of the *trans*-AI monomer are high in energy in static calculations, which explain why the latter are at odd with the experiment. In contrast, FPMD allows exploring G^- conformations for the monomer already, due to its entropic advantage, which results in a better agreement with the experiment.

One should be aware that the AMOEBA FF parameters were not fully optimised for small clusters in the gas phase, which could explain why some structures appear to be over-stabilised, as in the case of axial conformations. Keeping in mind that the electrostatic multipoles of *trans*-AI were extracted from QM methods, some small adjustments of the intramolecular parameters of *trans*-AI were also attempted for a better representation of the ring puckering flexibility, not altering the DMSO parameters, which were designed to represent the solvent as bulk. The small energy differences between *trans-A*I conformations predicted by the static approach were difficult to capture by the FF, even after such adjustment. However, inclusion of explicit polarisation effects gives a good description of the solute-solvent interactions provided that a few DMSO molecules are considered in the first solvation shell. Solvent mobility is well reflected by these various spots characterised from the FFMD simulations and the residence time of a DMSO at the OH or NH_2 groups is estimated to be between 70 and 100 ps for the 1:5 complexes. Furthermore, the simulations in a DMSO periodic cell confirm a residence time of 160 and 100 ps for the OH and NH_2 groups, respectively. Such stable complexes with one or two DMSO molecules can be used as appropriate starting points for FPMD for short-scale simulations or for static DFT calculations. FPMD is able to provide a good quality sampling of intramolecular configurations of the isolated *trans*-AI and the 1:1 complex, as can be seen in the sampling clocks (Figure 12). Consequently, FPMD performs very well on the timescales of molecular vibrations. The resulting spectrum from NVPT and TCF agrees well with the experiment, given that the sampling of the solvent is correctly accounted for. An important result of NVPT calculations is the locality of the VCD signal; DMSO does not actively take part in the local VCD generation. Two effects might contribute to the locality of the VCD signal. First, the interaction between the hydroxyl group and DMSO is weak and systems with stronger hydrogen bonding interactions like aqueous solutions might behave differently. Second, the vibrational modes are localised and there is no coupling between those located on DMSO and those located on *trans*-AI. However, we cannot guess *a priori* which solvent will show induced VCD and answering this question will be the subject of future investigation. The locality of the VCD signal found here may justify the use of QM/MM methods for this type of systems where the electronic wave function of the chiral solute only is considered. This should hold also for other molecules solvated in DMSO as long as they do not contain functional groups that carry a strong dipole moment (*e.g*. carbonyl groups).

The advantage of combining several theoretical methods is to be able to cover different time scales. Long timescales have been shown to be necessary to achieve satisfactory sampling

and obtain reliable conformations but also to allow exchanges of solvent molecules around *trans*-AI. FFMD provides a realistically ergodic sampling which has been emphasised by specific spots representing the DMSO positions around the *trans*-AI solute. In contrast, FPMD trajectories cannot provide a satisfactory sampling for the solvation process due to the limited simulation time.

Conclusion

This work confirms the suggestion of previous studies that VCD is an appealing probe of solvation, being much more sensitive to the surrounding molecules than IR absorption.[10-13, 21, 32, 109] The chosen system, (*1S,2S*)-*trans*-1-amino-2-indanol in DMSO, is especially challenging for VCD spectroscopy. Because of its flexibility, many structures coexist within tiny energy differences that strongly depend on the method used. The results of the static calculations highlight that the spectra obtained with a continuum model alone cannot account for the experimental spectra. The bands that correspond to a signature of hydrogen bonding are missing, which further supports the cluster approach with one or two DMSO molecules. Static calculations based on the cluster-in-a-liquid model capture the main feature of the experimental VCD spectra for the 1:1 complex. Solvation by DMSO does not introduce structures that have not been found as minima of the potential energy surface of the monomer, but modifies their relative energies. A bottleneck of the approach is the relative weights associated to each cluster included in the model, which strongly depend on the set of minimum energies, but also on the associated entropy, especially in terms of solvent fluctuations. Conformations higher in energy that are explored in FFMD or FPMD, like axial conformations, may contribute to the spectrum more than their static energy suggests. Independently of the method used, obtaining weights from cluster energies is limited by the fact that full solvation is not considered. This results in an energetically favoured bidentate 1:1 complex, bi-**eq**g^{+}G^{-}, in the static calculations. Overestimating its contribution to the spectrum deteriorates the agreement between experiment and calculations. The picture of a frozen bidentate complex thus appears oversimplified because solvation is a dynamical effect and the fluxionality of the structures should be considered. This result illustrates the difficulties of using cluster models for the description of solvation in bifunctional molecules with interaction of the solvent of medium strength. Bifunctional compounds showing stronger intramolecular interactions, like cis-AI in which there is an intramolecular hydrogen bond, might behave in a different way, and their study is currently in progress. Yet, DMSO mobility is probably overestimated using a cluster model for MD simulations, a problem that should be solved by considering full solvation.

The solvent influences only indirectly the calculated spectra through structural effects because of the locality of the VCD signal. This locality of the result is an important finding that further supports the interpretation of vibrational spectra from the sole perspective of

the solute conformational space and justifies the use of QM/MM calculations, with only a cluster of limited size being treated quantum mechanically.

As with any MD method, ergodicity is an important issue. The limited length of the FPMD trajectories, due to computational costs, makes it important to define properly the starting structure. In this respect, combining FFMD for an exhaustive exploration of the PES with FPMD for the calculation of the VCD spectra has proved to be efficient. The VCD spectra obtained by the NVPT analysis of FPMD built on structures found by the cluster-in-a-liquid approach or the FFMD trajectories is in good agreement with the experiment, as shown in Figure 14, thus confirming the validity of this two-scale approach.

The results obtained here stress the importance of solvent mobility onto the results, which is hardly taken into account by small clusters. Therefore, a thorough FFMD study seems essential to fully understand VCD, as a complementary tool to the static cluster picture. This is why development of polarisable force fields amenable to very large sizes, close to full solvation, is necessary. Innovative tools for calculating VCD spectra based on these polarisable FFMD trajectories are currently under development.

Acknowledgments

We acknowledge the use of the computing centre MésoLUM of the LUMAT research federation (FR LUMAT 2764), the OCCOGEN supercomputer administered by CINES, as well as computer time allowance by DI Univ. Paris-Sud. We thank Mrs Anael Danielou for experimental help and Mr. Cédric Lambin for preliminary theoretical results. This work has been supported by the French National Research Agency (ANR) Dichroprobe project (Grant ANR-18-CE29-0001-01).

References

1. L. A. Nafie, *Chirality*, 2020, **32**, 667-692.
2. G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner and A. Moscowitz, *Journal of the American Chemical Society*, 1974, **96**, 251-252.
3. L. A. Nafie, *Vibrational Optical Activity: Principles and Applications*, John Wiley & Sons Ltd., 2011.
4. S. R. Bhattacharya and T. Burgi, *Nanoscale*, 2019, **11**, 23226-23233.
5. T. P. Golub and C. Merten, *Chemistry-a European Journal*, 2020, **26**, 2349-2353.
6. A. Perez-Mellor, K. Le Barbu-Debus and A. Zehnacker, *Chirality*, 2020, **32**, 693-703.
7. A. Perez-Mellor and A. Zehnacker, *Chirality*, 2017, **29**, 89-96.
8. V. Declerck, A. Perez-Mellor, R. Guillot, D. J. Aitken, M. Mons and A. Zehnacker, *Chirality*, 2019, **31**, 547-560.
9. S. L. Ma, X. L. Cao, M. Mak, A. Sadik, C. Walkner, T. B. Freedman, I. K. Lednev, R. K. Dukor and L. A. Nafie, *Journal of the American Chemical Society*, 2007, **129**, 12364–12365.
10. C. Merten, *Physical Chemistry Chemical Physics*, 2017, **19**, 18803-18812.

11. J. Sadlej, J. C. Dobrowolski and J. E. Rode, *Chem. Soc. Rev.*, 2010, **39**, 1478-1488.
12. C. L. Covington, V. P. Nicu and P. L. Polavarapu, *Journal of Physical Chemistry A*, 2015, **119**, 10589-10601.
13. V. P. Nicu, J. Neugebauer and E. J. Baerends, *Journal of Physical Chemistry A*, 2008, **112**, 6978-6991.
14. T. B. Freedman, X. L. Cao, R. K. Dukor and L. A. Nafie, *Chirality*, 2003, **15**, 743-758.
15. A. Bouchet, T. Brotin, M. Linares, H. Agren, D. Cavagnat and T. Buffeteau, *Journal of Organic Chemistry*, 2011, **76**, 4178-4181.
16. T. Brotin, D. Cavagnat and T. Buffeteau, *Journal of Physical Chemistry A*, 2008, **112**, 8464-8470.
17. T. Buffeteau, D. Pitrat, N. Daugey, N. Calin, M. Jean, N. Vanthuyne, L. Ducasse, F. Wien and T. Brotin, *Physical Chemistry Chemical Physics*, 2017, **19**, 18303-18310.
18. Y. He, W. Bo, R. K. Dukor and L. A. Nafie, *Applied Spectroscopy*, 2011, **65**, 699-723.
19. T. Burgi and A. Baiker, *Journal of the American Chemical Society*, 1998, **120**, 12920-12926.
20. C. Merten and Y. Xu, *Angewandte Chemie-International Edition*, 2013, **52**, 2073-2076.
21. C. Merten, F. Li, K. Bravo-Rodriguez, E. Sanchez-Garcia, Y. Xu and W. Sander, *Physical Chemistry Chemical Physics*, 2014, **16**, 5627-5633.
22. B. Mennucci, C. Cappelli, R. Cammi and J. Tomasi, *Chirality*, 2011, **23**, 717-729.
23. A. S. Perera, J. Thomas, M. R. Poopari and Y. J. Xu, *Frontiers in Chemistry*, 2016, **4**, 9.
24. B. Kirchner, C. Spickermann, S. B. C. Lehmann, E. Perlt, J. Langner, M. von Domaros, P. Reuther, F. Uhlig, M. Kohagen and M. Brüssel, *Computer Physics Communications*, 2011, **182**, 1428-1446.
25. M. von Domaros, S. Jähnigen, J. Friedrich and B. Kirchner, *Journal of Chemical Physics*, 2016, **144**, 064305-064301-064305-064311
26. M. Losada and Y. J. Xu, *Physical Chemistry Chemical Physics*, 2007, **9**, 3127-3135.
27. G. Yang and Y. Xu, *Journal of Chemical Physics*, 2009, **130**, 164506.
28. M. R. Poopari, P. Zhu, Z. Dezhahang and Y. Xu, *Journal of Chemical Physics*, 2012, **137**, 194308.
29. K. Bünnemann and C. Merten, *Physical Chemistry Chemical Physics*, 2017, **19**, 18948-18956.
30. L. Weirich and C. Merten, *Physical Chemistry Chemical Physics*, 2019, **21**, 13494-13503.
31. J. Bloino, A. Baiardi and M. Biczysko, *International Journal of Quantum Chemistry*, 2016, **116**, 1543-1574.
32. A. S. Perera, J. Cheramy, C. Merten, J. Thomas and Y. J. Xu, *Chem Phys Chem*, 2018, **19**, 2234-2242.
33. M. Heshmat, E. J. Baerends, P. L. Polavarapu and V. P. Nicu, *Journal of Physical Chemistry A*, 2014, **118**, 4766-4777.
34. V. P. Nicu, S. R. Domingos, B. H. Strudwick, A. M. Brouwer and W. J. Buma, *Chemistry-a European Journal*, 2016, **22**, 704-715.
35. J. Blasius and B. Kirchner, *The journal of physical chemistry. B*, 2020, **124**, 7272-7283.
36. K. Bünnemann, C. H. Pollok and C. Merten, *Journal of Physical Chemistry B*, 2018, **122**, 8056-8064.
37. L. Weirich, J. Magalhaes de Oliveira and C. Merten, *Physical Chemistry Chemical Physics*, 2019, **22**, 1525-1533.
38. K. Le Barbu-Debus, A. Scherrer, A. Bouchet, D. Sebastiani, R. Vuilleumier and A. Zehnacker, *Physical Chemistry Chemical Physics*, 2018, **20**, 14635-14646.
39. F. Wang and P. L. Polavarapu, *Journal of Physical Chemistry A*, 2000, **104**, 10683-10687.
40. J. Kessler, T. A. Keiderling and P. Bouř, *Journal of Physical Chemistry B*, 2014, **118**, 6937-6945.
41. T. Giovannini, G. Del Frate, P. Lafiosca and C. Cappelli, *Physical Chemistry Chemical Physics*, 2018, **20**, 9181-9197.
42. T. Giovannini, M. Olszowka and C. Cappelli, *Journal of Chemical Theory and Computation*, 2016, **12**, 5483-5492.

43. E. Orestes, C. Bistafa, R. Rivelino and S. Canuto, *Journal of Physical Chemistry A*, 2015, **119**, 5099-5106.
44. M. Srebro-Hooper and J. Autschbach, in *Annu. Rev. Phys. Chem.*, eds. M. A. Johnson and T. J. Martinez, 2017, vol. 68, pp. 399-420.
45. F. Filippone and M. Parrinello, *Chem. Phys. Lett.*, 2001, **345**, 179-182.
46. C. Herrmann, J. Neugebauer and M. Reiher, *New Journal of Chemistry*, 2007, **31**, 818-831.
47. M. A. J. Koenis, Y. Y. Xia, S. R. Domingos, L. Visscher, W. J. Buma and V. P. Nicu, *Chemical Science*, 2019, **10**, 7680-7689.
48. I. Shahriar, M. K. Bin Islam, M. Iqfath, A. Rahman and M. A. Halim, *Theoretical Chemistry Accounts*, 2019, **138**, 32-31-32-16.
49. M. E. Tuckerman, *Theory and Molecular Simulation*, Oxford University Press Inc, New-York, 2010.
50. J.-H. Choi and M. Cho, *Journal of Chemical Theory and Computation*, 2011, **7**, 4097-4103.
51. F. Calvo, *Chirality*, 2015, **27**, 253-261.
52. S. Yang and M. Cho, *Journal of Chemical Physics*, 2009, **131**, 135102.
53. K. Kwac, K. K. Lee, J. B. Han, K. I. Oh and M. Cho, *Journal of Chemical Physics*, 2008, **128**, 105106
54. R. Kubo, *Journal of the Physical Society of Japan*, 1957, **12**, 570-586.
55. S. Abbate, G. Longhi, K. Kwon and A. Moscowitz, *Journal of Chemical Physics*, 1998, **108**, 50-62.
56. R. Ramirez, T. Lopez-Ciudad, P. Kumar and D. Marx, *Journal of Chemical Physics*, 2004, **121**, 3973-3983.
57. D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian and C. Clavaguéra, *Journal of Chemical Theory and Computation*, 2014, **10**, 3190-3199.
58. G. S. Fanourgakis and S. S. Xantheas, *Journal of Chemical Physics*, 2008, **128**, 074506.
59. F. Thaunay, C. Jana, C. Clavaguéra and G. Ohanessian, *Journal of Physical Chemistry A*, 2018, **122**, 832-842.
60. B. Schropp, C. Wichmann and P. Tavan, *Journal of Physical Chemistry B*, 2010, **114**, 6740-6750.
61. A. Scherrer, R. Vuilleumier and D. Sebastiani, *The Journal of Chemical Physics*, 2016, **145**, 084101.
62. M. Thomas and B. Kirchner, *Journal of Physical Chemistry Letters*, 2016, **7**, 509-513.
63. D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, 2009.
64. P. J. Stephens, *Journal of Physical Chemistry*, 1985, **89**, 748-752.
65. J. R. Cheeseman, M. J. Frisch, F. J. Devlin and P. J. Stephens, *Chem. Phys. Lett.*, 1996, **252**, 211-220.
66. L. A. Nafie and T. B. Freedman, *Journal of Chemical Physics*, 1983, **78**, 7108-7116.
67. A. Scherrer, R. Vuilleumier and D. Sebastiani, *Journal of Chemical Theory and Computation*, 2013, **9**, 5305-5312.
68. A. Scherrer, F. Agostini, D. Sebastiani, E. K. U. Gross and R. Vuilleumier, *The Journal of chemical physics*, 2015, **143**, 074106-074106.
69. S. Jahnigen, A. Scherrer, R. Vuilleumier and D. Sebastiani, *Angewandte Chemie-International Edition*, 2018, **57**, 13344-13348.
70. G. Tarczay, G. Magyarfalvi and E. Vass, *Angewandte Chemie-International Edition*, 2006, **45**, 1775-1777.
71. *MacroModel version 9.8; ed. Schrödinger, LLC: New York, NY, 2010*.
72. E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *Journal of Chemical Theory and Computation*, 2016, **12**, 281-296.
73. P. Y. Ren, C. J. Wu and J. W. Ponder, *Journal of Chemical Theory and Computation*, 2011, **7**, 3143-3161.

74. Y. Shi, C. J. Wu, J. W. Ponder and P. Y. Ren, *Journal of Computational Chemistry*, 2011, **32**, 967-977.
75. F. Thaunay, G. Ohanessian and C. Clavaguéra, *Chem. Phys. Lett.*, 2017, **671**, 131-137.
76. A. J. Stone, *Chem. Phys. Lett.*, 1981, **83**, 233-239.
77. A. J. Stone and M. Alderton, *Mol. Phys.*, 1985, **56**, 1047-1064.
78. S. Grimme, S. Ehrlich and L. Goerigk, *Journal of Computational Chemistry*, 2011, **32**, 1456-1465.
79. K. Le Barbu-Debus, N. Guchhait and A. Zehnacker-Rentien, *Physical Chemistry Chemical Physics*, 2007, **9**, 4465-4471.
80. K. Le Barbu-Debus, F. Lahmani, A. Zehnacker-Rentien and N. Guchhait, *Chem. Phys. Lett.*, 2006, **422**, 218-225.
81. J. Tomasi, B. Mennucci and R. Cammi, *Chemical Reviews*, 2005, **105**, 2999-3093.
82. C. Fabri, T. Szidarovszky, G. Magyarfalvi and G. Tarczay, *Journal of Physical Chemistry A*, 2011, **115**, 4640-4649.
83. http://cccbdb.nist.gov/vibscalejust.asp).
84. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. J. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A. , J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. S. Fox, *Journal*, 2009, Gaussian 09, Revision A.02.
85. J. Bogaerts, F. Desmet, R. Aerts, P. Bultinck, W. Herrebout and C. Johannessen, *Physical Chemistry Chemical Physics*, 2020, **22**, 18014-18024.
86. S. Nosé, *Journal of Chemical Physics*, 1984, **81**, 511-519.
87. J. W. Ponder, TINKER - Software Tools for Molecular Design, version 8 . http://dasher.wustl.edu/tinker, (accessed May 2Oth, 2019).
88. Collaborative data science https://plot.ly 2015).
89. R Core Team, R: A language and environment for statistical computing. (2018) https://www.R-project.org/).
90. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Computer Physics Communications*, 2005, **167**, 103-128.
91. The CP2K developers group www.cp2k.org 2020).
92. A. D. Becke, *Physical Review A*, 1988, **38**, 3098-3100.
93. C. T. Lee, W. T. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789.
94. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *Journal of Chemical Physics*, 2010, **132**, 154104
95. S. Goedecker, M. Teter and J. Hutter, *Physical Review B*, 1996, **54**, 1703-1710.
96. C. Hartwigsen, S. Goedecker and J. Hutter, *Physical Review B*, 1998, **58**, 3641-3662.
97. M. Krack, *Theoretical Chemistry Accounts*, 2005, **114**, 145-152.
98. J. VandeVondele and J. Hutter, *Journal of Chemical Physics*, 2007, **127**, 114105.
99. G. Bussi, D. Donadio and M. Parrinello, *Journal of Chemical Physics*, 2007, **126**, 014101.
100. E. Schwegler, J. C. Grossman, F. Gygi and G. Galli, *Journal of Chemical Physics*, 2004, **121**, 5400-5409.
101. *Journal*, http://www.cpmd.org/.
102. N. Troullier and J. L. Martins, *Physical Review B*, 1991, **43**, 1993-2006.
103. S. Jahnigen, Chirpy https://hartree.chimie.ens.fr/sjaehnigen/chirpy).
104. S. van der Walt, S. C. Colbert and G. Varoquaux, *Computing in Science & Engineering*, 2011, **13**, 22-30.

105. J. D. Hunter, *Computing in Science & Engineering*, 2007, **9**, 90-95.
106. A. Bouchet, J. Klyne, S.-i. Ishiuchi, O. Dopfer, M. Fujii and A. Zehnacker, *Physical Chemistry Chemical Physics*, 2018, **20**, 12430-12443.
107. A. Bouchet, J. Klyne, G. Piani, O. Dopfer and A. Zehnacker, *Physical Chemistry Chemical Physics*, 2015, **17**, 25809-25821.
108. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 12068-12073.
109. C. Merten, R. McDonald and Y. Xu, *Inorganic Chemistry*, 2014, **53**, 3177-3182.