UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVI CICLO DEL DOTTORATO DI RICERCA IN CHIMICA

Interplay between chirality and dynamics of complex systems: a novel computational approach

Settore scientifico-disciplinare: **CHIM/02**

DOTTORANDA
MARTA MONTI

COORDINATORE
PROF. ENZO ALESSIO

SUPERVISORE DI TESI
PROF. MAURO STENER

ANNO ACCADEMICO 2022/2023

"Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933.
Now it is our turn to study statistical mechanics."

- D. L. Goodstein, *States of Matter* (**1985**)

Acknowledgments

I would like to start thanking my supervisor, Prof. Mauro Stener, for everything he taught me. I couldn't have asked for a more brilliant, patient and kind person to guide me through this first experience as a researcher. The same goes for Prof. Massimiliano Aschi who has always had time to call me, follow me through my work, answer all my questions and give me a lot of advice, despite the distance. Thanks to both of you for being my mentors.

There are several other scientists I have had the pleasure to work with, who left me with way more than a scientific paper. I will start thanking Prof. Daniele Toffoli and Prof. Emanuele Coccia for the work done together and for the support during emotional times. Many thanks to Prof. Hannu Häkkinen and all his group at the University of Jyväskylä for hosting me and making me feel so welcome. Thank you Hännu for all our discussions and the nice words you had for my work. I am truly honored. Many thanks also to Prof. Stefano Corni, Dr. Giorgia Brancolini, Dr. Alessandro Fortunelli for the projects we shared. I would like to make a special mention to Erica Scarel, Prof. Silvia Marchesan, and Prof. Ali Hassanali. I am really glad to conclude my PhD experience with our work together.

Naturally, all the ups have come with their downs and I am sure I couldn't have survived them without my amazing colleagues. Plus, I know that it's not easy to share a room with me sometimes, so I'll always be grateful to all the friends I met in that office, because they still invite me out for drinks despite my personality. In particular, I want to thank Elena, Matteo, Chiara, Pierpaolo and Leonardo, you really made my days there (take care of my PC!). Most of all, I have to thank Marco who has been my first colleague, mentor and friend. I wouldn't have made it without you man, thanks for everything.

Since this experience has consumed most of my being, I also need to acknowledge all the people outside of academia who supported me. I will start with my parents, Ombretta and Angelo, and my brothers, Mirko and Erika. Thanks for all your understanding and help. Thanks to my best friend Giulia who has been there from day 1. Everything we shared, fought for/over has brought us here, and I'm really glad I had the chance to share this journey with you. Thanks to all my friends in Rome who are and always be my home, no matter how far I will go. In particular, I have to acknowledge Riccardo for always listening and cheering for me. Thanks to all my friends in Trieste as well, particularly to Erica, Sara, Roberto, and CAB who shared most of this journey with me.

Finally, I have to express my eternal gratitude to Valerio. For all the times I have called you to celebrate, to laugh, to cry, for all the times you have been there for me, I will always be grateful. Lastly, as the poet Snoop Dogg says:

"I wanna thank me for believing in me
I wanna thank me for doing all this hard work
I wanna thank me for having no days off
I wanna thank me for never quitting"

iv

Abstract

The aim of this PhD project was to set-up a computational procedure for accurate and economic Electronic Circular Dichroism (ECD) calculations of complex systems, such as biomolecules or nanostructures. This goal was achieved combining classical molecular dynamics (MD) simulations, essential dynamics (ED) analysis, and state-of-the-art time-dependent density functional theory (TDDFT) calculations. The procedure was tested on several classes of systems, switching from small peptides in water to gold nanoclusters soluble in different solvent environments. By means of this protocol, we were able to extract the most probable conformers, explicitly include the solvent, and obtain a final statistically averaged ECD for each system under investigation. In all the cases, a good qualitative agreement between the experimental and calculated ECD spectra was obtained, thus confirming the reliability of the procedure. It is worth noting that, for the first time, conformational effects and explicit water molecules have been included in the ECD calculation of thiolate-protected gold nanoclusters. The method was also used to investigate the direct role of the aqueous solvent on chiroptical properties and *vice versa*, showing that the water itself can assume a chiral arrangement due to the solute-solvent interactions. Moreover, the scheme was employed to study the correlation between the conformational landscape in solution and the solid-state evolution of heterochiral Phe-based dipeptides. Therefore, this affordable, yet accurate, scheme for the computation of ECD spectra has shown its ability to reproduce different experimental situations, giving insight into the chiroptical properties of complex systems.

Keywords: ECD spectra, molecular dynamics, essential dynamics, TDDFT

Contents

viii

Chapter 1

Introduction

The chiroptical properties of complex systems, such as biomolecules or nanoclusters, are deeply investigated because of their current and potential applications which range from medicine to optoelectronics. [1, 2] A powerful tool to study such properties is the Electronic Circular Dichroism (ECD), that is defined as the differential absorption of the left- and right-handed circularly polarized light (see Figure 1.1). Indeed, the high sensitivity of this spectroscopy allows to obtain information about the absolute configuration and the conformational preference of chiral compounds. [3] However, the experimental ECD spectra can be difficult to interpret, thus requiring a theoretical support. At the same time, to compute ECD spectra, one needs not only a good electronic scheme but also the inclusion of those aspects which make it such a sensitive technique, *i.e.* conformational and solvent effects. [4, 5] The combination of these aspects could result in an unaffordable computational procedure, hence cost and accuracy must be properly balanced. Before going into the details of this project, an overview on chirality and its role in biology and nanoscience is provided. The chapter will be concluded by introducing the state-of-the-art of the computational procedures employed to study chiroptical properties, with a focus on ECD.
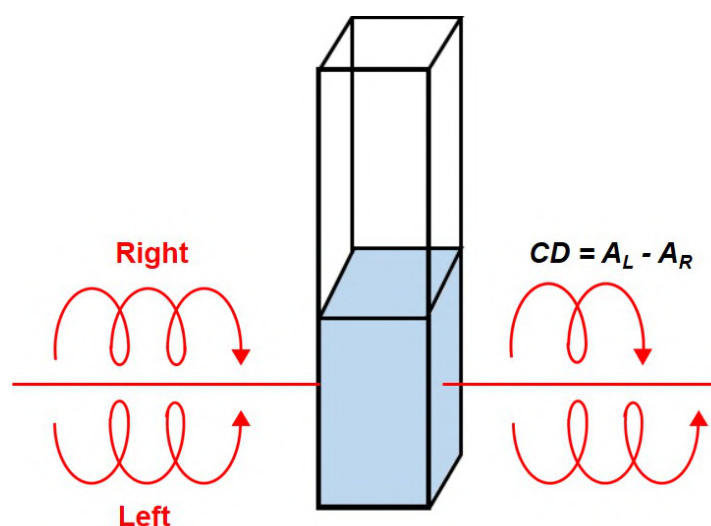

FIGURE 1.1: Schematic representation of the differential absorption of the left- and right-handed circularly polarized light.

1.1   An overview on chirality

Chirality is defined as the geometrical property of an object which is non-superimposable with its mirror image by any translation or rotation. [6] The two mirror images of a chiral compound are known as enantiomers. This property can be observed at different scales going from elementary particles up to macroscopic objects, even astronomical ones as shown in Figure 1.2. [7] A classic example of enantiomeric pairs is the human hands, which actually give the origin to the word chirality (*cheir* = "hand" in Greek). [8]
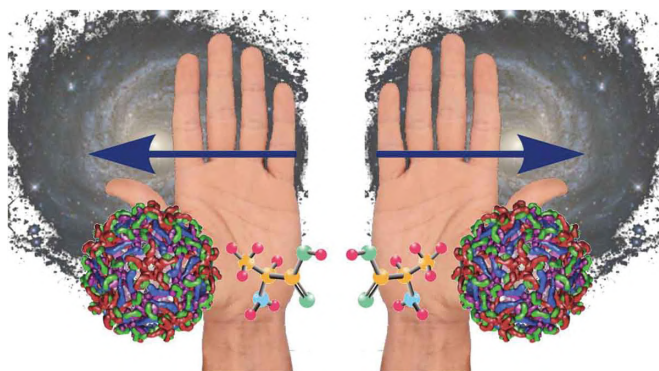

FIGURE 1.2: Representation of different chiral systems existing in our universe
(adapted from ref. [9])

Chiral materials are of particular interest for their unique optical activity, that is their ability to rotate the plane of polarized light. The enantiomer which rotates the plane clockwise is called dextrorotatory (+ enantiomer), while the enantiomer which rotates the plane anticlockwise is called levorotatory (- enantiomer). Several types of interactions between light and chiral compound can be considered, from electronic absorption, to vibrational absorption, or fluorescence. Each of these interactions can be studied with a different spectroscopic technique, *i.e.* ECD, [10] vibrational CD (VCD), [11] or circularly polarized luminescence (CPL). [12] Employing these techniques, one can obtain conformational and three-dimensional structural information on the selected material in different electronic states. [13]

As a result, chiral properties can have a profound impact in many fields such as physics, chemistry, biology, or medicine. [9] For instance, chirality plays a significant role for the detection of diseases. [14] It also has a major impact on drug design since the biological activity of a substance can greatly vary with the selected enantiomer. [15] A famous example here is the case of thalidomide, widely used around the 1960s to treat morning sickness in pregnant women. Marketed as racemic (50:50 mixture of the two enantiomers), it actually interconverts in our body producing the teratogenic *S*-enantiomer which causes several malformations in newborns. [16, 17] In the chemical and material fields, chirality has allowed advances in areas such as enantioselective synthesis, [18] catalysis, [19] and chiral sensing. [20] Additionally, chiral bio-systems have a tremendous impact on the life as we know it. [21] Indeed, it is believed that chirality played a key role on the origin of life and its evolution. [22] The biological importance of chirality will be further discussed in the next section 1.2.

1.2 Chirality in biological systems

The phenomenon of chirality has a strong impact in biochemistry due to its presence in most of the biomolecules. For instance, out of the 20 essential amino acids, only one is achiral. However, what makes this property truly relevant is the natural selection of only one of the two enantiomeric forms in living systems. Indeed, only L-enantiomers occur for amino acids, whereas only D-enantiomers are observed for sugars in biological systems. [23] The criteria for the L,D notation, which is commonly used for biomolecules, are shown in the following Figure 1.3

Amino Acids

L-alanine   D-alanine

Sugars

L-glucose   D-glucose
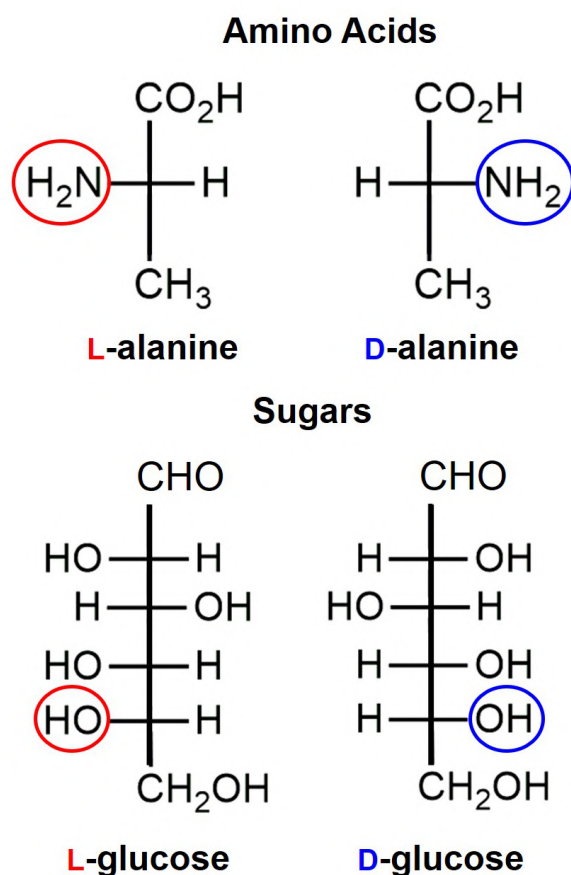
FIGURE 1.3: Examples of the L,D notation for amino acids (top panel) and sugars (bottom panel).

The chirality of biomolecules has a significant impact on their shape, structure and functionalities. For instance, L-amino acids give rise only to right-handed helical proteins. Furthermore, the body is able to metabolize right-handed sugars but not left-handed ones. Similarly, proteases hydrolyze only peptide bonds in proteins formed by all L-amino acids. [23] Absorption, metabolism, and excretion are all so selective processes in the human body that one enantiomer can produce the desired effect, while the other can be completely ineffective or even toxic. As previously discussed, this selectivity is indeed exploited in medicine or drug

design. It is worth mentioning that around the middle of last century, the presence of D-amino acids was uncovered in plants, [24] animals, [25] as well as in the human body. [26] Several studies have shown a correlation between the aging of the organism and its quantity of D-amino acids, hence their potential role in age-related disorders (*e.g.*, arteriosclerosis, Alzheimer's disease, or cataracts). [27, 28]

The specific chirality of biomolecules has also gained much attention due to its relation with the origin of life. The key question of why life is built exclusively on L-amino acids and D-sugars still remains unanswered, although several hypotheses have been proposed. [29–31] One of the first models presented to explain the evolution towards homochirality, which still is a milestone in this field, is known as Frank's model. [32] The idea behind it is that a substance can act, at the same time, as a catalyst in its own self-production and as a suppressor in the synthesis of its enantiomer, thus enabling the evolution from racemic mixture towards enantiopure molecules.

1.3   Chirality in (gold) nanosystems

The theory of chirality is well established for molecules, while it is still evolving for nanostructures. Nowadays, chiral nanomaterials are receiving major attention thanks to the improvement in their synthesis and characterization, as well as for their impressive range of applications. [33] Furthermore, being in the middle between chiral molecules and chiral macroscopic objects, these nanostructures can help in gaining insight into the origin of chirality. [9] Another advantage of chiral nanosystems is represented by the possibility of finely tuning their properties according to their size, shape, and composition. [34–36] Up to now, chirality has been observed in several classes of nanomaterials, *e.g.*, nanoribbons, [37] nanowires, [38] nanoalloys, [39] metal nanoclusters (NCs), [40] and both semiconductor [41] and carbon nanoparticles. [42]

Chirality in nanostructures can result from a chiral shape, a chiral assembly of interacting nanosystems, and a nanoparticle-molecule interaction as shown in Figure 1.4. The first of these sources of chirality (a in Figure 1.4) is observed for instance in twisted nanosystems (*e.g.*, nanorods, gold clusters and ribbons) [43–45] or helical structures. [46] Here, for metal-based nanosystems, the CD signals are observed in the plasmonic region (UV energy range), thus where the electronic collective excitation occurs. The second case, b in Figure 1.4, results from the chiral aggregation of achiral interacting nanoparticles. The CD response is particularly strong with plasmonic nanoparticles since their assembly leads to a plasmon coupling. [47] For the latter case c instead, we can consider two effects to explain the resulting chirality: (i) chiral molecules transfer the chirality to achiral metal cores, [48] (ii) achiral ligands form chiral absorption patterns on the metal. [49]

Sources of chirality in nanomaterials

a) Chiral shape   b) Chiral assembly of   c) Chiral nanoparticle-
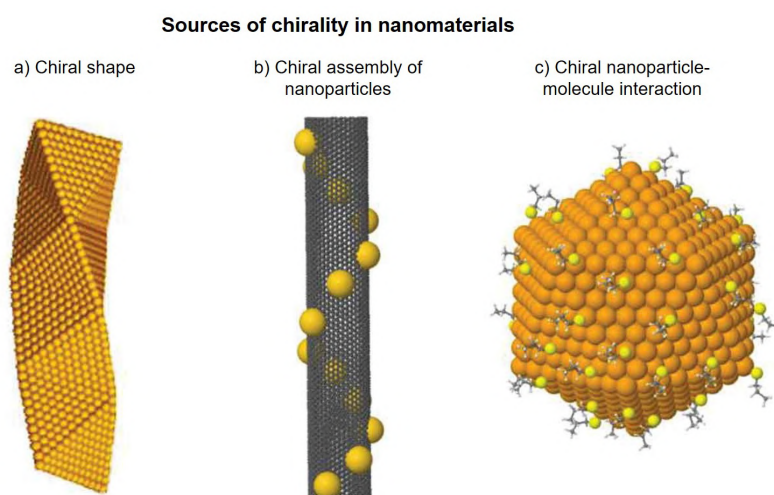   nanoparticles   molecule interaction

FIGURE 1.4: Three examples of chirality in nanosystems: a) Chiral shape found in the helical structure of a gold nanowire, b) chiral assembly obtained by the absorption of gold nanoparticles on a carbon nanotube, and c) induction of chirality due to the asymmetrical arrangement of organic molecules on a symmetrical metal NC.

Among the chiral nanomaterials, metal NCs have gained much attention due to their atomic precision, *i.e.* the size which corresponds to a discrete 'magic' number of atoms. [50] Properties such as high molar absorptivity and photostability contribute to the interest towards these systems. [13] In most of the studies on NCs, the silver or gold-based metal cores [40, 51, 52] are protected by a layer of ligands, *e.g.* thiolates, phosphines, and others [53–55] to avoid aggregation and guarantee the NC stability. From now on, we will only focus on chiral thiolate-protected gold NCs (RS-AuNCs), which were the object of this PhD project.

Chirality in RS-AuNCs was first detected in 1998 by Whetten and collaborators who worked on clusters with 20-40 Au atoms protected by L-glutathione (GSH). [48] Over the last 20 years, significant advances have been made on the synthesis, total structural characterization, as well as on the experimental and theoretical investigation of the chiroptical properties of RS-AuNCs. [44, 56–61] The three mechanisms considered above to elucidate chirality in nanosystems can be adapted here. Indeed, three models have been developed to explain the source of chirality in RS-AuNCs. In the first model, known as intrinsic chiral core model, RS-AuNCs have a metal core that is intrinsically chiral or distorted by the interaction with the ligands. An example of chiral bare NC is the anionic $Au_{34}{}^{-}$ cluster shown in Figure 1.5 - a. [62–64] When the metal core has a symmetric arrangement instead, the chiral response can arise from both chiral and achiral ligands. In the former case, the ligands induce the chirality on the metal core by trapping its electrons in a dissymetric electric field (dissymetric field model). [65] In the latter case, the achiral ligands form a chiral adsorption pattern onto the metal structure (footprint model). [51] For instance, $Au_{25}(SR)_{18}$ is achiral since it has a symmetric Au_{13} core surrounded by symmetric RS-Au-SR-Au-SR staples. However, chiroptical properties result for this motif when using chiral ligands (*e.g.,* L-GSH) as shown in Figure 1.5 - b. Instead,

$Au_{38}(SR)_{24}$ owes its chirality to the asymmetric arrangement of the Au-S staples (see Figure 1.5 - c), despite the non-chirality of its thiolate ligands and the high symmetry of the gold kernel. Therefore, the ECD features of RS-AuNCs arise from the response of the metal core, Au-S staples, and protective layer. Factors as the flexibility of both the gold-sulphur interface and the passivating ligands can have a drastic impact on the chiroptical features, hence they must be taken into account to simulate ECD spectra.
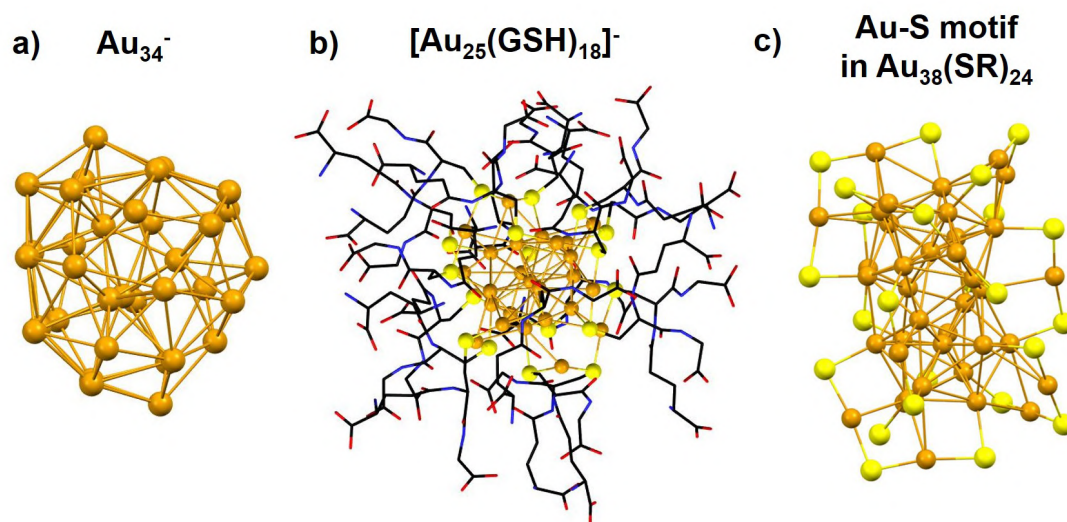

FIGURE 1.5: Structure of: a) the anionic Au_{34}^- cluster, b) the $[Au_{25}(GSH)_{18}]^-$ cluster, and c) the Au-S staple motif in $Au_{38}(SR)_{24}$. Au, and S atoms are represented with a stick and ball model in orange, and yellow, respectively. The C, O, and N atoms of the GSH ligands are represented with a stick model in black, red, and blue, respectively. H atoms are omitted for sake of clarity.

1.4   Computational protocols for ECD calculations

Among the techniques available to study chiroptical properties (*e.g.*, ECD, VCD, or CPL), we focused on ECD which is helpful to ascribe the absolute configuration of chiral compounds. The ECD spectroscopy is also a useful tool to investigate the conformational preference of a solute in its electronic ground-state, as well as its interaction with the solvent. [3] For these reasons, the theoretical modeling of ECD has attracted much interest over the last decades and several approaches have been proposed. In order to summarize the state-of-the-art of the computational procedures for ECD, there will be a first focus on schemes useful for organic and bio-molecules, followed by approaches more aimed at the RS-AuNCs treatment.

Anyway, a good electronic scheme is essential for all the ECD calculations, despite the nature of the selected system. The time-dependent density functional theory (TDDFT) [66] has gained a leader role in this field due to its satisfactory compromise between accuracy and computational cost. The quality of TDDFT calculations has especially increased with the arrival of hybrid exchange-correlational (xc) functionals as B3LYP, [67, 68] and range-separated ones as CAM-B3LYP, [69] or wB97X-D. [70] However, the cost of the calculation increases with

the accuracy of the xc functional, *e.g.* the inclusion of the exact Hartree-Fock (HF) non-local exchange, hence its choice is strongly related to the system size. The most popular TDDFT formalism was proposed by Casida, [71] whose equations are usually solved by means of the Davidson iterative algorithm [72] in quantum-chemical codes. Such approach is efficient when a relatively small number of low excitation energies is extracted but becomes unaffordable for large systems. Therefore, the Casida's formulation and its computational implementation work fine mostly for small-/medium-size molecules.

Bearing in mind the high sensitivity of ECD towards molecular flexibility, [3] it becomes clear that the accurate electronic scheme must be combined with a proper inclusion of the conformational effects. It is worth noting that conformers are also sensitive to the solvent presence, in particular to polar ones. Indeed, especially in the latter case, the solvent can be engaged in a wide range of weak inter-molecular interactions with the solute, *e.g.* H-bonds, charge-dipole transfers, or dipole-dipole interactions. Therefore, solvent effects should be included, at least at a partial extent, in a physically coherent computational scheme for ECD. [4, 5]

Many groups have worked on this topic, and several computational approaches have been proposed in the literature. The problem of treating properly large flexible molecules has been recently assessed by Grimme and co-workers, who extended the simplified TDDFT (s-TDDFT) method to the calculation of the optical rotation (OR). The sTDDFT is a very economic approach which can be used on a large amount of conformers obtained, for instance, by a preliminary molecular dynamics (MD) search. [73] In 2021, Grimme's group also proposed an automatic algorithm for the OR calculation of large and flexible biological molecules. Here, meta-dynamics, clustering through principal component analysis (PCA) and quantum-chemical methods (*i.e.* DFT, sTDDFT) are combined to obtain a final statistically weighted OR. [74] For a thorough conformational search, the Monte Carlo (MC) algorithm can be exploited as an alternative to MD simulations. Pescitelli and co-workers have adopted this strategy in several works combining the MC sampling with DFT optimizations, and TDDFT calculations for the spectra of (bio-)organic compounds. [75–77] Moreover, Barone *et al.* recently proposed to identify the low-energy minima of flexible molecules through the island model evolutionary algorithm. [78] This unsupervised search of conformers is combined again with quantum-chemical calculations to compute the optical properties.

The role of the solvent, especially water, has also been widely assessed in recent works. The aqueous solvent presents some peculiarities which must be considered to formulate a proper model. Indeed, despite its small size, water has a high dipole moment and the ability to form hydrogen bonds, both as an acceptor and a donor. [79] Therefore, models that seek to describe water and the solute-solvent interactions should accurately take into account its high polarity and H-bonds. [80] While shaping a solute-solvent model, one should also consider the instantaneous polarization of the system, hence its instantaneous configuration. At the same time, it is clear that the inclusion of all these aspects, together with the explicit treatment of the solvent molecules, can result in an unaffordable computational procedure.

Over the last years, different approaches have been proposed to treat the external environment. Among them, we find the focused models which treat the solute at a Quantum Mechanical (QM) level, whereas the solvent at a classical mechanics one. [81] These models work when the chiroptical signals arise from the solute, while the solvent acts as a perturbation. An immediate advantage is the possibility of treating large systems due to the lower level of accuracy required for the solvent. However, this approach fails when the solute imprints a chirality to the surrounding achiral environment (*i.e.,* chirality imprinting or transfer). [82, 83] In this case, a layer of water molecules should be included in the QM region to overcome the limitation. The most employed focused model is the Polarizable Continuum Model (PCM), [81, 84] which is suitable for weak interactive solvents. Indeed, QM/PCM calculations are usually performed on minimum-energy structures, that are considered as representative of the whole conformational ensemble, neglecting the solvation dynamics. As an alternative, the QM methods can be combined with Molecular Mechanics (MM) approaches to retain the atomistic nature of the aqueous environment. [85, 86] An example here is the QM/Fluctuating Charge approach developed by Cappelli and co-workers to specifically model the optical properties of molecules in aqueous solution. [80, 87] Recently, Barone and colleagues proposed a model that integrates the ONIOM/EE procedure (QM/MM) with the Perturbed Matrix Method (PMM) for accurate simulations of the chiroptical spectra in solution. [88] Instead, Crawford *et al.* modeled solvent effects combining classical MD for the sampling with TDDFT calculations for the spectra. For the selected configurations, the water molecules are treated explicitly on a QM level if within the solute's cybotactic region, or implicitly if in the bulk region. [89]

The considerations on the conformational and solvent impact on ECD can be extended to the RS-AuNCs case. The same goes for the relevance of employing an accurate electronic scheme to calculate the spectra. The simulation of the ECD spectra of RS-AuNCs is clearly challenging because of their size and complex nature. Indeed, as stated above, the Casida's scheme becomes less and less efficient increasing the dimension and density of the states of the system. Here, it could be used to extract only the lowest part (1-2 eV) of the absorption spectra, losing the features that arise from the passivating ligands (E > 3 eV). One way to overcome this limitation is to use a model of the real system, for instance replacing the thiolate ligands with smaller ones as -SH, and -SCH$_3$. This strategy is adopted in several works to facilitate the investigation of the electronic/optical properties. [61, 90, 91] An alternative strategy is to adopt electronic schemes that are more suitable to compute the optical properties of large systems. In this regard, it is worth mentioning the first-principles method developed by Noguez and co-workers, which combines TD perturbation and density functional theories to calculate the CD of nanostructures. [92] Moreover, Stener's group developed a new algorithm to solve the TDDFT equations, that is particularly suitable for simulating the absorption spectra of RS-AuNCs. [93] This method extracts the photoabsorption and ECD spectra from the imaginary part of the polarizability and rotator strength tensor, respectively, at any given photon energy. This way, the bottleneck of Davidson diagonalization is avoided and the whole energy range

of the spectrum can be spanned.

Despite the electronic improvements, the ligands mobility and its influence on the chiroptical properties have been mostly neglected in theoretical studies due to the huge computational cost. Indeed, the ECD of RS-AuNCs is typically calculated on the resolved X-ray structure in vacuum. [94] An attempt of inclusion was made by Sementa and co-workers, who explored the conformational space of a $Au_{38}(SR)_{24}$ NC to highlight the relation between ligands flexibility and optical properties. [95] Furthermore, Bürgi and collaborators have recently investigated the dynamic effects of AuNCs by combining VCD spectroscopy with theoretical calculations (MD simulations and DFT). [96–98] The computational cost has also brought to neglect the explicit inclusion of the solvent in the ECD calculations of RS-AuNCs. The simulation of optical properties in gas-phase can be less dramatic here since most of the nanoclusters are soluble in non-polar organic solvents. However, the solute-solvent interactions affect the conformational changes of RS-AuNCs, as demonstrated by Akola and colleagues. [99] Therefore, the organic solvent should at least be considered to find the low-energy conformers. Instead, for the clusters soluble in water, *e.g.* Au clusters protected by L-GSH, an explicit inclusion of the solvent, at least at a partial extent, is required. Recently, Akola *et al.* investigated how water affects the electronic properties of $[Au_{25}(GSH)_{18}]^-$ by using a QM/MM approach. [100] To the best of our knowledge, the works presented in this thesis are the first ones considering conformational changes and explicit water molecules in the ECD calculations of RS-AuNCs.

Chapter 2

Theory

2.1 Molecular Dynamics

Computer simulations are performed to help understanding the properties of molecules and their microscopic interactions. Indeed, we can see them acting as a bridge between the microscopic length and the macroscopic experimental world. Simulations can be used not only to predict and understand properties but also to evaluate a model/theory by direct comparison with experimental results. Furthermore, they can help gaining insight into experiments that are difficult to conduct in a laboratory (*e.g.*, extreme temperature or pressure conditions). Naturally, the accuracy of the predictions is subjected to the limitation of the computational cost. [101] The two main families of computer simulations are Molecular Dynamics (MD) and Monte Carlo (MC), which can also be combined in a family of hybrid techniques. In this work, we use and explain the theory of MD which gives a route to the dynamical properties of a system. In MD simulations, the atoms move and interact for a fixed time period giving a view of the dynamic evolution of the system. The resulting MD trajectory can be used then to calculate macroscopic properties ($\bar{A}$) which are defined by the statistical mechanics, [102] within the ergodic hypothesis, as the time averages of the corresponding microscopic properties (A):

$$\bar{A} = \frac{1}{T} \int_{t}^{t+T} A(\tau)d\tau \tag{2.1}$$

All the relevant aspects of classical MD simulations are discussed in the following sections.

2.1.1 Equations of motion: definition and integration

The time evolution of a set of N interacting particles with positions $\mathbf{r}_i$ and masses m_i, with $i = 1, ..., N$, can be followed by integrating their equations of motions, *i.e.* the second order differential equations according to Newton's second law:

$$\mathbf{F}_i = m_i \frac{d^2\mathbf{r}_i}{dt^2} = -\frac{\partial U(\mathbf{r}^N)}{\partial \mathbf{r}_i} \tag{2.2}$$

$\mathbf{F}_i$ is the force acting on the $i - th$ particle due to its interaction with other particles. The set of forces is derived from the potential energy term, $U(\mathbf{r}^N)$, that is described in detail in

section 2.1.6. Instead, $\frac{d^2\mathbf{r}_i}{dt^2}$ is the acceleration of the $i-th$ particle. It is worth noticing that the equations of motion can be similarly described by the Newtonian (Eq. 2.2) and Hamiltonian formalism. In the latter case, starting from the following system of equations:

$$\begin{cases} H(\mathbf{r},\mathbf{p}) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}^N) \\ \dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} \\ \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} = \mathbf{F}_i \end{cases} \tag{2.3}$$

where $\mathbf{r}$ and $\mathbf{p}$ represent positions and momenta, respectively, and calculating the second derivative of $\mathbf{r}$ with respect to t:

$$\ddot{\mathbf{r}}_i = \frac{\dot{\mathbf{p}}_i}{m_i} = \frac{\mathbf{F}_i}{m_i} \tag{2.4}$$

the Newton's equation of motion (Eq. 2.2) is exactly obtained.

In MD simulations, Eq. 2.2 is solved numerically with a discretization of the time. Therefore, for each particle i, we can go from the position $\mathbf{r}_i(t)$ at time t to the position $\mathbf{r}_i(t+\Delta t)$ at time $t + \Delta t$ (Δt equals to the time step) but without knowing its evolution in the middle. The time step, Δt, should be selected as large as possible to rapidly sample the phase-space and reduce the computational cost. [103] However, Δt must be smaller than the highest frequency simulated motions, hence it is limited by the fast forces that occur among the atoms. A typical value for Δt is around 2 fs in classical MD simulations.

Let us expand in Taylor series the position for the particle i at $(t+\Delta t)$ and $(t-\Delta t)$, respectively:

$$\begin{cases} \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{1}{2}\frac{\mathbf{F}_i}{m_i}\Delta t^2 + ... \\ \mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \mathbf{v}_i(t)\Delta t + \frac{1}{2}\frac{\mathbf{F}_i}{m_i}\Delta t^2 - ... \end{cases} \tag{2.5}$$

Summing up the two expressions in Eq. 2.5, the well-known Verlet integrator [104] is obtained:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) + \frac{\mathbf{F}_i}{m_i}\Delta t^2 - \mathbf{r}_i(t - \Delta t) + O(\Delta t^4) \tag{2.6}$$

Therefore, knowing the previous and current positions, and the current acceleration, we can calculate the next position of the particle i.

The Verlet integrator satisfies two of the three properties required by a stable MD propagation algorithm: time-reversibility and preservation of the density in the phase-space. However, the energy is not conserved with this algorithm. Another deficiency lies in the description of the velocity, and correlated kinetic energy and temperature, since this observable is not explicitly given in Eq. 2.6. This limitation can be overcome by estimating the mean-value of the velocity with the position terms. However, such approach introduces an error of order

$O(\Delta t^2)$ and $O(\Delta t)$ for $\mathbf{v}_i(t)$ and $\mathbf{v}_i(t + \Delta t)$, respectively.

Improvements of the position-Verlet integrator, such as the Leap-Frog [105] and the velocity-Verlet [106] algorithms, have been proposed. These two algorithms, briefly explained below, are equivalent to the Verlet integrator, *i.e.* they produce the same trajectory. In the Leap-Frog method, the position $\mathbf{r}_i$ and the velocity $\mathbf{v}_i$ are computed at integer and half integer time steps, respectively, hence "jumping" over each other as leapfrogs:

$$\begin{cases} \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{1}{2}\Delta t)\Delta t \\[2mm] \mathbf{v}_i(t + \frac{1}{2}\Delta t) = \mathbf{v}_i(t - \frac{1}{2}\Delta t) + \dfrac{\mathbf{F}_i(t)}{m_i}\Delta t \end{cases} \tag{2.7}$$

In the velocity-Verlet algorithm, positions and velocities evolve simultaneously as:

$$\begin{cases} \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \dfrac{\mathbf{F}_i(t)}{2m_i}\Delta t^2 \\[2mm] \mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t)\Delta t + \dfrac{\mathbf{F}_i(t) + \mathbf{F}_i(t + \Delta t)}{2m_i}\Delta t \end{cases} \tag{2.8}$$

In practice here, the velocities are fist computed at half-time step, then the positions, forces, which are only $\mathbf{r}_i$-dependent, and velocities are evaluated at $t + \Delta t$. Time-reversibility and density preservation in the phase-space are still retained with these algorithms. The Leap-Frog method is the main integrator in program packages as GROMACS, [107] while the velocity-Verlet is more used in the NAMD package. [108] Other families of integration methods, as the predictor-corrector algorithms [109] or symplectic integrators, [110] are less employed because of their higher complexity and computational cost.

2.1.2 Periodic Boundary Conditions

By integrating the Newtonian equations of motion, the atoms of our system are free to move through the simulation box. Naturally, the size of the simulated system is limited by the computational cost, hence only a portion of the real sample can be computed. This approximation consequently entails the necessity to avoid the border effects. The way to achieve this goal is by applying the Periodic Boundary Conditions (PBC) which are schematized in Figure 2.1. In practice, the simulation box is treated as surrounded by infinite replica over the three dimensions of the space. This is simply accomplished by considering that if a particle moves out of the box, an image particle moves in to replace it (see Figure 2.1). This treatment introduces artifacts in disordered systems such as liquids, while it works better for the highly ordered crystals. Furthermore, the box size should be large enough to avoid that a particle interacts with itself. In the case of short-range interactions, a common rule employed is the *minimum image convention* in which a particle interacts only with the nearest image of all the other particles. [101]
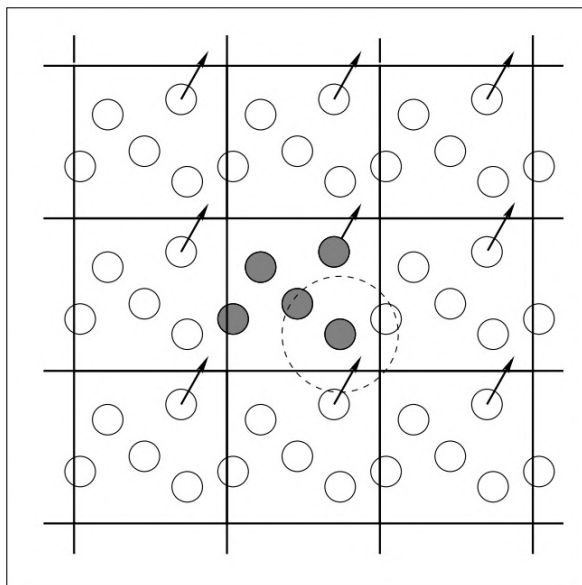
FIGURE 2.1: Exemplification of the PBC: when a particle moves out of the box,
an image particle moves in as a replacement.

2.1.3 Cutoff and Neighbor Lists

In principle, the computation of non-bonded interactions within the PBC and *minimum image convention* requires a huge number of pairwise calculations. Indeed, for each atom i, we should loop over all the j-atoms to find the minimum image distance, r_{ij}. A first assumption can be made imposing that the potential is equal to zero when $r_{ij} > rcut$, thus the distance between two atoms is bigger than the potential cutoff. This way, the calculation time is reduced but still proportional to $O(N^2)$ since all the distinct pairs need to be examined $\left(\frac{1}{2}N(N-1)\right)$. Therefore, the computational cost would become unmanageable for large systems. Methods of neighbor searching are usually employed in order to speed up the program. A technique for using lists of nearby pairs of atoms was proposed by Verlet. [104] In this case, a sphere of radius r_{cut} is surrounded by an external layer to form a larger sphere of radius r_{list}, as shown in Figure 2.2 - a. At the beginning of the simulation, a neighbor list is constructed for each atom i containing all the j-atoms for which $r_{ij} < r_{list}$. Over the next steps, only the pairs within this list are checked in the force routine. Along the simulation, the list is reconstructed to avoid that unlisted pairs move within the interaction range. The Verlet list approach is suitable for small systems, while techniques such as the cell-list method are preferable for larger systems (*e.g.* more than 1000 atoms). In the latter approach, the simulation box is divided into a number of cubic cells whose side, r_{cell}, is greater than r_{cut} (see panel b in Figure 2.2). Here, the search for neighbors of an atom happens only within the cell of the atom itself and those adjacent. [101]
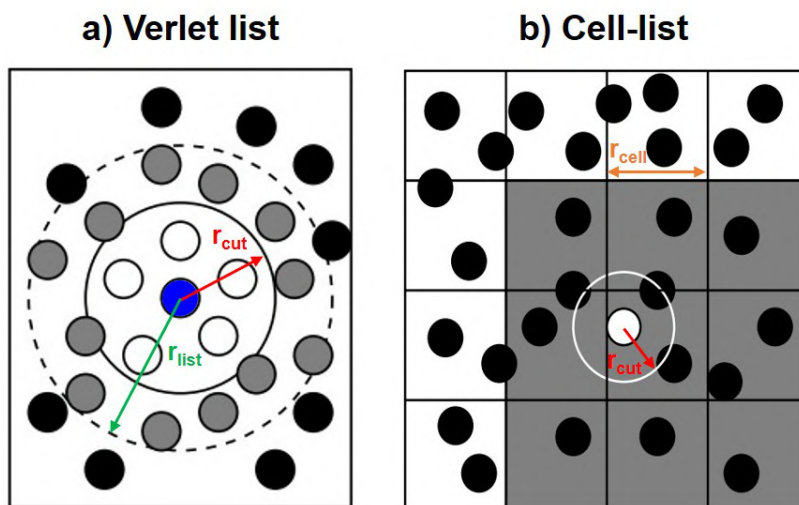
FIGURE 2.2: Examples of neighbor searching methods: a) Verlet list method, and b) cell-list method.

2.1.4 Ensembles

The integration methods introduced above can be employed to work in different statistical ensembles. For instance, in principle we could run MD simulations in the microcanonical ensemble where the number of particles N, volume V, and energy E are kept constant. However, an exact NVE situation is practically never reached unless a very small integration step is used. In addition, most of the experiments are conducted at fixed temperature or pressure, thus making the microcanonical ensemble less suitable for reproducing the experimental conditions. The ensembles where we fix N, V, T, and N, P, T are defined as the canonical and isothermal-isobaric ensemble, respectively. To simulate the NVT or NPT ensembles, it is necessary to couple the system with a thermostat for the temperature control and/or a barostat for pressure control, respectively. This aspect will be further discussed in section 2.1.5, whereas more details on the thermodynamic ensembles will be given right below.

The NVE ensemble is the collection of all the microstates where N, V, and, E are constant and can be described by a time-independent Hamiltonian $H(\mathbf{r}, \mathbf{p})$. Only the microstates corresponding to the fixed energy are accessible and have equal probabilities.

Let us introduce now the partition function, Q, which describes how the probability is partitioned among the available microstates. For the microcanonical ensemble, Q_{NVE} can be written as:

$$Q_{NVE} = \frac{1}{N!} \frac{1}{h^{3N}} \int \delta(H(\mathbf{r,p}) - E) \mathbf{drdp} \tag{2.9}$$

where N is the number of particles, h is the Planck's constant, E is the total energy, and δ is the Dirac delta function. The Q_{NVE} function is also connected to the thermodynamic potential (Boltzmann's expression of the entropy): [111]

$$S = k_B \ln Q_{NVE} \tag{2.10}$$

The *NVT* ensemble collects instead the available states of a system in contact with a thermal bath. Here, the energy is no longer constant but fluctuates in order to maintain the desired temperature. Therefore, in the canonical ensemble all values of energy are allowed with a probability density that corresponds to the Boltzmann's distribution:

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int e^{-\beta H} \mathbf{dr} \mathbf{dp} \tag{2.11}$$

where $\beta = 1/k_B T$. Here, the thermodynamic potential is connected to the Helmholtz free energy, A, as follows:

$$A = -k_B T \ln Q_{NVT} \tag{2.12}$$

where $A = U - TS$.

The *NPT* ensemble can be used to reproduce experiments conducted at both fixed temperature and pressure. The expression for the partition function, Q_{NPT}, resembles the previous one (Eq. 2.11), but it also takes into account the fluctuation of the volume as:

$$Q_{NPT} = \frac{1}{N!} \frac{1}{h^{3N}} \frac{1}{V_0} \int \int e^{-\beta(H+PV)} \mathbf{dr} \mathbf{dp} dV \tag{2.13}$$

The thermodynamic potential associated to this ensemble is proportional to the Gibbs free energy:

$$G = -k_B T \ln Q_{NPT} \tag{2.14}$$

where $G = U + PV - TS$.

Other thermodynamic ensembles, such as the grand canonical (μVT) ensemble, will not be discussed in this work.

2.1.5   Thermostats and Barostats

As stated in section 2.1.4, some modifications must be introduced in order to switch from the *NVE* ensemble to the *NVT* or the *NPT* one. MD simulations in the canonical ensemble require the contact between the system and an external thermal bath to keep constant the temperature. Several approaches have been proposed to achieve this goal, each of them with its strengths and limitations. It is worth noticing that a good thermostat should be able to guarantee the correct fluctuations of the kinetic energy, K. We know from the statistical mechanics that the average kinetic energy value ($\langle K \rangle$) is equal to:

$$\langle K \rangle = \frac{3}{2} N k_B T \tag{2.15}$$

while its variance, σ_K^2, to:

$$\sigma_K^2 = \frac{2}{3N} \langle K \rangle^2 \tag{2.16}$$

where $\langle K \rangle^2$ are the fluctuations of K.

Two of the first proposed algorithms actually reproduce the wrong energy fluctuations, although they have other advantages such as working well for the equilibration of the system. One of these methods is the velocity rescaling approach which, as suggested by the name, rescales the velocities to obtain the correct K at each step of the simulation. Therefore, for a particle i, the new velocity, $\mathbf{v}_i^{NEW}$, will be:

$$\mathbf{v}_i^{NEW} = \alpha \mathbf{v}_i^{OLD} \tag{2.17}$$

with the rescaling factor, α, equals to:

$$\alpha = \sqrt{\frac{\bar{K}}{K}} \tag{2.18}$$

$\bar{K}$ is the kinetic energy at the target temperature, while K is the instantaneous kinetic energy. The second method was proposed by Berendsen and co-workers in 1984. [112] The Berendsen thermostat shows similarities with the velocity rescaling one, but it corrects the velocities more slowly. Here, the new velocity is defined as:

$$\mathbf{v}_i^{NEW} = \sqrt{\frac{c_1 K + c_2 \bar{K}}{K}} \mathbf{v}_i^{OLD} \tag{2.19}$$

with the two coefficients equal to:

$$\begin{cases} c_1 = e^{-\frac{\Delta t}{\tau}} \\ c_2 = 1 - c_1 \end{cases} \tag{2.20}$$

Underlying the Berendsen approach is the assumption of a weak coupling between the system and the heat bath, whose coupling constant is the term τ shown in Eq. 2.20. This latter parameter, which is in time units, describes the strength of the coupling, therefore the larger the value, the weaker the coupling. Considering the difference between the kinetic energy at the new time step and the current kinetic energy (ΔK), and a time step small enough with respect to τ, the following differential equation results:

$$\frac{\mathrm{d}K}{\mathrm{d}t} = \frac{\bar{K} - K}{\tau} \tag{2.21}$$

Eq. 2.21 shows that the convergence of the instantaneous K towards the target $\bar{K}$ is slower. However, as mentioned above, the scheme does not give the canonical distribution, thus it should only be used for the pre-production (*i.e.* equilibration).

Bussi and co-workers proposed an extension of the Berendsen thermostat which enforces the correct distribution for K by adding a stochastic term. [113] Here, the idea is to force the kinetic energy towards the canonical distribution, instead of forcing it towards the average energy at the target temperature. Such goal is achieved by including a stochastic term in Eq.

2.21:

$$dK = \left(\frac{\bar{K} - K}{\tau}\right) dt + 2\sqrt{\frac{K\bar{K}}{N_f \tau}} dW \tag{2.22}$$

Here, N_f is the number of degrees of freedom, dW is the Wiener noise, and τ is the arbitrary parameter in time units that acts as the coupling constant in the Berendsen thermostat. The Bussi-Donadio-Parrinello thermostat, also termed as stochastic velocity rescaling algorithm, has been used in all the works presented in this PhD project.

Other famous schemes are the Andersen [114] and the Nosé-Hoover thermostats. [115, 116] The latter method has become quite popular since it allows a control of the temperature without using random numbers. Unfortunately, it can exhibit non-ergodic behavior even for simple systems as a collection of harmonic oscillators. This limitation can be overcome by introducing chains of thermostats, [117] but in this case the simplicity of the theory is lost.

Similarly to thermostats, we can define barostats as systems that keep constant the macroscopic pressure in the NPT ensemble. Among the several methods proposed for the pressure coupling, we will briefly discuss the Berendsen [112] and the Parrinello-Rahman [118] barostats. The Berendsen algorithms for temperature and pressure coupling share the same logic, which is to weakly couple the system to an external bath. As explained above for the thermostat, an extra term is added to the equations of motion to guarantee that the instantaneous pressure of the system, P, tends to the target one, $\bar{P}$:

$$\frac{dP}{dt} = \frac{\bar{P} - P}{\tau_P} \tag{2.23}$$

Therefore, at each time step, the volume is isotropically scaled by the following factor η:

$$\eta = 1 - \frac{\gamma \Delta t}{\tau_P}(\bar{P} - P) \tag{2.24}$$

where γ is the isothermal compressibility factor. It is common to take γ as the isothermal compressibility of liquid water.

Unlike the Berendsen barostat, the Parrinello-Rahman algorithm proposes an anisotropic scaling, thus allowing the box shape to change as well. Such approach becomes particularly useful to simulate the phase changes in solids. Let us start considering a cell with arbitrary shape and box vectors, $\vec{a}$, $\vec{b}$, and $\vec{c}$, collected in a matrix $\mathbf{h}$. The pressure coupling between the system and the external bath is achieved by transforming the equations of motion as follows:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \mathbf{G}^{-1}\frac{d\mathbf{G}}{dt}\frac{d\mathbf{r}_i}{dt} \tag{2.25}$$

where $\mathbf{G} = \mathbf{h}^T\mathbf{h}$ is the metric tensor.

2.1.6 Interatomic interactions

As introduced in section 2.1.1, the forces acting on the particles, $\mathbf{F}_i$, are derived from the corresponding potential energy term, $U(\mathbf{r}^N)$. The potential energy function can be split into three components: *bonded* term, *non-bonded* term, and the term due to the presence of the external field:

$$U(\mathbf{r}^N) = U_b + U_{nb} + U_{ext.field} \tag{2.26}$$

The bonded term, U_b, includes the 2-, 3-, and 4-body interactions between atoms that are bonded covalently. Therefore, it takes into account the contributions arising from bonds, angles, and both proper and improper dihedral angles. The bond stretching between two atoms, i and j, is usually described with a harmonic potential, $U_{bo}(r_{ij})$:

$$U_{bo}(r_{ij}) = \frac{1}{2}k_{bo,ij}\left(r_{ij} - r_{ij}^0\right)^2 \tag{2.27}$$

where r_{ij}^0 is the distance at the equilibrium. Clearly, the harmonic approximation is valid only when the system is close to the equilibrium configuration. Instead, an option to describe anharmonic bond stretching potentials is the Morse potential, U_M, whose form is: [119]

$$U_M(r_{ij}) = D_{ij}\left[1 - e^{-\beta_{ij}\left(r_{ij} - r_{ij}^0\right)}\right]^2 \tag{2.28}$$

D_{ij} represents the depth of the asymmetric well, while β_{ij} defines its steepness.

The angle bending between a triplet of $i - j - k$ atoms is also described by a harmonic potential function, $U_a(\Theta_{ijk})$:

$$U_a(\Theta_{ijk}) = \frac{1}{2}k_{a,ijk}\left(\Theta_{ijk} - \Theta_{ijk}^0\right)^2 \tag{2.29}$$

There are situations where the interaction between i and k needs to be considered explicitly. In this case, we add in Eq. 2.29 a harmonic correction on the $i - k$ distance which is known as Urey-Bradley term, U_{UB}:

$$U_{UB}(\Theta_{ijk}) = \frac{1}{2}k_{UB,ijk}\left(r_{ik} - r_{ik}^0\right)^2 \tag{2.30}$$

The torsion angles involving the i, j, k, and l atoms are classified as proper or improper dihedral angles. According to the IUPAC convention, the proper dihedral angle ϕ_{ijkl} is the angle between the ijk and the jkl planes, with the atoms connected as $i - j - k - l$. The potential term for proper dihedrals, $U_d(\phi_{ijkl})$, can be described by the sum of cosine functions as:

$$U_d(\phi_{ijkl}) = \sum_n k_{d,ijkl}\left[1 + \cos\left(n\phi_{ijkl} - \delta\right)\right] \tag{2.31}$$

where n is a non-negative integer constant that indicates the periodicity, while δ is the phase shift angle that acts as the equilibrium angle (ϕ^0_{ijkl}) when $n = 0$. Alternatively, $U_d(\phi_{ijkl})$ can defined by the Fourier series:

$$U_d(\phi_{ijkl}) = \frac{1}{2}\left[F_1\left(1 + \cos\phi\right) + F_2\left(1 - \cos(2\phi)\right) + F_3\left(1 + \cos(3\phi)\right) + F_4\left(1 - \cos(4\phi)\right)\right]$$
(2.32)

Eq. 2.32 can be also interconverted in the Ryckaert-Bellemans function, U_{RB}, which is often used for alkanes:

$$U_{RB}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos(\psi))^n, \quad \psi = \phi - \pi$$
(2.33)

by adopting the following definition for the coefficients C_n:

$$C_0 = F_2 + \frac{1}{2}(F_1 + F_3)$$

$$C_1 = \frac{1}{2}(-F_1 + 3F_3)$$

$$C_2 = -F_2 + 4F_4$$

$$C3 = -2F_3$$

$$C4 = -4F_4$$

$$C5 = 0$$
(2.34)

The improper dihedral angle, ξ_{ijkl}, is still defined as the angle between the planes ijk and jkl, but with the atom i connected to the other three atoms. The improper torsional angles are fundamental to keep groups planar and to retain enantiomers in their selected chiral form. Similarly to bonds and angles, the ξ_{ijkl} dihedrals are described by means of harmonic potential functions:

$$U_{id}(\xi_{ijkl}) = \frac{1}{2}k_{id,ijkl}\left(\xi_{ijkl} - \xi^0_{ijkl}\right)^2$$
(2.35)

It is worth noticing that if ξ^0_{ijkl} is equal to $\pm\pi$, and ξ_{ijkl} oscillates around ξ^0_{ijkl}, the discontinuity of the harmonic potential will become tricky. However, in practice, this is never a problem since the discontinuity is chosen at a distance of 180° from ξ^0_{ijkl}.

The non-bonded term, U_{nb}, should consider all the potential terms between non-bonded atoms. However, it is conventional to approximate U_{nb} by considering only the pairwise potential, thus the interaction of each pair of particles. These non-bonded interactions are described considering van der Waals and electrostatic terms. The van der Waals terms of the i, j atomic pairs can be properly modeled by the Lennard-Jones (LJ) potential:

$$U_{LJ}(r) = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \tag{2.36}$$

where ε is the depth of the potential well, r_{ij} is the distance between the particles, and σ is the distance at which the potential is zero. σ is also known as van der Waals radius since it gives a measurement of how close the two particles can get.

The electrostatic interactions between point charges can be modeled instead by the Coulomb potential:

$$U_C = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.37}$$

Both the potentials need long-range corrections (LRC) due to the cutoffs that are introduced with the PBC. For the Lennard-Jones potential, the LRC are applied only to the attractive term (second term in Eq. 2.36) because the repulsive one (first term in Eq. 2.36) decays so fast that it becomes negligible at the cutoff distance, r_{cut}. Therefore, the corrections for the LJ potential become:

$$U_{LJ,LRC}(r) = -\frac{8\pi}{3} N\rho\varepsilon\sigma^6 r_{cut}^{-3} \tag{2.38}$$

with ρ equals to the density.

Unfortunately, a similar integration does not work for the Coulomb potential. In this case, the interactions between the particle and its images in the periodic cells should be explicitly included in the calculation of the potential:

$$U_{C,tot} = \frac{1}{4\pi\varepsilon_0} \frac{1}{2} \sum_{\mathbf{n}} \sum_{ij} \frac{q_i q_j}{r_{ij,\mathbf{n}}} \tag{2.39}$$

The vector $\mathbf{n}$ is the index vector of the box that we assume as cubic for simplicity. Summing up all the contributions in Eq. 2.39, a slow and conditional convergence is obtained. However, the direct sum approach is clearly too expensive to be computed. A first alternative was proposed by Ewald in 1921. [120] Here, the idea is to divide the long-range interaction in a short-range contribution and a long-range contribution without singularities. The former term is evaluated in the real space, while the latter contribution in the Fourier space. This splitting results then in a quick convergence of the two terms. However, the Ewald summation scales as N^2 or $N^{3/2}$, at its best, thus becoming inefficient for large systems. The performance has been improved with the development of the Particle-mesh Ewald (PME) methods. [121, 122] The PME techniques interpolate the atomic charges on grid points, which are then transformed with a 3D Fast Fourier Transform algorithm. The PME algorithm scales as $N\log(N)$, significantly speeding up the computation of long-range interactions. Therefore, this method is widely used in large-scale MD simulations.

All the potential terms discussed in this section are collected together in the so-called Force Field (FF). Over the years, several FFs have been proposed to simulate in particular organic

and bio-molecular systems. Parameters compatible with the *standard* FFs have also been developed to treat other classes of systems, such as nanoclusters. An overview on the most used FFs, some of their extensions, as well as on the models that are usually employed to treat the water solvent, is reported in the following section.

2.1.7 Force Fields and Water Models

In the context of MD simulations, the FF refers to the ensemble of functional forms and parameters describing the potential energy terms discussed in section 2.1.6. Such parameters can be derived from experiments, quantum-mechanical calculations, or their combination. The quality of this derivation is usually assessed by the ability of the FF to reproduce the thermodynamic properties of bulk solutions. It is possible to distinguish all-atom and united-atom FFs. In the former case, the interatomic potential parameters are defined for each atom type, including light atoms as hydrogen. In the united-atom case, atoms of the same chemical group (*e.g.* H and C atoms of methyl groups) are treated as one interaction center.

AMBER, [123] CHARMM, [124] and OPLS [125] are three well-known FFs that were developed to simulate, mostly, bio-molecules. All these FFs adopt similar potential functions, with CHARMM being the only one where the Urey-Bradley term is added to account for the 1-3 interactions. AMBER, which stands for Assisted Model Building with Energy Refinement, calculates the partial atomic charges by means of the restrained electrostatic potential (RESP) fitting. In this FF, the 1-4 scaling factor is set to 0.5 for the Lennard-Jones potential, while it is equal to 0.833 for the Coulomb potential. In the CHARMM FF (Chemistry at HARvard Macromolecular Mechanics), the partial charges are empirically adjusted in order to reproduce certain experimental properties. Here, the 1-4 scaling factor is not included, hence it is unitary. The Optimized Potentials for Liquid Simulations (OPLS) FF employs partial atomic charges that are empirically optimized and a 1-4 scaling factor of 0.5 for both the non-bonded potential functions.

These FFs can be adjusted to treat other classes of systems. This is the case of thiolate-protected gold nanoclusters which are composed of a metallic core, Au-S staples, and organic tails. While the parameters of the organic ligands could still be found in the available *standard* FFs, this is usually not true for both the core and Au-S architectures. The process to parameterize these two motifs is similar to that employed for organic and bio-systems. Indeed, the experimental data are combined with quantum-chemical calculations, while the partial atomic charges are usually obtained through a RESP fitting procedure. One of the first FFs for RS-AuNCs was developed by Corni and collaborators, who calculated OPLS-compatible parameters for NCs of different size and geometry. [126] A couple of years later, Häkkinen and co-workers similarly proposed the parameters for different RS-AuNCs, but compatible with the AMBER FF. [127]

Several models have been proposed to account for the potential parameters of the water molecules, since most of the MD simulations are performed in aqueous environment. These models can be distinguished by: (i) the number of sites, thus the interaction points, (ii) the

rigidity or flexibility, and (iii) the inclusion of polarization effects. Normally, water models have three or four interaction points. Three-site models have the three interaction points corresponding to the three atoms of water. Each site has a point charge, and the Lennard-Jones parameters are also included for the oxygen atom. Famous examples of three-site models are the TIP3P, [128] SPC, [129] and SPC/E [130] water models. All these models use a rigid geometry which matches that of the actual water molecule. An exception is made by the SPC family which employs a HOH angle of 109.47° (ideal tetrahedral shape), instead of the observed one of 104.5°. The SPC/E model adds an average polarization correction in order to improve the density and diffusion constant of the SPC model. A variant of the rigid SPC model is the flexible SPC (F/SPC) model [131] which gives the correct density and dielectric permittivity of water. Four-site models consider a 4th interaction point by adding a dummy atom with a negative charge along the bisector of the HOH angle. This way, the electrostatic distribution around the molecule is improved. A well-known example in this category is the TIP4P model which is more expensive than the corresponding three-site TIP3P model, but it gives both better density and quadrupole moment values. [128]

2.2   Essential Dynamics

MD simulations are commonly used to span the conformational landscape of large and complex systems. However, the study of the conformational transitions, which are defined by the internal generalized degrees of freedom, can be very challenging. Moreover, the ergodic sampling over long time-scales can be limited by the presence of large free energy barriers, which define the conformational transitions. A solution is offered by the collective coordinates, which identify a low-dimensional subspace where the significant motions are expected to take place. Among the different approaches that can identify these collective motions, the most widely used are the normal mode analysis (NMA) [132, 133] and the principal component analysis (PCA). [134, 135] NMA is a harmonic analysis, particularly employed to probe the motions in biological systems. The NMA technique assumes that the conformational energy surface is defined by a single parabolic energy minimum. Here, the independent normal modes are obtained by diagonalizing the mass-weighted Hessian matrix of the energy minimum configuration. Despite its wide range of applications, its harmonic approximation of the conformational energy surface is known to be incorrect. Indeed, especially for proteins, many evidences have proved that the conformational energy landscape has several minima separated by energy barriers of various heights. [136, 137]

PCA can overcome some of the NMA limitations since it does not rely on the assumption of a harmonic potential. Therefore, it can be used to study the degree of anharmonicity in an MD simulation. This is particularly relevant for the protein case, where the principal modes of the collective motions are mostly dominated by anharmonic fluctuations. [135, 138] In the PCA method, the collective variables are obtained by diagonalizing a non-mass-weighted correlation matrix for a set of observables. Considering the covariance matrix of the atomic

coordinates, the largest part of the positional fluctuations results to be concentrated in correlated motions which are restrained in a subspace of few degrees of freedom. This subspace is termed as *essential* subspace. Outside of this subspace, the degrees of freedom are associated to much less relevant fluctuations and the motion can be considered as physically constrained. This way, the *essential* subspace is the only one required to describe the anharmonic internal motions, thus the conformational transitions. [139] This PCA application is known as essential dynamics (ED) analysis. [135]

Let us consider the dynamics of an equilibrated system composed of N atoms at a temperature T, and let us assume that its trajectory is available from an MD simulation. The overall roto-translational motions are removed due to their irrelevance to the object of this analysis, *i.e.* the internal motions. This operation can be done by least square fitting to a reference structure. Now, the internal motion can be described by a trajectory $\mathbf{x}(t)$, where $\mathbf{x}$ is the column vector containing the $3N$ atomic coordinates. The covariance matrix C of the positional deviations can be then expressed as:

$$C = \langle (\mathbf{x} - \langle \mathbf{x} \rangle) (\mathbf{x} - \langle \mathbf{x} \rangle)^T \rangle \tag{2.40}$$

where $\langle \mathbf{x} \rangle$ is the position averaged over time. Therefore, C is a symmetric matrix which can be diagonalized by an orthogonal coordinate transformation matrix T:

$$\mathbf{x} - \langle \mathbf{x} \rangle = T\mathbf{q} \tag{2.41}$$

$$C = T\Lambda T^T \tag{2.42}$$

The columns of T correspond to the eigenvectors belonging to the eigenvalues (λ_i) of the diagonal matrix $\Lambda = \langle \mathbf{q}\mathbf{q}^T \rangle$. These eigenvectors represent the directions along which the internal motion occurs, while the eigenvalues λ_i represent the corresponding mean square fluctuations. The eigenvectors are usually sorted to have the corresponding eigenvalues in a descending order. This way, the first eigenvectors and their eigenvalues are associated to the largest positional deviations and fluctuations, respectively. Therefore, a limited subset of the first n (with $n << 3N$) eigenvalues account for most of the fluctuations we are interested in.

The total $\mathbf{q}$- space can be now divided into an *essential* subspace ($q_1, ..., q_n$) and the remaining space ($q_{n+1}, ..., q_{3N}$). It can be demonstrated that the coordinates of the latter space effectively behave as constraints, without a significant contribution to the positional fluctuations. [135] Therefore, they can be set $= 0$ to approximate the mechanics in the *essential* subspace. The dynamics of the system of interest can be subsequently reformulated in this subspace by applying the transformation matrix T to the equations of motion of $\mathbf{x}$:

$$M'\ddot{\mathbf{q}} = T^T M \ddot{\mathbf{x}} = -T^T \nabla_x V(\mathbf{x}) = -\nabla_q V(\mathbf{q}) \tag{2.43}$$

where M is the diagonal matrix of the atomic masses, M' is the transformed mass tensor ($M' = T^T M T$), and V is the potential. In Eq. 2.43, the potential gradient was transformed by

using the following relation:

$$\frac{\partial V(\mathbf{x})}{\partial x_i} = \sum_J \frac{\partial V(\mathbf{q})}{\partial q_j} \frac{\partial q_j}{\partial x_i} = \sum_j T_{ij} \frac{\partial V(\mathbf{q})}{\partial q_j} \tag{2.44}$$

Now, V can be expressed using only the coordinates in the *essential* subspace since those of the remaining space were set as identically equal to 0. Therefore, the equations of motion are entirely restricted to the *essential* subspace for the approximated constrained system. Moreover, in this subspace such equations require only the upper left $n \times n$ block of M'. This means that the system dynamics can be actually reduced to a very low-dimensional *essential* subspace, thus bringing to a significant simplification of its study. [135] Therefore, using a quite restricted set of new observables, as defined by the *essential* eigenvectors, the properties of the original ones can be mostly reconstructed. [139]

Defining μ_i as the $i - th$ column of T ($i - th$ eigenvector of C), then the original trajectory can be projected on this (essential) eigenvector, obtaining the $i-$principal component, p_i:

$$p_i = \mu_i \left(\mathbf{x} - \langle \mathbf{x} \rangle \right) \tag{2.45}$$

It is worth noting that the variance $\langle p_i^2 \rangle$ is equal to the λ_i eigenvalue.

In all the works of this PhD project, the first pair of eigenvectors and associated eigenvalues were considered to account for most of the internal fluctuations, *i.e.* conformational transitions. This way, the resulting principal components provide us with a 2D conformational landscape relatively easy to investigate for the extraction of the most probable conformers. The selected structures will be then employed in the further quantum-chemical calculations.

Over the years, the ED analysis has been extensively used to characterize functional motions as well as the free energy surface of various biological systems. [140–142] However, there has been some debate regarding the possibility of robustly characterizing the essential coordinates over the time-scale that is accessible to MD simulations. Nevertheless, several works, including ours, have demonstrated that a stable eigenspace of principal components can be reached in the range of nanoseconds. [143, 144] Moreover, this post-processing tool can be used not only to analyze classical MD trajectories but also those generated with other approaches, such as enhanced sampling techniques (*e.g.,* parallel tempering, [145] bias exchange, [146] umbrella sampling, [147] and metadynamics [148]). As suggested by their name, the enhanced sampling methods increase the efficiency of sampling by addressing the problem of reaching all the relevant conformational sub-states. It is common indeed that the molecular phenomena of interest occur on a time-scale which cannot be reached by simply employing brute force MD simulations. However, such approaches have not been tested in this thesis since the systems under investigation have relatively short relaxation times, thus making classical MD the best compromise between accuracy and computational cost. It is also worthy of note that techniques such as the umbrella sampling require a predefined reaction coordinate to calculate the free energy surface, which is extremely difficult to define for systems such as

those studied in this thesis. This problem is entirely avoided instead with the ED analysis, providing us with a straightforward strategy to perform the conformational analysis.

2.3   Density Functional Theory

The Density Functional Theory (DFT) is a well-consolidated quantum-mechanical approach for the description of the ground-state electronic structure of many-body systems as atoms, molecules, or solids. [149] Over the last few decades, DFT has been largely employed to calculate the electronic structure of medium and large systems due to its good compromise between accuracy and computational cost. Indeed, while in *ab initio* methods, such as Hartree-Fock (HF) or post HF, the electronic structure is described by the multi-electronic wave function, Ψ, in DFT, it is rigorously described by the electronic density, $\rho(\mathbf{r})$. Therefore, the $3N$ spatial variables needed to treat N-electron systems in *ab initio* methods are reduced to only 3 when working with $\rho(\mathbf{r})$. Furthermore, the computational cost of DFT scales as $O(N^3)$, where N is the number of atoms of the system, while it scales at least as $O(N^4)$ for *ab initio* approaches.

DFT is based on the two theorems proposed by Hohenberg and Kohn (HK) in 1964, [150] which prove the one-to-one correspondence between the ground-state density and the external potential, V_{ext}. However, the original HK theorems work only for non-degenerate ground states and V-representable densities. Some years later, these limitations were overcome by the constrained-search formulation introduced by Levy and Lieb. [151, 152]

Let us start by defining the fundamental object of DFT, that is, the functional, F, as the application from the space of functions to the space of scalars (real or complex numbers):

$$F[\rho] : f \rightarrow \mathbb{R} \tag{2.46}$$

In the specific case of DFT, it is useful to consider the application from the space of wave functions to the space of electronic densities, as shown in the following figure.
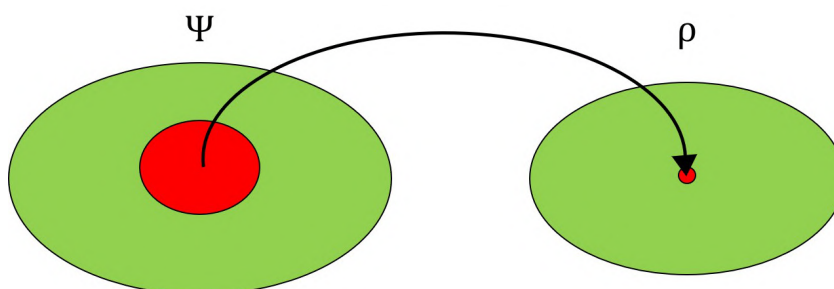

FIGURE 2.3: Graphical representation of the application from the wave function space (domain) to the electron density space (co-domain).

The functional formalism, in agreement with the Levy-Lieb formulation, will be used to treat the terms of a generic Hamiltonian, within the Born-Oppenheimer approximation, [153]

for a system of N interacting electrons:

$$H = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N} V_{ext}(\mathbf{r}) = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} \tag{2.47}$$

The first term in Eq. 2.47, $\hat{T}$, corresponds to the kinetic energy of electrons, the second one, $\hat{V}_{ee}$, represents the exact potential energy between the pairs of electrons, while the latter term, $\hat{V}_{ext}$, includes both the electron-nucleus and the external potentials. Since the former contribution is usually predominant, V_{ext} assumes the following expression:

$$V_{ext}(\mathbf{r}) = \sum_{K}^{Nucl} -\frac{Z_K}{|\mathbf{r}_i - \mathbf{R}_K|} \tag{2.48}$$

The first functional that will be considered, $F[\rho]$, takes into account the first two Hamiltonian terms as follows:

$$F[\rho] = \min_{\Psi \to \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle = \langle \Psi_{min} | \hat{T} + \hat{V}_{ee} | \Psi_{min} \rangle \tag{2.49}$$

The expectation value of the external potential energy can be expressed instead in its classical form:

$$\langle \Psi | \hat{V}_{ext} | \Psi \rangle = \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \tag{2.50}$$

Eq. 2.49 and Eq. 2.50 can be then collected together in a single functional, $E[\rho]$:

$$E[\rho] = F[\rho] + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \tag{2.51}$$

Starting from the definition of the two functionals $F[\rho]$ and $E[\rho]$, the variational theorem can be demonstrated for DFT. It can be proved that: (i) the energy functional $E[\rho]$ associated to any electron density is greater than, or equal to, the ground-state energy, (ii) the energy functional $E[\rho]$ associated to the electron density of the ground-state corresponds to the energy of the ground-state:

$$E[\rho] \geq E_{GS} \tag{2.52}$$

$$E[\rho_{GS}] = E_{GS} \tag{2.53}$$

The connection between variational principle and DFT was actually stated in the second HK theorem. [150]

Minimizing $E[\rho]$ with respect to ρ, both the density and the energy of the ground-state are obtained. Unfortunately, this approach cannot be used in practice since $F[\rho]$ is unknown.

However, it furnishes a formal justification of the theory and a rigorous definition of the objects introduced above. Two main strategies can be adopted to solve the problem of minimizing $E[\rho]$: (i) the Thomas-Fermi (TF, direct) approach, and (ii) the Kohn-Sham (KS, indirect) approach (see section 2.3.1).

The TF approach was conceived by Thomas and Fermi in 1927, [154, 155] more than 30 years before the formulation of the HK theorems. This statistical method was used to approximate the distribution of electrons in an atom as uniform (*i.e.* homogeneous electron gas model). Despite several limitations, the method has the great merit of being the first one to treat the ground-state energy as a functional of the electronic density instead of the wave function. The TF approach is considered indeed as the precursor of the modern DFT. Therefore, the functional for the kinetic energy, $T[\rho]$:

$$T[\rho] = \langle \Psi_{min} | \hat{T} | \Psi_{min} \rangle \tag{2.54}$$

and the functional for the potential of electronic pairs, $V_{ee}[\rho]$:

$$V_{ee}[\rho] = \langle \Psi_{min} | \hat{V}_{ee} | \Psi_{min} \rangle \tag{2.55}$$

can be calculated within this approximated model, obtaining the following TF functional for the total energy:

$$E_{TF}[\rho] = \frac{3}{10}(3\pi^2)^{2/3} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int \rho(\mathbf{r}) V_{ext} d\mathbf{r} \tag{2.56}$$

The first term in Eq. 2.56 corresponds to the TF functional of the kinetic energy ($T_{TF}[\rho]$), with $\rho(\mathbf{r})$ being the density of the homogeneous gas in that point. The second term instead is the classic Coulomb repulsion energy among pairs of electronic densities ($J[\rho]$). $E_{TF}[\rho]$ can be now minimized with the constraint that the density gives the correct number of electrons, which is constant in the real system. The accuracy of the TF model is limited by the inaccurate representation of the kinetic energy and the neglect of the exchange and correlation energy. Currently, the computational methods are mostly based on the KS formulation of DFT that is discussed below.

2.3.1 Kohn-Sham approach

In 1965, Kohn and Sham proposed their indirect approach which put the HK theorems into practical use, thus making the DFT calculations affordable. [156] The KS approach considers a fictitious system of non-interacting electrons, that is $V_{ee} = 0$, whose electronic density is equal to that of the real system (interacting electrons). Therefore, the impracticable problem of treating a many-body system of interacting electrons within an external potential is reduced to a tractable one, where the non-interacting electrons move within an effective potential. This fictitious potential is termed as KS potential, V_{KS}, and it is chosen in a way that guarantees the

equivalence between the fictitious and the real electronic density.

Starting from these assumptions, the KS Hamiltonian, H_{KS}, can be written as a sum of mono-electronic Hamiltonians:

$$H_{KS} = \sum_{i=1}^{N} \left(-\frac{1}{2} \nabla_i^2 \right) + \sum_{i=1}^{N} V_{KS}(\mathbf{r}_i) = \sum_{i=1}^{N} h_{ks}(i) \tag{2.57}$$

The ground-state wave function of the KS system can be constructed with a single Slater determinant as:

$$\Psi_{KS} = \frac{1}{\sqrt{N!}} \det|\chi_1 \chi_2 ... \chi_N| \tag{2.58}$$

where N is the number of electrons, and χ_i is the $i-th$ spin-orbital satisfying the time-independent Schrödinger equation:

$$h_{KS}\chi_i = \varepsilon_i \chi_i \tag{2.59}$$

Solving Eq. 2.59, the (KS) orbitals and the orbital energies, ε_i, are obtained. The KS density can be now calculated from the spatial part, $\varphi_i(\mathbf{r})$, of the spin-orbitals:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \varphi_i^*(\mathbf{r}) \varphi_i(\mathbf{r}) \tag{2.60}$$

As stated above, V_{KS} must be chosen in a way that gives the KS density equals to the real one. Similarly to the TF model, this goal is achieved by minimizing the energy functional $E[\rho]$ with the following constraint on the density:

$$\int \rho(\mathbf{r})d\mathbf{r} = N \tag{2.61}$$

where N corresponds again to the number of electrons of the system.

Let us start from the real, interacting system, whose energy functional (see Eqs. 2.51, 2.54, and 2.55) can be also formulated as:

$$E[\rho] = T_0[\rho] + J[\rho] + E_{xc}[\rho] + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r})d\mathbf{r} \tag{2.62}$$

Here, $T_0[\rho]$ is the non-interacting kinetic energy functional:

$$T_0[\rho] = -\frac{1}{2} \sum_{i=1}^{N} \langle \chi_i | \nabla_i^2 | \chi_i \rangle \tag{2.63}$$

while $J[\rho]$ is the classical electrostatic interaction that was already shown in Eq. 2.56. Instead, $E_{xc}[\rho]$ is the exchange and correlation functional which is defined as follows:

$$E_{xc}[\rho] = T[\rho] - T_0[\rho] + V_{ee}[\rho] - J[\rho] \tag{2.64}$$

This functional is unknown, but it can be approximated with different models, *e.g.* electron gas models, depending on the properties of the system one is interested in. Moreover, this term is smaller with respect to the other objects, thus making the approximation, then the approach, properly working. Models that are usually employed to treat the $E_{xc}[\rho]$ functional will be investigated in detail in the next section 2.3.2. The constrained (Eq. 2.61) minimization of the energy functional in Eq. 2.62 can be now performed exploiting the method of Lagrange multipliers:

$$\frac{\delta}{\delta\rho}\left(E[\rho] - \lambda \int \rho(\mathbf{r})d\mathbf{r}\right) = 0$$

$$\frac{\delta T_0}{\delta\rho} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{xc}}{\delta\rho} + V_{ext}(\mathbf{r}) = \lambda$$

$$(2.65)$$

The procedure is then repeated to minimize the energy functional of the fictitious KS system, $E_{KS}[\rho]$:

$$\frac{\delta}{\delta\rho}\left(E_{KS}[\rho] - \lambda \int \rho(\mathbf{r})d\mathbf{r}\right) = 0$$

$$\frac{\delta T_0}{\delta\rho} + V_{KS}(\mathbf{r}) = \lambda$$

$$(2.66)$$

The same solution is imposed for the two systems due to the equivalence of the electronic densities ($\rho_{KS} = \rho_{real}$). Consequently, the following expression is obtained for the KS potential:

$$V_{KS}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{xc}}{\delta\rho} + V_{ext}(\mathbf{r}) \qquad (2.67)$$

The mono-electronic KS equations in Eq. 2.59 can be now solved by putting this latter expression into the KS Hamiltonian (Eq. 2.57).

It is worth noticing that the KS potential depends on the electronic density, which in turn depends on the spatial orbitals $\varphi(\mathbf{r})$ that are obtained by solving the KS equations. Therefore, a Self Consistent Field (SCF) procedure must be employed to treat these mono-electronic equations, similar to that employed in the HF method. The SCF approach, whose flowchart is schematized in Figure 2.4, starts with an initial guess of the density, $\rho^{(0)}(\mathbf{r})$, which is generally chosen as the sum of the atomic densities to avoid oscillations along the procedure. The tested density is used to calculate the KS potential first and solve the mono-electronic equations later. The resulting molecular orbitals are used to calculate the new electronic density, which is then compared with that calculated in the previous iteration. The convergence is reached when the difference between the two density values is below a certain threshold. Meanwhile, the potential is updated from the orbitals of the previous cycle at each iteration.
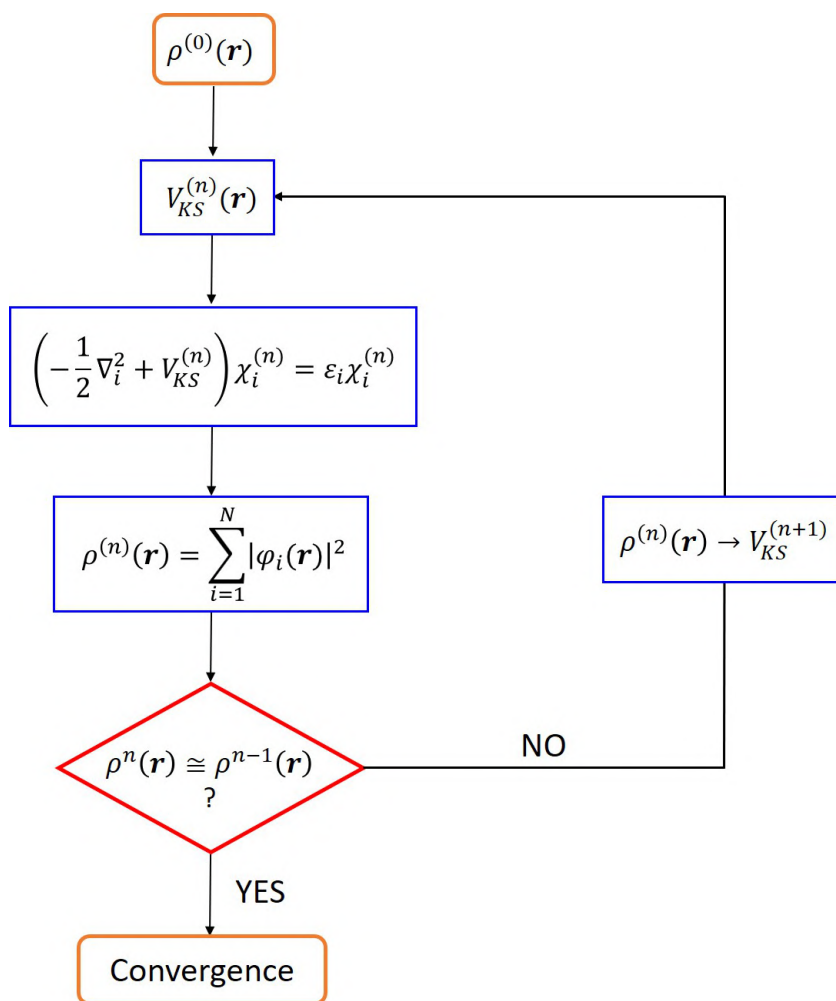
FIGURE 2.4: Flowchart of the SCF procedure employed to solve the KS equations.

The φ_i molecular orbitals of the KS equations are obtained with the Linear Combination of Atomic Orbitals (LCAO) formulation. In LCAO, the orbitals are expressed as linear combinations of basis functions:

$$\varphi_i(\mathbf{r}) = \sum_{k=1}^{basis} \nu_k(\mathbf{r}) C_{ki} \tag{2.68}$$

and the problem of obtaining the KS eigenfunctions and eigenvalues is reduced to a generalized diagonalization one:

$$H_{KS} C = S C E \tag{2.69}$$

In Eq. 2.69, S is the overlap matrix that is given in input, together with the KS Hamiltonian, while the coefficient matrix C, and the energy matrix E are obtained in output.

There are two different philosophies to build the H_{KS} matrix: the numerical integration and the analytical integration. The former approach, employed in the Amsterdam Modelling Suite (AMS) code, [157] uses Slater basis sets, whereas the latter approach, employed in the

Gaussian code, [158] works with Gaussian basis sets. Slater-type orbitals describe more naturally the atomic orbitals, but Gaussian-type orbitals allow faster calculations due to the use of analytical integrals. In the AMS code, which was used in all the works of this PhD thesis, a technique known as *density fitting* is adopted to improve the performance of the calculations. Here, the density is approximated with a fitting density, $\tilde{\rho}(\mathbf{r})$, to reduce the sums that need to be computed:

$$
\rho(\mathbf{r}) = \sum_{i=1}^{N} \varphi_i^*(\mathbf{r})\varphi_i(\mathbf{r})
$$

$$
= \sum_{i=1}^{N} \sum_{j=1}^{basis} \sum_{k=1}^{basis} C_{ji}^* C_{ki} \nu_j^*(\mathbf{r})\nu_k(\mathbf{r}) \tag{2.70}
$$

$$
\cong \tilde{\rho}(\mathbf{r}) = \sum_{m}^{fit} f_m(\mathbf{r})a_m
$$

In Eq. 2.70, f_m are the auxiliary fitting functions, while a_m are the coefficients defined to optimize $\tilde{\rho}(\mathbf{r})$. Therefore, they are obtained imposing the following condition:

$$
\min \int (\rho(\mathbf{r}) - \tilde{\rho}(\mathbf{r}))^2 \, d\mathbf{r} \tag{2.71}
$$

and constraining the fitting density:

$$
\int \tilde{\rho}(\mathbf{r})d\mathbf{r} = N \tag{2.72}
$$

2.3.2   Exchange and correlation functionals

As stated in the previous section 2.3.1, the exchange-correlation (xc) energy functional is the unknown term of the KS equations. Therefore, the xc potential, V_{xc}, must be approximated using a model, which is chosen according to the system and properties under investigation. The fact that both exchange and correlation keep electrons apart has brought to describe the xc contribution in terms of a hole that surrounds each electron, preventing the others from approaching it. The so-defined *xc hole* can be explained considering the joint probability of finding, simultaneously, an electron at point $\mathbf{r}_1$ and another electron at point $\mathbf{r}_2$. Therefore, the *xc hole* can be described in terms of the pair-correlation function, $h_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ as:

$$
h_{xc}(\mathbf{r}_1, \mathbf{r}_2) = h_x^{s1=s2}(\mathbf{r}_1, \mathbf{r}_2) + h_c^{s1,s2}(\mathbf{r}_1, \mathbf{r}_2) \tag{2.73}
$$

where the exchange term (*Fermi hole*) acts only on electrons with the same spin value, while the correlation term (*Coulomb hole*) affects all electrons, independently from their spin value.

Starting from the pair density function, $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$:

$$
\begin{aligned}
\rho_2(\mathbf{r}_1, \mathbf{r}_2) &= \frac{N(N-1)}{2} \int |\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)|^2 \mathrm{d}\mathbf{r}_3 \mathrm{d}\mathbf{r}_4 ... \mathrm{d}\mathbf{r}_N \\
&= \frac{1}{2} \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) [1 + h_{xc}(\mathbf{r}_1, \mathbf{r}_2)]
\end{aligned}
\tag{2.74}
$$

the *xc hole*, $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$, can be defined as:

$$
\begin{aligned}
\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) &= \rho(\mathbf{r}_2) h_{xc}(\mathbf{r}_1, \mathbf{r}_2) \\
&= \frac{2\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - \rho(\mathbf{r}_2)
\end{aligned}
\tag{2.75}
$$

$\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ represents the probability of finding two electrons within the volume elements $\mathrm{d}\mathbf{r}_1$ and $\mathrm{d}\mathbf{r}_2$, simultaneously. Instead, $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ represents the difference between the conditioned and unconditioned probability. It can be also written in terms of correlation function as:

$$
\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = g(\mathbf{r}_1, \mathbf{r}_2) - \rho(\mathbf{r}_2)
\tag{2.76}
$$

Now, the expectation value of the repulsive electron-electron potential becomes:

$$
\begin{aligned}
\langle V_{ee} \rangle &= \frac{1}{2} \left[\int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 + \int \int \frac{\rho(\mathbf{r}_1)\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \right] \\
&= J[\rho] + E_{xc}[\rho]
\end{aligned}
\tag{2.77}
$$

By introducing models to derive an explicit expression for $E_{xc}[\rho]$, it is possible to define the xc energy density, ε_{xc}, which can be further split into its exchange and correlation contributions:

$$
\varepsilon_{xc} = \varepsilon_x + \varepsilon_c
\tag{2.78}
$$

The oldest model for the xc potential is the Local Density Approximation (LDA) which describes the electrons with the Uniform Electron Gas model. Therefore, the whole space is divided into small volume elements in which the electronic density is treated as constant. Subsequently, the xc energy functional depends only on the value of the density at each point of the space and is obtained by integrating ε_{xc} at a given density $\rho(\mathbf{r})$:

$$
E_{xc}^{LDA}[\rho(\mathbf{r})] = \int \varepsilon_{xc}^{LDA}[\rho(\mathbf{r})]\rho(\mathbf{r})\mathrm{d}\mathbf{r}
\tag{2.79}
$$

The exchange term of ε_{xc}^{LDA} has the following analytical form:

$$
\varepsilon_x^{LDA}[\rho(\mathbf{r})] = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \rho^{1/3}(\mathbf{r})
\tag{2.80}
$$

and the exchange potential, v_x^{LDA}, becomes:

$$v_x^{LDA}[\rho(\mathbf{r})] = - \left(\frac{3}{\pi}\right)^{1/3} \rho^{1/3}(\mathbf{r}) \tag{2.81}$$

Actually, a first attempt to formulate a solution for the local v_x was made by Slater, who proposed the X_α method in 1951. [159] In order to reduce the computational complexity of the HF method, he introduced the following expression for the exchange potential:

$$v_{X\alpha}[\rho(\mathbf{r})] = -\alpha\frac{3}{2} \left(\frac{3}{\pi}\right)^{1/3} \rho^{1/3}(\mathbf{r}) \tag{2.82}$$

The Slater's exchange parameter, α, was imposed equal to one in the original method. For the correlation term of ε_{xc}^{LDA}, the solution is much less trivial. It can be approximated by fitting a set of discrete values, previously obtained via a Quantum Monte Carlo procedure, by means of complicated analytical expressions. [160]

Despite its simplicity, the LDA method works very well in many cases, in particular for geometry optimizations or to describe material properties as lattice constants of solids, vibrational frequencies, dipole moments, and Fermi surfaces of bulk materials. However, it has a tendency to overestimate polarizabilities, as well as to underestimate electron affinities and ionization potentials. Moreover, the resulting KS eigenvalues are too low in magnitude because the xc potential goes to zero exponentially instead of following the correct $-\frac{1}{r}$ behavior at long distances.

Some of the LDA limitations are overcome by models within the Generalized Gradient Approximation (GGA), which accounts for the non-homogeneity of the electronic density by considering its gradient, $\nabla\rho$. Herein, the xc energy functional assumes the following expression:

$$E_{xc}^{GGA}[\rho(\mathbf{r})] = \int \varepsilon_{xc}^{GGA}(\rho, \nabla\rho)\rho(\mathbf{r})\mathrm{d}\mathbf{r} \tag{2.83}$$

GGA functionals are non-local since the xc energy density ε_{xc} at a given point $\mathbf{r}$ is defined by both the electronic density in that point ($\rho(\mathbf{r})$) and the electronic density in the neighboring points ($\rho(\mathbf{r} + \Delta\mathbf{r})$).

Among the widely used GGA functionals, we can mention the forms proposed by Becke (B88), [161] Perdew, Burke, and Enzerhof (PBE), [162] and van Leeuwen, and Baerends (LB94). [163] The latter one was employed in several works of this PhD project due to its ability of reproducing the correct asymptotic Coulomb decay behavior.

GGA allows improved calculations of the energy-related properties with respect to LDA, especially for systems characterized by fast variations of the electronic density. However, both GGA and LDA perform poorly on materials with localized and strongly correlated electrons (*e.g.* transition metal oxides, and rare-earth elements).

Variations and refinements of the GGA family have been proposed, such as meta-GGA functionals that use higher-order derivatives of the electronic density for its approximation in

the ε_{xc} energy. [164] Other examples of more accurate GGAs are the hyper-GGA and the Generalized Random Phase approximation. [165, 166] Hyper-GGA functionals include the exact exchange energy density, while in the latter approximation the contribution of the unoccupied orbitals is taken into account.

A relevant family of functionals was introduced by Becke in 1993 [167] and is termed as hybrid xc functionals. These potentials incorporate a portion of the exact non-local HF exchange providing a good description of short-range electron-electron interactions, hence a very accurate description of short-range properties. Hybrid functionals can be divided into purely theoretical and semi-empirical ones, depending on whether their parameters are fitted or not with experimental data. Nowadays, one of the most used functional is B3 (*i.e.* Becke's three parameter exchange functional) in combination with the LYP correlation functional. [168] The resulting B3LYP functional belongs to the semi-empirical family and possesses the following expression:

$$E_{xc}^{B3LYP} = E_x^{LDA} + \alpha_0 \left(E_x^{HF} - E_x^{LDA} \right) + \alpha_x E_x^{B88} + \alpha_c E_c^{LYP} + (1 - \alpha_c) E_c^{VWN} \qquad (2.84)$$

α_0, α_x, and α_c are the three parameters defining the potential and are approximately equal to 0.2, 0.7, and 0.8, respectively. E_x^{B88}, and E_c^{LYP} are the B88 exchange, [161] and the Lee, Yang, and Parr correlation [168] GGA functionals, while E_x^{LDA} and E_c^{VWN} are both LDA functionals. B3LYP was used in several occasions in this PhD project due to its good accuracy, although it may require a significant computational effort because of the HF portion.

Hybrid functionals as B3LYP are often described as *global* hybrids because the portion of the exact exchange is constant all over the space. However, there are situations where it may be useful to go beyond this description, for instance when non-covalent systems are treated or weak interactions and charge-transfer excitations play a major role. In this cases, one can employ range-separated (RS) xc functionals that divide the space of electron-electron interactions into short-, and long-range regions by splitting the Coulomb operator into two parts:

$$\frac{1}{\mathbf{r}_{12}} = SR + LR = \frac{1 - \omega(\gamma \mathbf{r}_{12})}{\mathbf{r}_{12}} + \frac{\omega(\gamma \mathbf{r}_{12})}{\mathbf{r}_{12}} \qquad (2.85)$$

Here, $\mathbf{r}_{12}$ is the inter-electronic distance ($|\mathbf{r}_1 - \mathbf{r}_2|$), while ω is a continuous switching function that goes to 0 when $\mathbf{r}_{12}$ goes to 0 and goes to 1 when $\mathbf{r}_{12}$ becomes large. The rapidity of this switching is defined by the parameter γ.

Exact and DFT exchange have been applied in different ways in range separation, depending on the system of interest. For instance, it useful to employ HF exchange at short-range, and DFT exchange at long-range in the description of periodic systems, [169] while the opposite separation is adopted for molecular systems. [170]

The switching function ω mostly assumes the form of an error function (erf), which is quite attractive for its practical implementation. The erf function can be also combined with additional Gaussian functions. An alternative to the erf is the exponential function $e^{-\gamma \mathbf{r}_{12}}$, that combined with the $1/r_{12}$ operator gives the Yukawa potential. The Yukawa potential is a more

natural choice as a screened Coulomb potential, but it is not compatible with the Gaussian basis set which is widely employed in most of the quantum chemistry packages. In the AMS code, that works with Slater-type orbital functions, the Yukawa potential is used as switching function according to ref. [171].

Among the several RS functionals proposed, it is worth mentioning the CAM-B3LYP [69] and wB97X-D [70] xc functionals. For CAM-B3LYP, Eq. 2.85 is adjusted by introducing two extra parameters, α and β:

$$\frac{1}{\mathbf{r}_{12}} = \frac{1 - [\alpha + \beta \text{erf}(\gamma \mathbf{r}_{12})]}{\mathbf{r}_{12}} + \frac{\alpha + \beta \text{erf}(\gamma \mathbf{r}_{12})}{\mathbf{r}_{12}} \tag{2.86}$$

where the relations $0 \leq \alpha + \beta \leq 1$, $0 \leq \alpha \leq 1$, and $0 \leq \beta \leq 1$ should be satisfied. CAM-B3LYP is a Coulomb attenuated version of the hybrid B3LYP and comprises of 0.19 and 0.65 HF exchange at short-range and long-range, respectively. The intermediate region is described through the standard erf with $\gamma = 0.33$. Instead, wB97X-D, widely employed in this PhD work, is a RS version of Becke's 97 functional with additional dispersion corrections. This functional includes about 22% and 100% of exact exchange at short- and long-range, respectively. The wB97X-D functional has proven to be very accurate in describing non-covalent interactions, as solute-solvent ones, because of its corrections.

2.4 Time-Dependent Density Functional Theory

Although some DFT extensions for the treatment of excited states have been proposed (*e.g.* Transition Potential (TP) method), this theory only provides a description of the electronic ground-state. Therefore, its rigorous time-dependent (TD) extension, *i.e.* TDDFT, is introduced to study the excited states, hence to calculate spectral properties. The formal foundation of TDDFT is provided by the Runge-Gross (RG) theorem (1984), [172] which is the extension of the HK theorem (see section 2.3) to the time-dependent case. Indeed, this theorem demonstrates the one-to-one correspondence between TD densities and TD potentials. However, the RG theorem does not guarantee that the density of the full interacting system can be calculated using a much simpler non-interacting one, thus that a KS-like approach can be adopted. This is where the van Leeuwen theorem [173] steps in, opening the way towards a time-dependent KS approach. Indeed, this theorem guarantees the existence of non-interacting systems that reproduce the TD density of the real interacting systems.

The TDDFT equations will be discussed in their Casida's formulation [71] in section 2.4.2 and by means of the polTDDFT algorithm [93] in section 2.4.3. In both cases, the system is subjected to a small perturbation, hence not strongly deviating from its initial state. This means that the response can be calculated to the first order in the perturbation, working in the so-called linear response regime. The theoretical bases of the Linear Response Theory (LRT) are provided in the following section.

2.4.1 Linear Response Theory

Response Theory aims to describe how a system reacts to an external perturbation such as an applied electric or magnetic field. Considering a small TD perturbation acting on a system of N-interacting electrons, which satisfies the TD Schrödinger equation, the perturbed Hamiltonian becomes:

$$H(t) = H^{(0)} + \lambda H^{(1)}(t) \tag{2.87}$$

The superscripts in Eq. 2.87 indicate the order of the perturbation, while λ represents the parameter that switches it on. Similarly, the perturbed wave function can be written as:

$$\Psi^{'}(\mathbf{r}, t) = \Psi^{(0)}(\mathbf{r}) + \lambda \Psi^{(1)}(\mathbf{r}, t) \tag{2.88}$$

$H^{(0)}$ and $\Psi^{(0)}$ are the unperturbed Hamiltonian and wave function, respectively. In this approach, we assume to know all the solutions of the unperturbed system:

$$H^{(0)} \Psi_n^{(0)} = E_n^{(0)} \Psi_n^{(0)} \tag{2.89}$$

The zeroth-order wave functions can be factorized in their spatial and time-dependent components as:

$$\Psi_n^{(0)}(\mathbf{r}, t) = \varphi_n^{(0)}(\mathbf{r}) e^{-iE_n^{(0)}t} \tag{2.90}$$

These wave functions generally form a complete set that can be employed to expand the first-order perturbed wave function $\Psi^{(1)}$ as:

$$\Psi^{(1)}(\mathbf{r}, t) = \sum_{n \neq 0} c_n(t) \varphi_n^{(0)}(\mathbf{r}) e^{-iE_n^{(0)}t} \tag{2.91}$$

where $c_n(t)$ represents the $n - th$ time-dependent expansion coefficient. Now, the solution of the perturbed system becomes:

$$|\Psi_0^{'}\rangle = \varphi_0^{(0)}(\mathbf{r}) e^{-iE_0^{(0)}t} + \lambda \sum_{n \neq 0} c_n(t) \varphi_n^{(0)}(\mathbf{r}) e^{-iE_n^{(0)}t} \tag{2.92}$$

Considering the first-order in the expansion series and collecting the exponential terms, the TD Schrödinger equation assumes the following form:

$$H^{(1)} \Psi_0^{(0)} = i \sum_{n \neq 0} \dot{c}_n \varphi_n^{(0)} e^{-i\omega_{0n}t} \tag{2.93}$$

where $\omega_{0n} = E_n^{(0)} - E_0^{(0)}$, thus it is the difference between the energy of the n-excited state and of the ground-state. The $c_n(t)$ coefficients are obtained solving the following integral:

$$
\begin{aligned}
c_n(t) &= \int_{-\infty}^{t} \dot{c}_n \mathrm{d}t \\
&= -i \int_{-\infty}^{t} \langle \varphi_n^{(0)} | H^{(1)} | \varphi_0^{(0)} \rangle e^{i\omega_{0n}\tau} \mathrm{d}\tau \\
&= -i \langle n | A(\mathbf{r}) | 0 \rangle \int_{-\infty}^{t} F(\tau) e^{i\omega_{0n}\tau} \mathrm{d}\tau
\end{aligned}
\tag{2.94}
$$

where $H^{(1)} = A(\mathbf{r})F(t)$.

Let us focus now on studying the effect of the perturbation on a specific property of the system (the observable $B(t)$), that is the actual response of the system. This response to a perturbation A can be measured considering the difference between the expectation values of the perturbed and the initial functions:

$$
\delta \langle B(t) \rangle = \langle B(t) \rangle - \langle B \rangle_0 = \langle \Psi' | B | \Psi' \rangle - \langle \Psi^{(0)} | B | \Psi^{(0)} \rangle
\tag{2.95}
$$

By introducing the expression for the perturbed wave function and neglecting higher (> 1) perturbative terms (see Eq. 2.88), Eq. 2.95 becomes:

$$
\delta \langle B(t) \rangle = \langle \Psi^{(0)} | B | \lambda \Psi^{(1)} \rangle + \langle \lambda \Psi^{(1)} | B | \Psi^{(0)} \rangle
\tag{2.96}
$$

Now, the perturbed wave function can be expressed in terms of its expansion (see Eq. 2.91), explicating the TD coefficients as shown in Eq. 2.94:

$$
\delta \langle B(t) \rangle = -i \sum_{n \neq 0} \langle 0 | B | n \rangle \langle n | A | 0 \rangle \int_{-\infty}^{t} F(\tau) e^{-i\omega_{0n}(t-\tau)} \mathrm{d}\tau + c.c.
\tag{2.97}
$$

where $c.c.$ stands for complex conjugate.

It is convenient to define the time-correlation function, R, as:

$$
R(BA|t-\tau) = -i \sum_{n \neq 0} \left[\langle 0 | B | n \rangle \langle n | A | 0 \rangle e^{-i\omega_{0n}(t-\tau)} - \langle n | B | 0 \rangle \langle 0 | A | n \rangle e^{i\omega_{0n}(t-\tau)} \right]
\tag{2.98}
$$

in order to express Eq. 2.97 in the following compact way:

$$
\delta \langle B(t) \rangle = \int_{-\infty}^{t} R(BA|t-\tau) F(\tau) \mathrm{d}\tau
\tag{2.99}
$$

Therefore, the linear response function is the variation of the expected value $\langle B \rangle$ at time t caused by the perturbation A at time τ. Note that this function is defined only for $t > \tau$, according to the principle of causality. Moreover, the response function depends only on the distance between t and τ but not on their actual values. The dependency on the specific time

value is included in the TD potential, $F(t)$.

In LRT, it is actually more convenient to work in the frequency-domain considering the frequency components of the TD potential, which can be obtained by means of a Fourier transform: $F(t) \rightarrow f(\omega)$. Before that, let us find an expression for $F(t)$ ensuring that the perturbation acts on an unperturbed state. This goal is simply achieved by introducing an adiabatic "switch on" factor, $e^{\varepsilon t}$, such that:

$$\lim_{t \to -\infty} e^{\varepsilon t} = 0 \tag{2.100}$$

while for finite t, $e^{\varepsilon t} \approx 1$. This way, the perturbation is gradually switched on, avoiding transient effects.

The Fourier transform of $F(t)$ is given by:

$$F(t) = e^{\varepsilon t} \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i\omega t} \mathrm{d}\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega + i\varepsilon)t} \mathrm{d}\omega \tag{2.101}$$

Substituting this expression in Eq. 2.99, the response becomes:

$$\delta\langle B(t)\rangle = \frac{1}{2\pi} \int_{-\infty}^{t} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega + i\varepsilon)\tau} R(BA|t-\tau) \mathrm{d}\tau \mathrm{d}\omega \tag{2.102}$$

It is possible to integrate over τ to keep only the frequency dependence:

$$\begin{cases} \int_{-\infty}^{t} e^{-i(\omega + i\varepsilon + \omega_{0n})\tau} e^{i\omega_{0n}t} \mathrm{d}\tau = -\dfrac{e^{-i(\omega + i\varepsilon)t}}{i(\omega + i\varepsilon + \omega_{0n})} \\[4mm] \int_{-\infty}^{t} e^{-i(\omega + i\varepsilon - \omega_{0n})\tau} e^{-i\omega_{0n}t} \mathrm{d}\tau = -\dfrac{e^{-i(\omega + i\varepsilon)t}}{i(\omega + i\varepsilon - \omega_{0n})} \end{cases} \tag{2.103}$$

Putting these results in Eq. 2.102, it is obtained:

$$\delta\langle B(t)\rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega + i\varepsilon)t} \sum_{n \neq 0} \left[\frac{B_{0n}A_{n0}}{\omega + i\varepsilon - \omega_{0n}} - \frac{B_{n0}A_{0n}}{\omega + i\varepsilon + \omega_{0n}} \right] \mathrm{d}\omega \tag{2.104}$$

where the sum is the Fourier transform of the time-correlation function, R, and is defined as the dynamic polarizability:

$$\prod(BA|\omega) = \sum_{n \neq 0} \left[\frac{B_{0n}A_{n0}}{\omega + i\varepsilon - \omega_{0n}} - \frac{B_{n0}A_{0n}}{\omega + i\varepsilon + \omega_{0n}} \right] \tag{2.105}$$

Finally, the expression of the linear response in the frequency-domain can be written in the following compact form:

$$\delta\langle B(t)\rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(\omega)e^{-i(\omega+i\varepsilon)t}\prod(BA|\omega)\mathrm{d}\omega \tag{2.106}$$

Therefore, the linear response function is the variation of the expected value $\langle B\rangle$ at time t induced by the perturbation A at frequency ω. If the spectral dependency, $f(\omega)$, relative to the perturbation A is known, it is possible to predict $\langle B\rangle$ at any time.

It is worth noticing that the dynamic polarizability in Eq. 2.105 contains two poles in $\omega = -i\varepsilon - \omega_{0n}$, and $\omega = \omega_{0n} - i\varepsilon$. Moreover, it is interesting to look into the resonant case, that is, $\omega = \omega_{0n}$:

$$\lim_{\varepsilon\to0^+}\prod(BA|\omega_{0n}) \cong \frac{\langle 0|B|n\rangle\langle n|A|0\rangle}{i\varepsilon} \tag{2.107}$$

where the dynamic polarizability becomes an imaginary number.

Considering the electric dipoles induced by the oscillating electric field of an electromagnetic radiation, one obtains:

$$Im\big[\lim_{\varepsilon\to0^+}(\mu\mu|\omega_{0n})\big] \cong -\frac{|\langle n|\mu|0\rangle|^2}{\varepsilon} \tag{2.108}$$

Therefore, the polarizability, α, is an imaginary object proportional to the transition moment, and then to the oscillator strength. The problem can be solved in two different ways. The first approach involves the resolution of the response matrix by diagonalization, which gives a discrete spectrum, while in the second approach the spectrum is calculated point-by-point from the imaginary part of α, setting finite ε values. Such strategies will be discussed in detail in the following sections 2.4.2 and 2.4.3.

Let us conclude this section with a practical example, considering a monochromatic perturbation of frequency ω_0 induced by an electromagnetic radiation:

$$V(\mathbf{r},t) = A(\mathbf{r})F(t) = A(\mathbf{r})\cos(\omega_0 t) \tag{2.109}$$

Bearing in mind that the Fourier transform of $\cos(\omega_0 t)$ is: $\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$, the variation of the expected value of the observable B becomes:

$$\delta\langle B(t)\rangle = \frac{1}{2}\left[e^{i(\omega_0-i\varepsilon)t}\prod(BA|-\omega_0) + e^{-i(\omega_0+i\varepsilon)t}\prod(BA|\omega_0)\right] \tag{2.110}$$

2.4.2   Random Phase Approximation and Casida's scheme

The first order perturbation introduced above, $H^{(1)}(\mathbf{r},t)$, can be written within the first order TD perturbation theory as:

$$H^{(1)}(\mathbf{r},t) = A_\omega e^{-i\omega t} + A_{-\omega}e^{i\omega t} \tag{2.111}$$

Consequently, the time-dependent Schrödinger equation becomes:

$$\left[H^{(0)} + \lambda \left(A_\omega e^{-i\omega t} + A_{-\omega} e^{i\omega t} \right) \right] \left[\Psi_0^{(0)} e^{-iE_0 t} + \lambda \left(\Psi_0^{(1,-)} e^{-i(E_0+\omega)t} + \Psi_0^{(1,+)} e^{-i(E_0-\omega)t} \right) \right]$$

$$= i\frac{\partial}{\partial t} \left[\Psi_0^{(0)} e^{-iE_0 t} + \lambda \left(\Psi_0^{(1,-)} e^{-i(E_0+\omega)t} + \Psi_0^{(1,+)} e^{-i(E_0-\omega)t} \right) \right]$$

$$(2.112)$$

assuming that the perturbed wave function has only two components in time, namely, $e^{-i(E_0+\omega)t}$ and $e^{-i(E_0-\omega)t}$ associated with $\Psi_0^{(1,-)}$ and $\Psi_0^{(1,+)}$, respectively. The zeroth-order terms in Eq. 2.112 give the time-independent Schrödinger equation, while the first-order terms can be collected together according to their time behavior:

$$e^{-i(E_0+\omega)t} \left[H^{(0)} \Psi^{(1,-)} + A_\omega \Psi_0^{(0)} - (E_0 + \omega) \Psi_0^{(1,-)} \right] +$$

$$+ e^{-i(E_0-\omega)t} \left[H^{(0)} \Psi^{(1,+)} + A_{-\omega} \Psi_0^{(0)} - (E_0 - \omega) \Psi_0^{(1,+)} \right] = 0$$

$$(2.113)$$

Such expression is a linear combination of two TD linear independent functions, hence their coefficients (*i.e.*, quantities in the square brackets) must be null.

Let us now consider the HF case, where the unperturbed Hamiltonian represents the Fock operator ($H^{(0)} = \hat{F}$), and the ground-state wave function represents the occupied orbitals. The perturbation of the wave function must preserve the orthonormality condition over the occupied orbitals:

$$\langle \delta\varphi_i | \varphi_j \rangle + \langle \varphi_i | \delta\varphi_j \rangle = 0 \tag{2.114}$$

Since the two terms of this equation are equal to zero, the perturbed wave function is orthogonal to all the occupied orbitals, hence it must be a linear combination of virtual ones:

$$\delta\varphi_i = \sum_b^{virt} c_{ib} | \varphi_b \rangle \tag{2.115}$$

Eq. 2.115 can be adjusted such that the perturbed wave functions satisfy the time behaviors described above:

$$\Psi_i^{(1,-)} = \sum_b^{virt} X_{ib} \varphi_b \tag{2.116}$$

$$\Psi_i^{(1,+)} = \sum_b^{virt} Y_{ib}^* \varphi_b \tag{2.117}$$

By taking this formulation into Eq. 2.113 and doing the scalar product with another virtual orbital, φ_a, it is obtained:

$$\sum_b^{virt} \langle \varphi_a | F | \varphi_b \rangle X_{ib} + \langle \varphi_a | A_\omega | \varphi_i \rangle - (\varepsilon_i + \omega) \sum_b^{virt} X_{ib} \langle \varphi_a | \varphi_b \rangle = 0 \qquad (2.118)$$

$$\sum_b^{virt} \langle \varphi_a | F | \varphi_b \rangle Y_{ib}^* + \langle \varphi_a | A_{-\omega} | \varphi_i \rangle - (\varepsilon_i - \omega) \sum_b^{virt} Y_{ib}^* \langle \varphi_a | \varphi_b \rangle = 0 \qquad (2.119)$$

In both the equations, the only term of the sum that survives is that where $a = b$ because of the orbitals orthonormality. Moreover, transforming the second equation in its complex conjugate, the following expressions result:

$$\langle \varphi_a | A_\omega | \varphi_i \rangle + (\varepsilon_a - \varepsilon_i - \omega) X_{ia} = 0 \qquad (2.120)$$

$$\langle \varphi_i | A_\omega | \varphi_a \rangle + (\varepsilon_a - \varepsilon_i + \omega) Y_{ib} = 0 \qquad (2.121)$$

However, the TD external field affects the orbitals, hence modifying the Fock operator. In order to evaluate the $\langle \varphi_a | A_\omega | \varphi_i \rangle$ term, the first-order perturbed $\hat{F}$ operator ($F^{(1)}$) must be introduced:

$$F^{(1)} = \sum_j^{occ} \left(\langle \delta \varphi_j || \varphi_j \rangle + \langle \varphi_j || \delta \varphi_j \rangle \right) \qquad (2.122)$$

Then, the matrix elements become:

$$\langle \varphi_a | F^{(1)} | \varphi_i \rangle = \sum_j^{occ} \left(\langle \varphi_a \delta \varphi_j || \varphi_i \varphi_j \rangle + \langle \varphi_a \varphi_j || \varphi_i \delta \varphi_j \rangle \right)$$

$$= \sum_j^{occ} \sum_b^{virt} e^{-i\omega t} \left(Y_{jb} \langle \varphi_a \varphi_b || \varphi_i \varphi_j \rangle + X_{jb} \langle \varphi_a \varphi_j || \varphi_i \varphi_b \rangle \right) \qquad (2.123)$$

$$+ \sum_j^{occ} \sum_b^{virt} e^{i\omega t} \left(X_{jb}^* \langle \varphi_a \varphi_b || \varphi_i \varphi_j \rangle + Y_{jb}^* \langle \varphi_a \varphi_j || \varphi_i \varphi_b \rangle \right)$$

From this expression, it is derived that:

$$\langle \varphi_a | A_\omega | \varphi_i \rangle = \langle \varphi_a | z | \varphi_i \rangle + \sum_j^{occ} \sum_b^{virt} \left(X_{jb} \langle \varphi_a \varphi_j || \varphi_i \varphi_b \rangle + Y_{jb} \langle \varphi_a \varphi_b || \varphi_i \varphi_j \rangle \right) \qquad (2.124)$$

z in Eq. 2.124 is the electric dipole operator, while the X and Y coefficients are unknown. However, we know all the other terms of Eq. 2.120 and Eq. 2.121, which can be reformulated

as:

$$\sum_j^{occ} \sum_b^{virt} \left(X_{jb} \langle \varphi_a \varphi_j || \varphi_i \varphi_b \rangle + Y_{jb} \langle \varphi_a \varphi_b || \varphi_i \varphi_j \rangle \right) + (\varepsilon_a - \varepsilon_i - \omega) X_{ia} = -\langle \varphi_a | z | \varphi_i \rangle \quad (2.125)$$

$$\sum_j^{occ} \sum_b^{virt} \left(Y_{jb} \langle \varphi_i \varphi_b || \varphi_a \varphi_j \rangle + X_{jb} \langle \varphi_i \varphi_j || \varphi_a \varphi_b \rangle \right) + (\varepsilon_a - \varepsilon_i + \omega) Y_{ia} = -\langle \varphi_i | z | \varphi_a \rangle \quad (2.126)$$

These coupled equations can be also represented in the following matrix form:

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} - \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} -V \\ -V^* \end{pmatrix} \quad (2.127)$$

where $A_{ia,jb} = \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i) + \langle \varphi_a \varphi_j || \varphi_i \varphi_b \rangle$ and $B_{ia,jb} = \langle \varphi_a \varphi_b || \varphi_i \varphi_j \rangle$. Therefore, by solving this matrix problem, the coefficients X and Y for a given frequency ω can be extracted. Note that when the frequency corresponds to the excitation energy, we formally have values that go to infinity.

It is worth mentioning that this matrix object is connected with polarizability. In order to obtain the response of the system, let us start with the variation of the electric dipole as our observable:

$$\delta \langle B(t) \rangle = \sum_i^{occ} \left(\langle \delta \varphi_i | B | \varphi_i \rangle + \langle \varphi_i | B | \delta \varphi_i \rangle \right)$$
$$= \sum_i^{occ} \left(\langle \delta \varphi_i | z | \varphi_i \rangle + \langle \varphi_i | z | \delta \varphi_i \rangle \right) \quad (2.128)$$

As shown before in Eq. 2.115, the first-order perturbation of orbitals can be expressed as a linear combination of virtual orbitals. Therefore, previous Eq. 2.128 becomes:

$$\delta \langle B(t) \rangle = \sum_i^{occ} \sum_a^{virt} e^{-i\omega t} \left(Y_{ia} \langle \varphi_a | z | \varphi_i \rangle + X_{ia} \langle \varphi_i | z | \varphi_a \rangle \right)$$
$$+ \sum_i^{occ} \sum_a^{virt} e^{i\omega t} \left(X_{ia}^* \langle \varphi_a | z | \varphi_i \rangle + Y_{ia}^* \langle \varphi_i | z | \varphi_a \rangle \right) \quad (2.129)$$

By comparing this expression with Eq. 2.110, it is possible to identify:

$$\prod (BA_\omega | \omega) = \sum_i^{occ} \sum_j^{virt} \left(Y_{ia} \langle \varphi_a | z | \varphi_i \rangle + X_{ia} \langle \varphi_i | z | \varphi_a \rangle \right) \quad (2.130)$$

Therefore, δB goes to infinite (absorption) when the coefficients X and Y are the solutions of the following pseudo-eigenvalue equations:

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{2.131}$$

This formulation is the well-known Random Phase Approximation (RPA). The RPA problem can be simplified in two cases: (i) the molecular orbitals are real, hence $A = A^*$ and $B = B^*$, (ii) the B matrix can be neglected, leading to the more common eigenvalue problem:

$$AX = \omega X \tag{2.132}$$

The latter case is known as Tamm-Dancoff Approximation (TDA). [174, 175]

Replacing the Fock operator with the KS operator as $H^{(0)}$ leads instead to the TDDFT equations. The TDDFT sub-matrices, A and B, in Eq. 2.127 differ from those in the RPA formulation for the presence of the bielectronic term and the xc potential. Therefore, depending on the theory, A assumes the following form:

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle \varphi_a\varphi_j||\varphi_i\varphi_b \rangle$$

$$= \begin{cases} RPA : \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle \varphi_a\varphi_j|\varphi_i\varphi_b \rangle - \langle \varphi_a\varphi_j|\varphi_b\varphi_i \rangle \\[2mm] TDDFT : \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle \varphi_a\varphi_j|\varphi_i\varphi_b \rangle - \langle \varphi_a\varphi_j|K_{xc}|\varphi_i\varphi_b \rangle \end{cases} \tag{2.133}$$

while B becomes:

$$B_{ia,jb} = \langle \varphi_a\varphi_b||\varphi_i\varphi_j \rangle$$

$$= \begin{cases} RPA : \langle \varphi_a\varphi_b|\varphi_i\varphi_j \rangle - \langle \varphi_a\varphi_b|\varphi_j\varphi_i \rangle \\[2mm] TDDFT : \langle \varphi_a\varphi_b|\varphi_i\varphi_j \rangle - \langle \varphi_a\varphi_b|K_{xc}|\varphi_i\varphi_j \rangle \end{cases} \tag{2.134}$$

If A and B are both real matrices, the RPA problem is simplified to the following system of equations:

$$\begin{cases} AX + BY - \omega X = -V \\[2mm] BX + AY + \omega Y = -V \end{cases} \tag{2.135}$$

Now, subtracting the second equation from the first one, it is obtained:

$$(A - B)(X - Y) = \omega(X + Y) \rightarrow (X - Y) = \omega(A - B)^{-1}(X + Y) \tag{2.136}$$

Such formulation is quite interesting in TDDFT, where the $(A - B)$ matrix becomes diagonal, while this is not true for RPA. From Eq. 2.136, it follows that:

$$\left[(A + B) - \omega^2 (A - B)^{-1} \right] (X + Y) = -2V \tag{2.137}$$

The diagonal $(A - B)^{-1}$ matrix is usually defined as ε^{-1}.

This final expression leads to the famous Casida's formulation [71] of TDDFT:

$$\Omega F = \omega^2 F \tag{2.138}$$

where $\Omega = \varepsilon^{-1/2}(A + B)\varepsilon^{1/2}$, and $F = \varepsilon^{-1/2}(X + Y)$. This eigenvalue problem can be transformed as follows:

$$\left(\Omega - \omega^2 \right)^{-1} = F_{n0} \left(\frac{1}{\omega_{no}^2 - \omega^2} \right) F_{n0}^\dagger \tag{2.139}$$

In this scheme, the variation of our observable B becomes:

$$\delta \langle B(t) \rangle = V^\dagger (X + Y) = V^\dagger \varepsilon^{1/2} F = V^\dagger \varepsilon^{1/2} (-2) F_{n0} \left(\frac{1}{\omega_{no}^2 - \omega^2} \right) F_{n0}^\dagger \varepsilon^{1/2} V \tag{2.140}$$

When the frequency of the external potential is in resonance with one of the excited state energies, Eq. 2.140 can be simplified as:

$$\delta \langle B(t) \rangle \cong -\frac{2 \left| V^\dagger \varepsilon^{1/2} F_{n0} \right|^2}{\omega_{n0}^2 - \omega^2} \cong -\frac{\left| V^\dagger \varepsilon^{1/2} F_{n0} \right|^2}{(\omega_{n0} - \omega)\omega} \tag{2.141}$$

By replacing the numerator of Eq. 2.141 with $\omega |\langle n|z|0 \rangle|^2$, it results:

$$\delta \langle B(t) \rangle \cong \frac{|\langle n|z|0 \rangle|^2}{\omega - \omega_{n0}} \tag{2.142}$$

Therefore, the Casida's equations are solved by diagonalizing the matrix Ω, which is a four-indexed matrix defined within the 1h-1p space. The dimension of Ω corresponds to the product between the number of occupied and virtual KS orbitals, while its elements are:

$$\Omega_{ia\sigma,bj\tau} = \delta_{\sigma\tau}\delta_{ij}\delta_{ab} \left(\varepsilon_a - \varepsilon_i \right)^2 + 2\sqrt{\varepsilon_a - \varepsilon_i} K_{ia\sigma,bj\tau} \sqrt{\varepsilon_b - \varepsilon_j} \tag{2.143}$$

where i and j are the indeces that run over the occupied orbitals, a and b are those running over the virtual orbitals, and σ and τ are the indeces for the spin. Instead, $K_{ia\sigma,bj\tau}$ is the so-called coupling matrix which includes the Coulomb term, the xc term, and the xc kernel that can be approximated by using the Adiabatic LDA (ALDA) [176]:

$$K_{ia\sigma,bj\tau} = \int \int \varphi_{i\sigma}(\mathbf{r})\varphi_{a\sigma}(\mathbf{r}) \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}^{ALDA}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') \right] \varphi_{j\tau}(\mathbf{r}')\varphi_{b\tau}(\mathbf{r}')\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r} \tag{2.144}$$

Moreover, considering the Casida's eigenvectors, F_{n0}, that correspond to the ω_{n0} eigenvalues, it is possible to extract the intensities of the spectral lines (*i.e.*, oscillator strength or rotatory strength values). For instance, the expression for the oscillator strength (absorption spectrum) within the Casida's scheme becomes:

$$
\begin{aligned}
f_{n0} &= \frac{2}{3}\omega \left(|\langle n|x|0\rangle|^2 + |\langle n|y|0\rangle|^2 + |\langle n|z|0\rangle|^2 \right) \\
&= \frac{2}{3} \left(x^\dagger \varepsilon^{1/2} F_{n0} + y^\dagger \varepsilon^{1/2} F_{n0} + z^\dagger \varepsilon^{1/2} F_{n0} \right)
\end{aligned}
\tag{2.145}
$$

Therefore, excitation energies and intensities can be obtained by solving the Casida's equations. However, the direct solution of Eq. 2.138 is practically infeasible due to computational and storage requirements. The eigenvalue problem is usually solved by means of the Davidson's algorithm, [72] which is a very efficient iterative procedure for extracting the lowest excitation energies and intensities of the spectrum. This method works well for small-/medium-sized systems, while it becomes unsuitable for larger systems or when a significant number of eigenvalues is needed. Indeed, in these latter cases the procedure may become numerically unstable. Among the alternatives to solve TDDFT equations for large systems, such as metallic nanoclusters, it can be mentioned the polTDDFT algorithm [93] that will be discussed in the next section 2.4.3.

2.4.3   The polTDDFT algorithm

As introduced above, the Casida's formulation and its practical implementation (Davidson's algorithm) may suffer from some limitations. First of all, despite its robustness and efficiency, the Casida's scheme requires the diagonalization of very high-dimensional matrices ($\Omega = O(N_{occ} \times N_{virt})$). Furthermore, employing the Davidson's algorithm, we are able to extract only "few" low energy eigenvalues. This problem becomes particularly relevant when treating metallic clusters which show a very high-density of states. An efficient alternative here can be the polTDDFT algorithm proposed by Baseggio and co-workers. [93] This method, which still works in the linear response regime, employs the polarizability tensor to drastically reduce the size of the matrices that need to be diagonalized. As a result, polTDDFT allows faster calculations and, above all, to obtain spectra of large systems up to the required excitation energy. More in detail, polTDDFT extracts absorption and ECD spectra, point-by-point, from the imaginary part of the complex dynamic polarizability ($\alpha(\omega)$) and rotatory strength tensor ($\beta(\omega)$), respectively. Let us start by discussing the formalism for absorption spectra, then continuing with the formalism for ECD spectra, that are the main focus of this PhD project.

In polTDDFT, the absorption spectrum, $\sigma(\omega)$, is calculated point-by-point from the imaginary part of the complex dynamic polarizability as:

$$
\sigma(\omega) = \frac{4\pi\omega}{c} Im[\alpha(\omega)]
\tag{2.146}
$$

where c is the speed of light, and ω is the complex photon energy ($\omega = \omega_r + i\omega_i$). Therefore, the real part, ω_r, corresponds to the photon frequency, while the imaginary part, ω_i, corresponds to the line enlargement that naturally arises from the uncertainty between lifetime and energy of the excited states. Instead, the dynamic polarizability is a tensor, whose diagonal elements assume the following expression:

$$\alpha_{kk}(\omega) = \int \rho_k^{(1)}(\omega, \mathbf{r}) k \, d\mathbf{r} \tag{2.147}$$

where k can be one of the three x, y, z components. From now on, the z component will be taken as an example, but similar considerations can be applied to the other two directions. $\rho_k^{(1)}(\omega, \mathbf{r})$ in Eq. 2.147 is the TD first-order correction of the electronic density induced by the external TD perturbation (*i.e.*, incident radiation).

For the absorption spectrum (see Eq. 2.146), the average over the trace of the polarizability tensor is considered:

$$\alpha(\omega) = \frac{1}{3} \sum_{k=1}^{3} \alpha_{kk}(\omega) \tag{2.148}$$

Note that the trace of the tensor is invariant with respect to the selected system of coordinates, hence it does not require the diagonalization or choice of the principal axes.

Now, the only unknown term is $\rho_z^{(1)}(\omega, \mathbf{r})$ that can be actually calculated as follows:

$$\rho_z^{(1)}(\omega, \mathbf{r}) = \int \chi_{KS}(\omega, \mathbf{r}, \mathbf{r}') V_{SCF}^z(\omega, \mathbf{r}') \, d\mathbf{r}' \tag{2.149}$$

where $\chi_{KS}(\omega, \mathbf{r}, \mathbf{r}')$ is the dielectric susceptibility of the KS system of non-interacting electrons. Instead, $V_{SCF}^z(\omega, \mathbf{r}')$ is the effective perturbative potential accounting for the external potential, as well as for the Coulomb and xc terms:

$$V_{SCF}^z(\omega, \mathbf{r}') = V_{ext}^z(\omega, \mathbf{r}') + \int \frac{\rho_z^{(1)}(\omega, \mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \left.\frac{\partial V_{xc}}{\partial \rho}\right|_{\rho_0} \rho_z^{(1)}(\omega, \mathbf{r}) \tag{2.150}$$

$V_{SCF}^z(\omega, \mathbf{r}')$ includes the xc kernel functional that will be treated again with the ALDA approach, [176] as shown above.

Eq. 2.149 and Eq. 2.150 can be reworded in their operator form as:

$$\begin{cases} \rho_z^{(1)} = \chi_{KS} V_{SCF}^z \\ V_{SCF}^z = V_{ext}^z + K\rho_z^{(1)} \end{cases} \tag{2.151}$$

where K is the coupling matrix (see Eq. 2.144) that contains the Coulomb and xc terms:

$$K(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \delta(\mathbf{r} - \mathbf{r}') \left.\frac{\partial V_{XC}}{\partial \rho}\right|_{\rho_0} \tag{2.152}$$

The linearity of the operators allows us to do some algebra on this system of equations, thus obtaining:

$$[1 - \chi_{KS}K]\,\rho_z^{(1)} = \chi_{KS}V_{ext}^z \tag{2.153}$$

In order to solve this matrix problem, it is convenient to expand the induced density as a linear combination of frequency-independent fitting functions and frequency-dependent expansion coefficients (see section 2.3.1):

$$\rho_z^{(1)}(\omega, \mathbf{r}) = \sum_{\mu} f_\mu(\mathbf{r})b_\mu(\omega) \tag{2.154}$$

By substituting Eq. 2.154 in Eq. 2.153 and doing the scalar product with $\langle f_\nu|$, the matrix problem becomes:

$$\sum_{\mu} \left(\langle f_\nu|f_\mu\rangle - \langle f_\nu|K\chi_{KS}|f_\mu\rangle\right) b_\mu = \langle f_\nu|\chi_{KS}V_{ext}^z\rangle \tag{2.155}$$

and in terms of matrix elements:

$$\sum_{\mu} \left(S_{\nu\mu} - M_{\nu\mu}(\omega)\right) b_\mu(\omega) = d_\nu(\omega) \tag{2.156}$$

Finally, the non-homogeneous system can be formulated in the following compact matrix form:

$$[S - M(\omega)]\,b = d \tag{2.157}$$

where S is the overlap matrix between fitting functions, b is the unknown vectors of frequency-dependent expansion coefficients, and d is the frequency-dependent vector. The latter vector is known and can be expressed as follows:

$$d_\nu = \langle f_\nu|\chi_{KS}|z\rangle \tag{2.158}$$

with z being the $z - th$ component of the external oscillating dipole, V_{ext}^z. Instead, M is a frequency-dependent matrix that needs to be calculated for each wavelength. Unfortunately, this results in an impossible computational problem. Therefore, some approximations are needed in order to proceed.

The strength of the polTDDFT algorithm lies in the construction of a simplified expression of M, which is obtained by a linear combination of frequency-independent matrices, G^k, and frequency-dependent expansion coefficients, $s_k(\omega)$:

$$M(\omega) = \sum_{k} s_k(\omega)G^k \tag{2.159}$$

The matrices G^k are calculated only once since frequency-independent, while $M(\omega)$ is built by varying the coefficients $s_k(\omega)$ at each frequency (= photon energy). This way, the calculation

becomes fast and affordable.

The approximation made to treat M can be justified starting from the expression of the dielectric susceptibility:

$$\chi_{KS}(\omega, \mathbf{r}, \mathbf{r}') = \sum_i^{occ} \sum_a^{virt} \varphi_i(\mathbf{r}) \varphi_a(\mathbf{r}) \frac{4\varepsilon_{ia}}{\omega^2 - \varepsilon_{ia}^2} \varphi_i(\mathbf{r}') \varphi_a(\mathbf{r}')$$

$$= \sum_i^{occ} \sum_a^{virt} \Theta_{ia}(\mathbf{r}) \lambda_{ia}(\omega) \Theta_{ia}(\mathbf{r}')$$

(2.160)

where ε_{ia} is the difference between the energy of occupied and virtual orbitals. This object is approximated as constant because of the high density of excitation energies. It is worth noticing that in Eq. 2.160, $\lambda_{ia}(\omega)$ is the only frequency-dependent term.

Let us now discretize the scale of excitation energies into P ordered intervals, I_k, as schematized in the following figure.
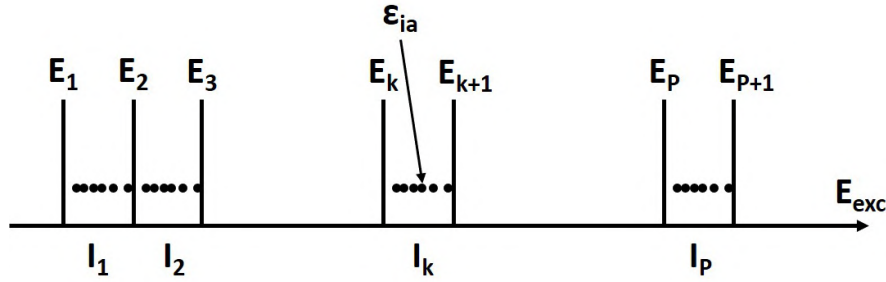
FIGURE 2.5: Graphical representation of the grid of energy.

Note that the minimum ε_{ia} corresponds to the HOMO-LUMO gap. If the I_k intervals are narrow enough, thus the density of excitation energies is high, ε_{ia} can be considered constant in each interval. Consequently, it can be approximated by using the mean of adjacent orbital energies as follows:

$$\overline{E}_i = \frac{E_i + E_{i+1}}{2}$$

(2.161)

and the dielectric susceptibility can be expressed as:

$$\chi_{KS}(\omega, \mathbf{r}, \mathbf{r}') = \sum_{k=1}^P \frac{4\overline{E}_k}{\omega^2 - \overline{E}_k^2} \sum_{\varepsilon_{ia} \in I_k} \Theta_{ia}(\mathbf{r}) \Theta_{ia}(\mathbf{r}')$$

$$= \sum_{k=1}^P s_k(\omega) \tilde{G}^k(\mathbf{r}, \mathbf{r}')$$

(2.162)

where the frequency-dependent term is outside the inner sum. This way, χ_{KS} is defined as a linear combination of frequency-independent objects and frequency-dependent coefficients. This expression of χ_{KS} allows us to construct the matrix $M(\omega)$ (see Eq. 2.159) in the following

way:

$$
\begin{aligned}
M_{\nu\mu}(\omega) &= \sum_{k=1}^{max} \langle f_\nu | s_k(\omega) \tilde{G}^k(\mathbf{r}, \mathbf{r}') K | f_\mu \rangle \\
&= \sum_{k=1}^{max} s_k(\omega) \sum_{\varepsilon_{ia} \in I_k} \langle f_\nu | \Theta_{ia}(\mathbf{r}) \rangle \langle \Theta_{ia}(\mathbf{r}') | K | f_\mu \rangle \\
&= \sum_{k=1}^{max} s_k(\omega) G^k_{\nu\mu}
\end{aligned}
\tag{2.163}
$$

Therefore, the coefficients $s_k(\omega)$ are equal to:

$$
s_k(\omega) = \frac{4\overline{E}_k}{\omega^2 - \overline{E}_k^2}
\tag{2.164}
$$

The $G^k_{\nu\mu}$ matrix can be conveniently split into two sub-matrices:

$$
G^k_{\nu\mu} = \sum_{\varepsilon_{ia} \in I_k} A^k_{\nu,\varepsilon_{ia}} B^k_{\varepsilon_{ia},\mu}
\tag{2.165}
$$

corresponding to:

$$
A^k_{\nu,\varepsilon_{ia}} = \langle f_\nu | \Theta_{ia}(\mathbf{r}) \rangle
\tag{2.166}
$$

$$
B^k_{\varepsilon_{ia},\mu} = \langle \Theta_{ia}(\mathbf{r}') | \frac{1}{|\mathbf{r} - \mathbf{r}'|} | f_\mu \rangle + \langle \Theta_{ia}(\mathbf{r}') | \delta(\mathbf{r} - \mathbf{r}') \frac{\partial V_{xc}}{\partial \rho} \Big|_{\rho_0} | f_\mu \rangle
\tag{2.167}
$$

A similar expression can be derived for the vector d_ν of the non-homogeneous system (see Eq. 2.158).

Now, all the elements needed to determine the unknown vector $b(\omega)$ in Eq. 2.157 are available, therefore it is also possible to obtain the induced density (Eq. 2.154). Once $\rho_z^{(1)}(\omega, \mathbf{r})$ is determined, we are finally able to calculate the complex dynamic polarizability (see Eq. 2.147) as:

$$
\alpha_{zz}(\omega) = \sum_\mu b_\mu(\omega) \int f_\mu(\mathbf{r}) z \, d\mathbf{r}
\tag{2.168}
$$

and to extract the absorption spectrum from its imaginary part (Eq. 2.146).

The polTDDFT approach can be easily extended to the calculation of ECD spectra. It should be noted that in the following discussion the dependence on the current density of the xc functionals will be neglected. As already mentioned, the CD is defined as the differential absorption of the left- and right-handed circularly polarized light by a cromophore:

$$
CD = A_L - A_R
\tag{2.169}
$$

Considering an electromagnetic radiation that propagates along the x direction, it is possible to express CD as:

$$CD = 2\gamma Im\left[\langle 0|\mu_y|n\rangle\langle n|m_y|0\rangle + \langle 0|\mu_z|n\rangle\langle n|m_z|0\rangle\right] \tag{2.170}$$

where μ_i and m_i are the $i-th$ components of the electric and magnetic dipole, respectively, while γ is a constant. Actually, this expression must be averaged over all the possible orientations since molecules are free to rotate in solution:

$$CD = \frac{4}{3}\gamma Im\left[\langle 0|\mu|n\rangle\langle n|m|0\rangle\right] \tag{2.171}$$

Eq. 2.171 is the well-known Rosenfeld equation [177] that can be written in terms of rotatory strength, R_{0n}, as:

$$R_{0n} = Im\left[\langle 0|\mu|n\rangle\langle n|m|0\rangle\right] \tag{2.172}$$

The R_{0n} values correspond to the intensity values of the ECD spectra, thus the objects we want to calculate. Moreover, the rotatory strength is correlated to the rotation tensor, β:

$$\beta = \frac{1}{3}\sum_{k=1}^{3}\beta_{kk} = \frac{2}{3}c\sum_{n\neq 0}\frac{R_{0n}}{\omega_{n0}^2 - \omega^2} \tag{2.173}$$

As already done for the polarizability α, only the average over the trace of the tensor is taken into account. Therefore, R_{0n} is obtained by calculating the tensor β, which in turn is derived from the dipole induced by the external electromagnetic radiation:

$$\mu_i^{(1)} = \sum_j \alpha_{ij}C_j - \sum_j \frac{\beta_{ij}}{c}\frac{\partial B_j}{\partial t} \tag{2.174}$$

where C and B are the electric and magnetic field, respectively. From Eq. 2.174, it is clear that β comes from the electric dipole induced by the external oscillating magnetic field.

Similarly to what has been done above for the polarizability, only the z component will be considered to define the first-order perturbed Hamiltonian:

$$H^{(1)}(t) = -\mathbf{m}\cdot\mathbf{B} = -m_z B_z\cos(\omega t) = -\frac{1}{2}\left(m_z e^{i\omega t} + m_z^{\dagger}e^{-i\omega t}\right)B_z \tag{2.175}$$

By using LRT, the magnetic polarizability can be expressed as:

$$\prod(\mu_z m_z|\omega) = \lim_{\varepsilon\to 0^+}\sum_{n\neq 0}\left(\frac{\langle 0|\mu_z|n\rangle\langle n|m_z|0\rangle}{\omega - \omega_{n0} + i\varepsilon} - \frac{\langle 0|m_z|n\rangle\langle n|\mu_z|0\rangle}{\omega + \omega_{n0} + i\varepsilon}\right) \tag{2.176}$$

where ε is the adiabatic switch factor introduced in section 2.4.1. Since m_z is purely imaginary, it is obtained that:

$$m_z^{\dagger} = -m_z \tag{2.177}$$

and:

$$\langle 0|m_z|n\rangle = -\langle n|m_z|0\rangle \tag{2.178}$$

This relation allows us to recast Eq. 2.176 in the following way:

$$\prod(\mu_z m_z|\omega) = \lim_{\varepsilon\to 0^+} \sum_{n\neq 0} \langle 0|\mu_z|n\rangle\langle n|m_z|0\rangle \left(\frac{1}{\omega - \omega_{n0} + i\varepsilon} + \frac{1}{\omega + \omega_{n0} + i\varepsilon}\right)$$
$$= -\int \rho_z^{(1)}(\omega,\mathbf{r})z\mathrm{d}\mathbf{r} \tag{2.179}$$

However, this result is just formal due to the impossibility of calculating the sum over the excited states. Such limitation can be overcome following the same strategy adopted above for the absorption case. Let us start expressing the induced density as a perturbation of the magnetic field:

$$\rho_z^{(1)}(\omega,\mathbf{r}) = \sum_i^{occ}\sum_a^{virt} \varphi_i\varphi_a\langle a|m_z|i\rangle \left(\frac{1}{\omega - \omega_{ai} + i\varepsilon} + \frac{1}{\omega + \omega_{ai} + i\varepsilon}\right)$$
$$+ \int \chi_{KS}(\omega,\mathbf{r},\mathbf{r}')V_{z,ind}(\omega,\mathbf{r}')\mathrm{d}\mathbf{r}' \tag{2.180}$$

$V_{z,ind}(\omega,\mathbf{r}')$ is the potential induced by $\rho_z^{(1)}(\omega,\mathbf{r})$ and assumes the following expression:

$$V_{z,ind}(\omega,\mathbf{r}') = \int \frac{\rho_t^{(1)}(\omega,\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}' + \left.\frac{\partial V_{xc}}{\partial\rho}\right|_{\rho_0}\rho_t^{(1)}(\omega,\mathbf{r}) \tag{2.181}$$

As in the previous Eq. 2.154, the induced density can be expanded as a linear combination of fitting functions:

$$\rho_z^{(1)}(\omega,\mathbf{r}) = \sum_\mu f_\mu(\mathbf{r})q_\mu(\omega) \tag{2.182}$$

subsequently obtaining the following matrix problem by simple substitution:

$$[S - M(\omega)]q = g \tag{2.183}$$

The objects in this expression are analogous to those in Eq. 2.157, except for the vector g which is defined as:

$$g_\mu = \sum_{k=1}^{max} t_k(\omega) \sum_{\varepsilon_{ia}\in I_k} A_{\mu,ia}^k(\mathbf{r},\mathbf{r}')\langle i|m_z|a\rangle \tag{2.184}$$

where the frequency-dependent coefficients, $t_k(\omega)$, assume the following expression:

$$t_k(\omega) = \frac{1}{\omega - \omega_{ai} + i\varepsilon} + \frac{1}{\omega + \omega_{ia} + i\varepsilon} \tag{2.185}$$

Therefore, once the objects needed to extract the absorption spectrum are calculated, it becomes very easy to extract the ECD spectrum too, since the only unknown object is the vector g_μ.

Now, in order to have a practical expression for calculating ECD spectra, it is necessary to make some adjustments in Eq. 2.174 and Eq. 2.179. Let us consider the z-component of a monochromatic radiation with frequency ω. The behavior of the magnetic field over time is described by the following equation:

$$B_z(t) = B_z \cos(\omega t) = B_z \left(\frac{e^{-i\omega t} + e^{i\omega t}}{2} \right) \tag{2.186}$$

while its derivative becomes:

$$\frac{\partial B_z(t)}{\partial t} = B_z \left(\frac{-i\omega e^{-i\omega t} + i\omega e^{i\omega t}}{2} \right) \tag{2.187}$$

Substituting Eq. 2.187 into:

$$-\frac{\beta_{zz}}{c} \frac{\partial B_z}{\partial t} = \frac{1}{2} \left(\prod(\mu_i m_i | \omega) e^{-i\omega t} + \prod(\mu_i m_i | -\omega) e^{i\omega t} \right) \tag{2.188}$$

the following expression for the diagonal component of the rotation tensor is obtained:

$$\beta_{zz} = -\frac{ic}{\omega} \prod(\mu_i m_i | \omega) \tag{2.189}$$

In the resonant case ($\omega \rightarrow \omega_{0n}$), β displays poles at excitation energies and its strength is extracted from the residue. Therefore, the previous equation can be reworded as:

$$\beta_{zz} = -\frac{ic}{\omega} \frac{\langle 0|\mu_z|n\rangle \langle n|m_z|0\rangle}{i\varepsilon} \tag{2.190}$$

Considering the average over all the orientations and closed-shell systems (only singlet-singlet transitions), β becomes:

$$\beta = -\frac{2icR_{0n}}{3\omega\varepsilon} \tag{2.191}$$

Therefore, a practical expression for the rotatory strength is finally reached:

$$R_{0n} = \frac{3\omega\varepsilon}{2c} Im[\beta(\omega)] \tag{2.192}$$

2.5 Relativistic effects

The inclusion of relativistic effects becomes necessary when describing the electronic structure of systems that contain heavy atoms. Thereupon, these effects have been included in all the works where metal clusters were treated (see sections 4.1.2, 4.1.3, and 4.2.1).

The relativistic effects arise with objects that move at a speed close to that of light, c. Translating this concept to chemical systems, these effects occur when heavy elements (*e.g.*, Au) are considered since electrons of the inner orbitals, particularly 1s orbitals, show an average speed comparable to c. As a result of the fundamental relation between mass and velocity, these inner orbitals are affected by a relativistic contraction. This phenomenon is actually extended to all s-type orbitals due to their orthogonality with the 1s orbital. [178] The contraction of the core orbitals improves the nuclear shielding, which in turn makes the valence orbitals less attracted to the nucleus, thus higher in energy. The combination of these two effects brings to a reduced gap between the 5d and 6s gold bands. [179]

The quantum-mechanical description of the electronic structure considering relativistic effects was proposed in 1928 by Dirac, [180] who extended the Schrödinger equation to the relativistic case:

$$\hat{H}_D \Psi(\mathbf{r}, t) = i\hbar \frac{\partial \Psi(\mathbf{r}), t}{\partial t} \tag{2.193}$$

The Dirac Hamiltonian, $\hat{H}_D$ is composed of four terms:

$$\hat{H}_D = c\alpha \cdot p + m_0 c^2 \beta \tag{2.194}$$

where p is the well-known moment operator: $p = -i\hbar \nabla$, and α is a vector whose components are traceless 4×4 diagonal matrices with the following form:

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{2.195}$$

σ_i represents the Pauli spin matrix of the $i - th$ component:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.196}$$

Instead, β is a traceless 4×4 diagonal matrix defined as:

$$\beta = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix} \tag{2.197}$$

where its elements are:

$$1_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.198}$$

In this case, the electronic spin must be included in the wave function which becomes:

$$\Psi(\mathbf{r}, t) = \begin{pmatrix} \varphi_\alpha(\mathbf{r}, t) \\ \varphi_\beta(\mathbf{r}, t) \\ \xi_\alpha(\mathbf{r}, t) \\ \xi_\beta(\mathbf{r}, t) \end{pmatrix} \tag{2.199}$$

where α and β refer to spin up and down, respectively, while φ and ξ are the so-called large and small components.

Introducing a potential, V, in the Hamiltonian, it is obtained:

$$\hat{H} = \hat{H}_D + V = c\alpha \cdot p + m_0 c^2 \beta + V \tag{2.200}$$

Therefore:

$$(E - c\alpha \cdot p - m_0 c^2 \beta - V)\Psi(\mathbf{r}, t) = 0 \tag{2.201}$$

However, the solution of the Dirac equation is not trivial since it is not bound from below, thus negative eigenvalues are given too. This means that the variational principle is no longer granted here due to the absence of a minimum for the solutions. Moreover, this Hamiltonian requires a special set of basis functions and a huge computational effort because of its four-components nature. In this respect, some approximations are needed for $\hat{H}_D$. For instance, it is possible to approximately decouple the large and small components through the Douglass-Kroll transformation, [181] scaling down to a two-components problem. Furthermore, if the Spin-Orbit (SO) coupling is neglected, only one component will remain. This formulation to describe the electronic structure is known as Scalar Relativistic (SR).

In this PhD project, the SR Zeroth Order Regular Approximated (SR-ZORA) [182] equations were used. The ZORA approach is based on reformulating the energy expression and expanding in the term $\frac{E}{2mc^2 - V}$, which remains small even close to the nucleus. Retaining only the zeroth-order term, one gets:

$$E = \frac{p^2 c^2}{2mc^2 - V} + V \tag{2.202}$$

which gives the following ZORA Hamiltonian:

$$H_{ZORA} = V + \sigma \cdot p \frac{c^2}{2c^2 - V} \sigma \cdot p \tag{2.203}$$

All the terms have been previously introduced. Therefore, this ZORA Hamiltonian incorporates relativistic effects, that are traditionally introduced only at the level of the Pauli Hamiltonian, but it can be used variationally without suffering from the singularities for $r \to 0$.

Chapter 3

Computational Procedure

In this section, the main steps of the computational approach proposed in this work are reported. Clearly, some of the following details must be adjusted depending on the system under study. Anyway, extensive information are reported in different publications of this PhD work (see sections 4.1.1, 4.1.2, 4.1.3, and 4.1.4).

The first step of the procedure regards the production of a trajectory by means of a classical MD simulation. Such simulations can be easily performed employing, for instance, the GROMACS package [107] and both a suitable FF and solvent model. Generally, we work in the NVT ensemble by keeping the temperature (T) constant with the stochastic velocity rescaling thermostat. [113] Moreover, the LINCS algorithm [183] is used to constrain all the bonds to improve the perfomances. The Particle Mesh Ewald method [121] is adopted to treat the long-range electrostatic interactions.

Before producing an MD trajectory, the (solute+solvent) box must be properly set up and equilibrated. First of all, the system of interest is inserted into a cubic box and solvated. The volume and number of solvent molecules are chosen in order to reproduce the concentration of the experimental ECD measurement in solution. Afterwards, the energy of the box is minimized using the steepest descent algorithm. This minimization is conducted at 0 K, hence it is necessary to heat before the MD production. In practice, T is increased of 50 K at each minimization until the experimental T value is reached (usually RT). In the second part of the equilibration, the box is adjusted to achieve the correct density at the T of interest and a pressure of 1 bar. In detail, the box volume is modified until its average pressure reaches that measured on a box of pure solvent with the same number of solvent molecules. The average pressure of the solvent box, $\bar{P}_{solv}$, is obtained by running a short simulation (*ca.* 5/10 ns) in the NVT ensemble at the experimental density of the pure solvent at the selected T (*e.g.*, d_{H_2O} at $25°C$). Afterwards, a first MD simulation of the whole box (5/10 ns again) is performed, giving an average pressure $\bar{P}_{syst}$. If $\bar{P}_{syst} > \bar{P}_{solv}$, the box will be expanded, while if $\bar{P}_{syst} < \bar{P}_{solv}$, the box will be reduced. The procedure is repeated until the condition $\bar{P}_{syst} \sim \bar{P}_{solv}$ is reached.

The equilibrated box is then employed for the MD production, whose time length changes with the system under investigation. Typically, an MD production covers hundreds of ns. The trajectory thus obtained is analyzed in terms of ED (see section 2.2).

First of all, the solute is centered along the MD to avoid non-physical fluctuations due to the use of PBC. Moreover, since the interest here is focused only on the internal motion, the

overall roto-translational one is removed with a least square fitting to a reference structure (usually the solute structure). Now, the first step of the ED analysis is the construction and diagonalization of the covariance matrix of the (solute) atomic coordinates. This step produces a set of eigenvectors that correspond to the eigendirections along which the system undergoes its internal motion, while their associated eigenvalues are the actual values of the mean square fluctuations. The first eigenvectors correspond to the highest eigenvalues, thus representing the directions that are essential to account for most of the conformational transitions. In this approach, the analysis is limited to the first two eigenvectors, termed as essential eigenvectors, to compromise between accuracy and computational cost. Therefore, the MD trajectory is now projected onto these two eigendirections, producing the so-called Principal Components (PCs) which provide us with a conformational landscape easy to investigate. In practice, a 2D histogram is built with these PCs, where each square represents a conformational basin that is more or less populated depending on how many PCs pairs fall within it. The generic conformational basin i is then associated with a certain number of pairs of projections, N_i, and a probability, $P(i)$, which is defined as:

$$P(i) = \frac{N_i}{N_{max}} \tag{3.1}$$

where N_{max} is the number of PCs that fall within the most spanned (probable) square. Considering $P(ref)(= 1)$ as the probability of the most populated basin and a negligible difference between the partial molar volumes (MD simulations in the NVT ensemble), a (standard) Gibbs free energy difference can be estimated between the $i - th$ and ref basins:

$$\Delta G^{\circ} = -RT \ln \frac{P(i)}{P(ref)} \tag{3.2}$$

Repeating the procedure for each square, the relative free energy landscape is obtained. This conformational space is then investigated to extract a set of representative structures (with $\Delta G^{\circ} < k_B T$).

In order to reproduce properly the experimental ECD, it is often necessary to explicitly include the solvent. In this case, suitable conformations of the solvation shell must also be extracted. Practically, each solute conformer, selected with the previous ED analysis, is inserted again into the equilibrated box to produce a new MD trajectory. These simulations are performed under the previous conditions (same temperature, density, *etc...*), but with the solute conformer kept frozen at the center. This way, only the conformational role of the solvent is taken into account. The hereafter termed constrained-MD simulations run for shorter times, in a range of 10-25 ns. Now, the trajectory will be analyzed again by means of ED to locate plausible solute-solvent clusters. It is worth noticing that the box contains a quite large number of solvent molecules which make the ED analysis infeasible. Therefore, we reduce the complexity of the problem considering only the solvent molecules around the solute. Such clusters

are built first constructing an internal reference frame centered in the solute geometrical center. The projection of the frozen solute coordinates onto the above unit vectors provides the axes of the ellipsoid best describing the solute shape. At each frame of the MD, it is possible now to extract the (pre-)selected N solvent molecules that show the lowest square distances within the metric of the so-built ellipsoid. Finally, a trajectory of the solute-(solvent)$_N$ cluster is obtained. The ED analysis is then repeated providing the conformational landscape of the solute-(solvent)$_N$ cluster. It is worth mentioning that here the eigenvalues should be higher with respect to those obtained for the solute, following a much less steep behavior. Similar results should be obtained for the PCs. Indeed, in the case of a non-covalent system, the overall fluctuation is usually spread over a larger number of internal degrees of freedom.

For each of the K selected solute conformers, we obtain a number $M(i)$ of solvent conformational basins, each with a probability $p(j,i)$ where $j = 1,...,M(i)$ and $i = 1,...,K$. Once again, we can easily calculate the standard Gibbs free energy (using $p(j,i)$) for all the conformers extracted from the corresponding solute-(solvent)$_N$ landscapes. The total Gibbs free energy can be calculated for each $k-$selected solute-(solvent)$_N$ cluster by summing up the values obtained with the two different ED analyses. Therefore, a total probability $P_{tot}(k)$ can also be associated to each (solute+solvent) conformation. Alternatively, $P_{tot}(k)$ can be obtained by multiplying the probability values resulting from the solute and solute-solvent investigations. Finally, all $P_{tot}(k)$ values are normalized so that their sum is equal to 1.

In the last step of the procedure, the set of selected conformers undergoes the quantum-chemical calculations. First, a partial geometry optimization in internal coordinates is performed at the DFT level. In detail, we constrain the internal degrees of freedom corresponding to the semi-classical motions (*i.e.*, torsional angles) to preserve the extracted conformation. The optimization is followed by the ECD calculation at the TDDFT level. Based on the size and nature of the system, different algorithms [72, 93] can be used to solve the TDDFT equations (see sections 2.4.2 and 2.4.3).

Note that the spectral intensities are calculated in terms of rotatory strength (R_{0n}, in cgs units), although the experimental spectra are more commonly provided in $\Delta\varepsilon$ or $[\Theta]$ units. When using the Casida's scheme, the following expression can be used to convert the calculated data in $\Delta\varepsilon$ units: [184]

$$\Delta\varepsilon(E) = \frac{1}{2.297 \times 10^{-39}} \frac{1}{\sigma\sqrt{\pi}} \sum_n R_{0n} E_{0n} e^{-\left(\frac{E-E_{0n}}{\sigma}\right)^2} \tag{3.3}$$

where $\sigma = \frac{HWHM}{\sqrt{\ln(2)}}$. HWHM is the Half Width at Half Maximum of the Gaussian functions employed to broaden the discrete lines. Instead, the polTDDFT calculation is intrinsically broadened by a Lorentzian function with a lifetime, ε, equals to the experimental HWHM. In this case, the intensity values are modified as:

$$\Delta\varepsilon(E_i) = \frac{1}{22.97} \frac{R_i E_i}{\pi\varepsilon} \tag{3.4}$$

In both cases, $[\Theta]$ can be obtained by using the following relation:

$$[\Theta] = \Delta\varepsilon \times 3298 \tag{3.5}$$

Each ECD is then weighted by the corresponding normalized $P_{tot}(k)$ and summed up to give the statistically averaged ECD. This final ECD can be then compared with the experimental reference and employed for all the investigations.

Chapter 4

Publications

4.1 Main Publications

4.1.1 A computational approach for modeling electronic circular dichroism of solvated cromophores

In this first work, the method at the base of this PhD project, developed in collaboration with Prof. Massimiliano Aschi, was presented. This novel computational procedure was tested on two biomolecules of different size, namely, the cationic GAG$^+$ tripeptide and the neutral (GVGVP)$_2$ decapeptide. Both the systems are soluble in aqueous solution, therefore the water molecules were treated explicitly. Moreover, different temperature conditions were tested, working at 30 °C and 90 °C for GAG$^+$, and at 15 °C and 85 °C for (GVGVP)$_2$. The ECD spectra calculated after extracting the most probable conformers of both the peptide and peptide-(solvent)$_N$ clusters (with $N = 30$ for GAG$^+$, and $N = 40$ for (GVGVP)$_2$) were compared with the corresponding experimental references for a qualitative evaluation. Despite a small shift in energy for the ECD spectra of (GVGVP)$_2$, all the calculations correctly reproduce the experimental features.

The comparison of the spectra revealed the ability of the method to properly capture conformational, solvent and temperature effects. Indeed, two different types of conformers were found for the GAG$^+$ tripeptide, *i.e.* pPII and β-strand conformations, whose population varies with the temperature. In the (GVGVP)$_2$ case, an equilibrium between random coil and β-turn conformers was observed, with the former ones being predominant at lower temperature values. The reliability of the conformational analysis was confirmed by the spectral features which are consistent with these distributions. Furthermore, the role of the solvent on both the conformational analysis and spectra calculation was discussed, corroborating the need of explicitly treating the water molecules for a correct reproduction of the experimental features. Indeed, calculating the ECD spectra with and without the extracted solvation shells, we were able to confirm that the absence of water brings to significant mismatches between the experimental and calculated features.

RESEARCH ARTICLE

A computational approach for modeling electronic circular dichroism of solvated chromophores

Marta Monti[1] 🔾 | Mauro Stener[1] 🔾 | Massimiliano Aschi[2] 🔾

[1]Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste, Italy

[2]Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, L'Aquila, Italy

Correspondence
Mauro Stener, Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy.
Email: stener@units.it

Massimiliano Aschi, Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, Via Vetoio, 67100 L'Aquila, Italy.
Email: massimiliano.aschi@univaq.it

Abstract

The present study consists in a novel computational protocol to model the UV-circular dichroism spectra of solvated species. It makes use of quantum-chemical calculations on a series of conformations of a flexible chromophore or on a series of chromophore/solvent clusters extracted from molecular dynamic simulations. The protocol is described and applied to the aqueous cationic tripeptide GAG^+ and to the aqueous neutral decapeptide $(GVGVP)_2$. The protocol has proven able to: (i) properly consider the conformational motion of solute in the given environment; (ii) give the actual statistical weight of each conformational state; (iii) provide a reliable quantum mechanical method able to reproduce the spectral features. Temperature effects on conformations and spectral properties are properly taken into account. The role of explicit solvent on the conformational analysis and the spectra calculation is discussed. The comparison of the calculated circular dichroism spectra with experimental ones recorded at different temperatures represents a strict validation test of the method.

KEYWORDS
CD spectroscopy, DFT, essential dynamics, molecular dynamics, TDDFT

1 | INTRODUCTION

The theoretical modeling of electronic circular dichroism (ECD) has attracted much interest in the last decades, due to its usefulness to ascribe the structure to absolute enantiomers of chiral compounds.[1] Indeed, if a reliable protocol would be available to calculate the ECD, a direct comparison of the calculated ECD with the experimental one would allow to easily ascribe the absolute structure of the enantiomer. The possibility to employ such a strategy in practice depends on the reliability of the computational protocol. At the moment, ECD spectra can be calculated with enough accuracy by means of the Time Dependent Density Functional Theory[2] employing accurate exchange-correlation functionals, typically hybrid ones such as B3LYP[3,4] or more elaborate ones, for instance including Range Separated functionals

such as CAM-B3LYP.[5] Such a scheme is generally adequate for the treatment of the electronic problem in organic compounds. However, it must be considered that the ECD is extremely sensitive to the conformations of the system under study,[1] and that the conformations themselves are further sensitive to the presence of the solvent, a situation which is further amplified when hydrogen-bonds are present such as in polypeptides. For this reason, a valid computational protocol for the calculation of the ECD not only needs an accurate electronic scheme, but also requires a physically coherent inclusion of the contributions of several conformations as well as of the solvent effects, at least at a partial extent.[6,7]

Therefore, many scientists have worked on this topic and the scientific literature in the field is quite wide. Grimme et al. suggested a simplified TDDFT (sTDDFT), which is so cheap to allow many

calculations on several snapshots taken from a preliminary molecular dynamics (MD)[8] calculation, or an automated procedure to treat flexible molecules.[9] From a different perspective, Caricato et al. have studied the problem of the gauge choice in optical rotation.[10–14] Mennucci et al. worked on excitonic models[15] while Cappelli on the chiroptical properties in aqueous solution.[16] Chiroptical properties in solution have been treated also by Barone et al.,[17] who recently suggested an automated procedure based on evolutionary algorithm.[18] Autschbach et al. wrote a recent review on optical activity.[19] Crawford has studied the effect of the basis set choice on optical rotation[20] and included solvent effects in a combined MD + DFT approach.[21] Finally, Ruud extended the methods for optical activity to the relativistic case.[22,23]

Most of the above methods treat the contributions of the different conformers of the chromophore by a preliminary MD search, which allows to properly span the associated conformational repertoire. Subsequently the ECD spectra are calculated, a posteriori, for all conformers and properly weighted in order to obtain a Boltzmann averaged spectrum, which can be directly compared with the experiment. Such a procedure is in principle very accurate provided four basic mutually interconnected conditions are fulfilled: first, a sufficiently accurate MD simulation can be carried out for an adequately long time, typically from dozens to several hundreds of nanoseconds depending on the size and conformational flexibility of the chromophore of interest; second, the chromophore itself can be safely defined; third, an exhaustive conformational analysis can be performed and, finally, a valid but still affordable quantum-chemical approach is available. Concerning the first argument, the reliability of the MD simulation—or more in general of whatever tool for the semiclassical span of a conformational space—obviously depends on the accuracy of the force-field. On the other hand, concerning the second argument, when considering chromophores in strong interaction with the solvent (e.g., through H-bond), it might be necessary to include explicit solvent molecules in the definition of the chromophore itself. It follows that, in any case, the geometrical identification and the subsequent extraction from the MD simulation of the relevant chromophore conformations (the third argument), irrespectively from its definition, might not represent a straightforward task. As a matter of the fact if we exclude the trivial case of rigid chromophores not strongly interacting with the solvent or the less trivial, although still affordable, cases of chromophores whose conformational transitions can be easily followed by a reduced set of internal coordinates (e.g., the Ramachandran plot, see for example Reference 24), the unbiased conformational analysis might represent the actual bottleneck of the whole procedure. For this purpose, the essential dynamics (ED)[25,26] analysis could represent a powerful tool for obtaining a reduced, and hence computationally affordable, number of internal degrees of freedom. ED, which consists in the construction of the positional covariance matrix (CM) of a chromophore—excluding the roto-translations—of arbitrary size, is usually employed to describe the fluctuations of a covalent framework (i.e., a single flexible chromophore), however recently it has been also extended for addressing the conformational repertoire of weakly interacting systems such as the

previously mentioned clusters of solvated chromophores.[27] In the present work, we provide a computational protocol for the calculation of ECD spectra of solvated species based on quantum-chemical calculations carried out on a number of conformations extracted through MD-ED analysis and also taking into account the possible effects of the inclusion of the explicit solvation shells in interaction with the chromophore. Note that with "conformations" in this study we identify well-defined spatial organizations not only of the chromophore but also of the cluster formed by the chromophore and a predefined number of solvent molecules. For this purpose, we decided to use, as test cases, the aqueous cationic tripeptide (GAG$^+$) and the aqueous neutral decapeptide (GVGVP)$_2$ (see Figure S1) whose experimental ECD data[28–31] are available also at temperature values different from the room temperature. Most importantly, reliable force-fields are available for these systems and, moreover their conformational features have been extensively investigated in the past.

The possibility to compare the calculated and the experimental ECD spectra also at different physical conditions certainly represents a stringent test for the presented computational protocol. The work is organized as follows: first, an outline of the employed theoretical methods is given (MD, ED, DFT, and TDDFT), then the results are presented according to a logical ordering starting from MD results, ED analysis, and ECD spectra for both systems. This part is repeated twice with and without explicit solvent molecules, in order to assess with more confidence the importance of the solvent effects. Finally, a direct comparison with the available experimental data is performed, in order to discuss the temperature effect and to identify all the possible drawbacks and source of errors.

2 | COMPUTATIONAL DETAILS

2.1 | Molecular dynamics

All the molecular dynamics (MD) simulations were carried out using the Gromacs package,[32] version 5.1.2. The peptides (either cationic GAG or GVGVPGVGVP) were described using the OPLS-AA force field.[33,34] For water solvent, the single point charge (spc) model was adopted.[35] A proper number of counterions (chloride ions) was also added for ensuring the electroneutrality of the box. It is important to remark that the choice of the force-field is essential for the final outcome. However, in this first study, we did not take into account the possible effects induced by the use different force fields on the quality of the final spectra.

The temperature was kept constant using the velocity rescaling model[36] and the LINCS algorithm[37] was employed to constrain all the bond lengths. Long range electrostatic interactions were computed by the Particle Mesh Ewald method[38] with 34 wave vectors in each dimension and a 4th order cubic interpolation and a cut-off of 1.0 nm was used. All the simulations were carried out in the NVT ensemble. The solute (either GAG$^+$ or (GVGVP)$_2$) was inserted in a cubic box preventively adjusted to correctly reproduce the correct density, in infinite dilution in water, at the temperature of interest and 1.0 bar of

pressure. This latter condition has been achieved by modifying the solute-solvent box in order to reach the same average pressure of a box of pure water (i.e., with the same number of spc particles corresponding to 600 and 800 for GAG^+ and $(GVGVP)_2$, respectively) preventively simulated in the NVT ensemble at the experimental density of the pure water at the temperature of interest. In this respect, for the pure water simulations, we utilized the following densities at the corresponding temperatures: 999.06 g/L (15°C), 995.61 g/L (30°C), 968.59 g/L (85°C), 965.30 g/L (90°C). The solvated peptide simulations (hereafter free-MD), after the initial mechanical and thermal equilibration disregarded from the analysis, were propagated for 40 and 100 ns for GAG^+ and $(GVGVP)_2$, respectively. Additional simulations (termed as constrained-MD, see below) were subsequently carried out upon keeping frozen the peptide in each of the different conformations previously extracted from the free-MD as described in the next paragraph.

2.2 | Essential dynamics

The conformational repertoire of the chromophore (irrespectively of its definition, see above) was accomplished by analyzing the MD simulations using ED whose basic features are widely described in the literature[25,26] and here only briefly outlined. In the case of the free-MD, the peptide positional CM is first constructed using, in the present case, the backbone atoms and then diagonalized. This operation produces a set of eigenvectors, that is, a set of eigendirections along which the peptide undergoes the internal fluctuations, with the associated eigenvalues corresponding to the actual value of the mean square fluctuations. Therefore, the eigenvectors characterized by the highest eigenvalues represent the eigendirections describing the fluctuations with largest amplitude necessary, that is, essential, to account for most of the peptide conformational transitions. If the number of these eigenvectors, hereafter termed as Essential Eigenvectors, is limited to one or two, the corresponding projection of the peptide trajectory on these eigenvectors, that is, the Principal Components, provides us with a conformational landscape (with respect to a single coordinate or to a plane) relatively easy to visualize and to analyze. In this respect, the regions of the conformational landscape more frequently spanned by the peptide represent the conformational basins.

This allows to directly calculate the statistical weight of the ith conformation hereafter termed as $P(i)$ and corresponding to the number of projected points falling within the ith basin divided by the total number of frames. From $P(i)$ we can also estimate the free energy difference between the ith basin and a reference basin (with probability P_{ref}) using the standard relation

$$P(i) = P_{ref} e^{\frac{-\Delta G_i^\circ}{RT}}, \tag{1}$$

ΔG_i° approximately corresponds to the (standard) Gibbs free energy difference between the ith basin (conformation) and the reference one, assuming: (i) a negligible difference between the corresponding partial molar volumes (it should be further remarked that our simulations are performed at constant volume and not at constant pressure) and (ii) negligible differences between the (quantum) vibrational partition functions of the ith conformation and the reference one. Subsequently, from all the basins, we can extract a number (hereafter indicated with K) of peptide structures in the range 0–3 kJ/mol, hence obtaining the statistically relevant peptide conformations representing the whole basin. Note that when we are dealing with a relatively large basin, more than one representative structure could be extracted (see Section 3). Each of the K extracted conformations characterized by the proper statistical weight and by the corresponding free energy (ΔG_i°, $i = 1, K$) are then used for the quantum-chemical calculations for determining the observable of interest (e.g., a spectral feature). Note that this latter step should be carried out after a preliminary constrained geometry optimization of the peptide (possibly at the same level of the quantum-chemical approach used for the calculation of the observable), starting from the MD-extracted structure, performed by keeping frozen the internal degrees of freedom corresponding to semi-classical motions, that is, typically the torsion angles.

In order to take into account also the possible contribution of the solvent molecules[28] in the observable of interest, we decided to adopt the same protocol just described for locating plausible peptide-solvent clusters to be used for subsequent quantum-chemical calculations. For this purpose, we performed a further analysis, again based on the ED described in the detail elsewhere[27] and here only briefly outlined.

For each of the K conformations previously extracted from the free-MD we first perform a MD simulation, with the same box and protocol used for free-MD, with the peptide kept frozen at the center of the box. We hereafter refer to these simulations as constrained-MD. For each of the K constrained-MD we then construct an internal reference frame, centered in the frozen peptide geometrical center. The projection of the frozen peptide coordinates onto the above unit vectors provides the axes of the ellipsoid best describing the peptide shape (in this case a fixed ellipsoid for each constrained-MD). At each of the constrained-MD trajectory, a pre-selected number (N) of solvent molecules showing the lower square distances in the metric of the above ellipsoid, are then extracted. By repeating this last step for all the frame of the constrained-MD we can obtain a trajectory of the peptide-(solvent)$_N$ cluster. The ED analysis, previously described for the single peptide and now carried out on such a cluster, as well as the consequent Principal Components, provides the conformational landscape of the peptide-(solvent)$_N$ cluster, that is, the location of a series of basins now corresponding to solvation shell conformations. Therefore, considering the ith constrained-MD, which showed a probability $P(i)$ in the free-MD, we can obtain, from the corresponding ED analysis, a number $M(i)$ of solvent conformational basins, each with probability $p(j,i)$, with $j = 1$ to $M(i)$ and $i = 1$ to K. Obviously the sum of all the $p(j,i)$ is equal to $P(i)$. Also in this case the relative free energy of each peptide-(solvent)$_N$ cluster conformation can be easily calculated from the standard expression previously reported and now making use of the obtained $p(j,i)$. Each of these conformations is finally extracted and optimized through quantum-chemical calculations using internal constraints for maintaining both the peptide conformation of the ith basin and the relative peptide-water positions and orientations.

The obtained $M(1) + M(2) + \cdots + M(K)$ clusters, with the related statistical weight, $p(i,j)$, are then used for calculating the observable of interest, that is, the spectral intensities (see the next paragraphs) are considered as the weighted sum of the spectra of each conformational state (either the single chromophore of the cluster). Given the size and the shape of the peptides we have used 30 and 40 solvent molecules for cationic GAG$^+$ and (GVGVP)$_2$, respectively. An additional analysis of the conformational space is considered in Figure S2 while the atomic coordinates of the clusters have been reported in Tables S1–S4.

2.3 | Quantum chemical calculations

Once a set of significant number of representative conformations has been identified by the ED analysis, a partial geometry optimization in internal coordinates of their MD non-equilibrated geometries is performed at the density functional theory (DFT) level.[39] The DFT calculations have been performed employing the ADF engine of the AMS code.[40] All the geometry optimizations, with internal constraints as described in previous Section 2.2, have been performed employing a basis set of Slater-type orbitals of triple-zeta quality (STO-TZP) and the range-separated exchange-correlation (xc) functional wB97X-D. Such functional includes both the dispersion and non-local exchange asymptotical corrections, allowing a suitable description of the physics of the system (i.e., intramolecular, intermolecular hydrogen bonds and charge-transfer excitations).[41]

Finally, the calculations of the UV-electronic circular dichroism (UV-ECD) and UV-Photoabsorption spectra have been performed at the Time-Dependent DFT (TDDFT) level analyzing the lowest 70 and 100 excited-states for GAG$^+$ and (GVGVP)$_2$, respectively. In order to allow an easy comparison with the experimental data,[28–31] the calculated spectra have been reported in terms of ε (photoabsorption) and $\Delta\varepsilon$ (ECD). The individual UV-ECD spectra relative to the most relevant conformations in presence of the explicit solvent are reported in Figures S3–S8. To this end, a convolution of all the spectra has been realized by employing Gaussian functions of appropriate half width at half maximum (HWHM; i.e., 0.4 eV for GAG$^+$, and 0.3 eV for (GVGVP)$_2$) in order to have the best consistency with the experimental resolution. Since the calculated spectra consist of discrete lines in terms of oscillator strength (f_{0k}, in atomic units) for photoabsorption and rotatory strength (R_{0k}, in cgs units) for ECD, the following expressions have been employed to calculate ε and $\Delta\varepsilon$[42,43]:

$$\varepsilon(E) = 2.870 \times 10^4 \frac{1}{\sigma\sqrt{\pi}} \sum_k f_{0k} e^{-\left(\frac{E-E_{0k}}{\sigma}\right)^2}, \quad (2)$$

$$\Delta\varepsilon(E) = \frac{1}{2.297 \times 10^{-39}} \frac{1}{\sigma\sqrt{\pi}} \sum_k R_{0k} E_{0k} e^{-\left(\frac{E-E_{0k}}{\sigma}\right)^2}, \quad (3)$$

where in expressions (2) and (3) E_{0k} is the calculated excitation energy (in eV) from the ground state to the kth excited state, and σ is related to HWHM according to the following relation (4):

$$\sigma = \frac{\text{HWHM}}{\sqrt{\ln(2)}}. \quad (4)$$

This procedure allows a direct comparison with the experimental spectra not only in terms of excitation energy but also in terms of absolute intensity.

The UV-CD spectra of the most probable conformations at low temperature values (i.e., GAG$^+$(H$_2$O)$_{30}$ at 30°C, and (GVGVP)$_2$(H$_2$O)$_{40}$ at 15°C) were also calculated with the hybrid xc functional CAM-B3LYP[44] to evaluate the quality of our choice (wB97X-D) which was based on the data available in the literature.[30] The effect of the functional on the calculations is briefly discussed in Figure S9. We also evaluated the mean-field produced by the solvent using the conductor like screening model (COSMO) of solvation[45] as implemented in ADF. Such scheme was adopted for the tripeptide conformers (extracted from the ED analysis) at 30°C, considering both the GAG$^+$ by itself and the GAG$^+$ solvated with 30 water molecules. The resulting statistically weighted ECD were compared and discussed in Figure S10.

The procedure just presented has been schematized in the flow-chart shown in Figure 1 to help the visualization and comprehension of the multistep scheme we proposed.

3 | RESULTS AND DISCUSSION

Although the main focus of the present study is to provide a computational strategy for addressing the spectra of relatively complex species in solution, we deserve the initial part of this section to briefly outline the preliminary conformational/structural analysis of the species of interest.

3.1 | Free-MD simulations and peptides conformational states

In Figure 2, we report the spectrum of the eigenvalues from the diagonalization of the peptide (backbone) covariance matrix as obtained from the free-MD simulations of the two peptides at the temperatures of interest. We first observe that, as expected, the spectrum for (GVGVP)$_2$ is characterized by much larger eigenvalues (approximately 50 times larger than GAG$^+$) indicating that peptides with a larger size undergo much higher fluctuations (see also the Figure 3). At the same time, it is also worth of comment that in both the systems the change of temperature does not induce a drastic change in the shape, that is, in the trace of the covariance matrix and hence in the whole backbone fluctuation. As a matter of fact when passing from the lower to the higher temperature a sharp although not dramatic increase of the highest eigenvalues is systematically observed. Finally, we also wish to remark that, as usual for relatively small peptides,[26] a large fraction (between 65% and 75%) of the whole peptide fluctuation, corresponding to the trace of the covariance matrix, can be safely described using only two eigenvectors of the backbone covariance matrix.

FIGURE 1 Flowchart of the procedure proposed in this work. The three methods (i.e., MD simulations, ED analysis, and quantum-chemical calculations) are reported together with the specific conditions we used to realize the tests on the two peptides

Free-MD Simulations
40 ns (GAG$^+$) or 100 ns (GVGVP$_2$) MD in water at 15, 30, 85 and 90°C

ED Analysis
(1) ED analysis of the peptide backbone conformational space
(2) Extraction of a number (K) of peptide conformers (with their own probability P_i)

Constrained-MD Simulations
(1) 10 ns MD (same protocol and box of the free-MD), for each of the K extracted structures, with the peptide kept frozen at the center

ED Analysis
(1) Selection of the number (N) of solvent molecules
(2) ED analysis of the peptide-(solvent)$_N$ clusters conformational spaces
(3) Extraction of a number ($M(K)$) of clusters (with their own probability $p_{j,i}$)

Quantum-Chemical Calculations
(1) Constrained geometry optimizations (DFT)
(2) UV-ECD and UV-photoabsorption calculations (TDDFT)
(3) Statistical weighting of the spectra using $p_{j,i}$

FIGURE 2 Eigenvalues of the GAG$^+$ backbone covariance matrix at $T = 30°C$, and $T = 90°C$ (left panels); (GVGVP)$_2$ backbone covariance matrix at $T = 15°C$ and $T = 85°C$ (right panels)
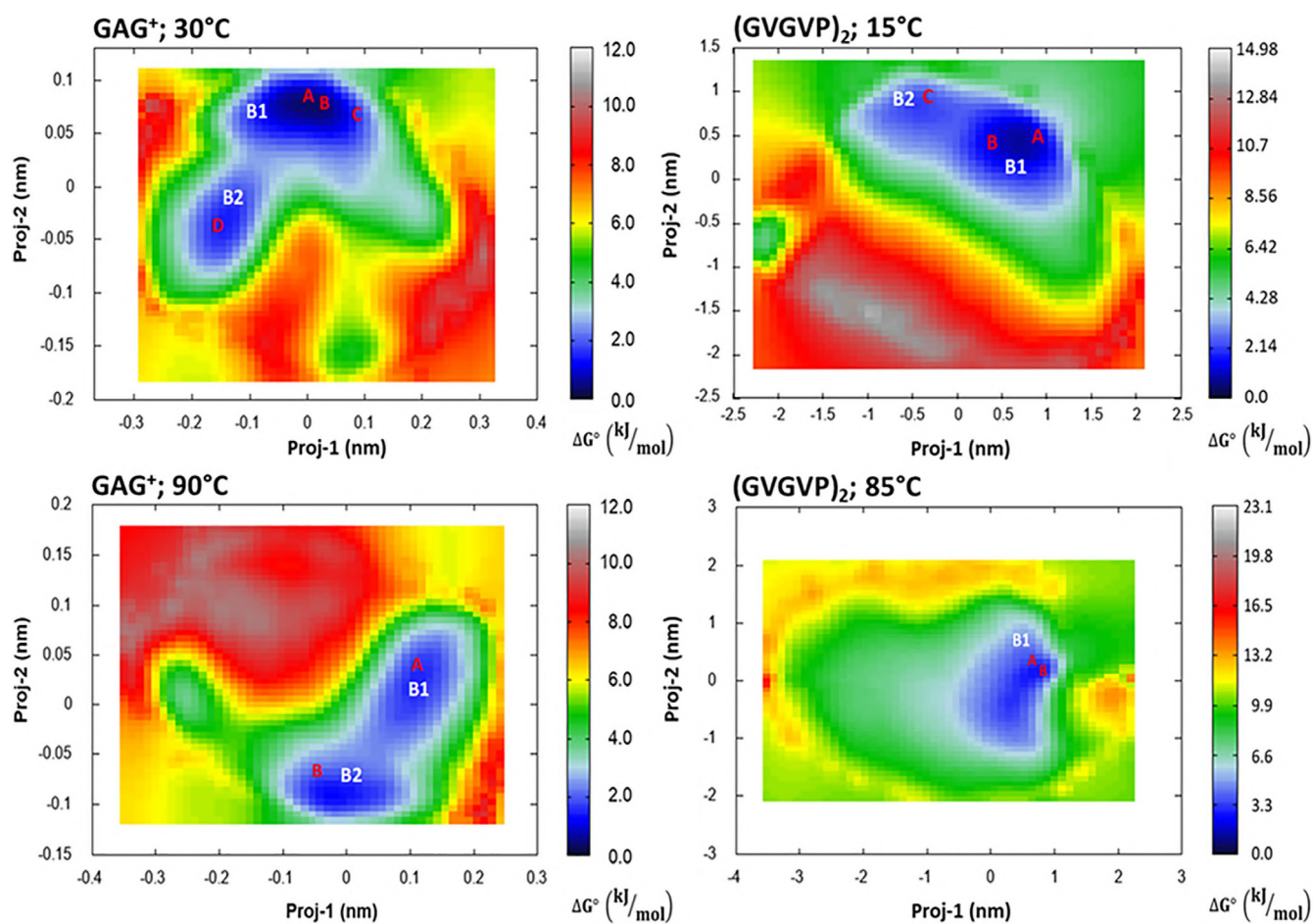
FIGURE 3 Free energy landscape (Equation 1 in kJ/mol) in the essential plane of GAG$^+$ at $T = 30°$C, and $T = 90°$C (left panels); (GVGVP)$_2$ at $T = 15°$C, and $T = 85°$C (right panels). The scale of the energy values is also reported as a vertical-colored bar

TABLE 1 Relative free energies, at the given temperatures, for GAG$^+$ and (GVGVP)$_2$ basins extracted from free-MD simulations (see also Figures 3 and 4)

Peptide	Temperature (°C)	Basin – solute structure	Relative free energy (kJ/mol)
GAG$^+$	30	B1 – A	0.7
GAG$^+$	30	B1 – B	0.0
GAG$^+$	30	B1 – C	2.4
GAG$^+$	30	B2 – D	1.6
GAG$^+$	90	B1 – A	1.6
GAG$^+$	90	B2 – B	0.0
(GVGVP)$_2$	15	B1 – A	1.8
(GVGVP)$_2$	15	B1 – B	2.9
(GVGVP)$_2$	15	B2 – C	2.7
(GVGVP)$_2$	85	B1 – A	0.0
(GVGVP)$_2$	85	B1 – B	2.9

The corresponding Principal Components were then used for obtaining the free energy landscapes reported in the Figure 3.

In the same Figure 3, we have also shown the conformational basins whose relative free-energies, calculated using Equation (1), are reported in Table 1. The corresponding basin representative structures, extracted for subsequent quantum-chemical calculations are reported in the Figure 4.

First of all, it is interesting to observe, in the Figure 3, the different Principal Components values further reflecting the already remarked differences in the peptides whole fluctuation: the proj-1 and proj-2 domains become much wider when passing from GAG$^+$ to (GVGVP)$_2$ and, although to a lesser extent, upon the temperature increase. In the case of GAG$^+$, we also systematically observe the presence of a bimodal distribution, that is, the
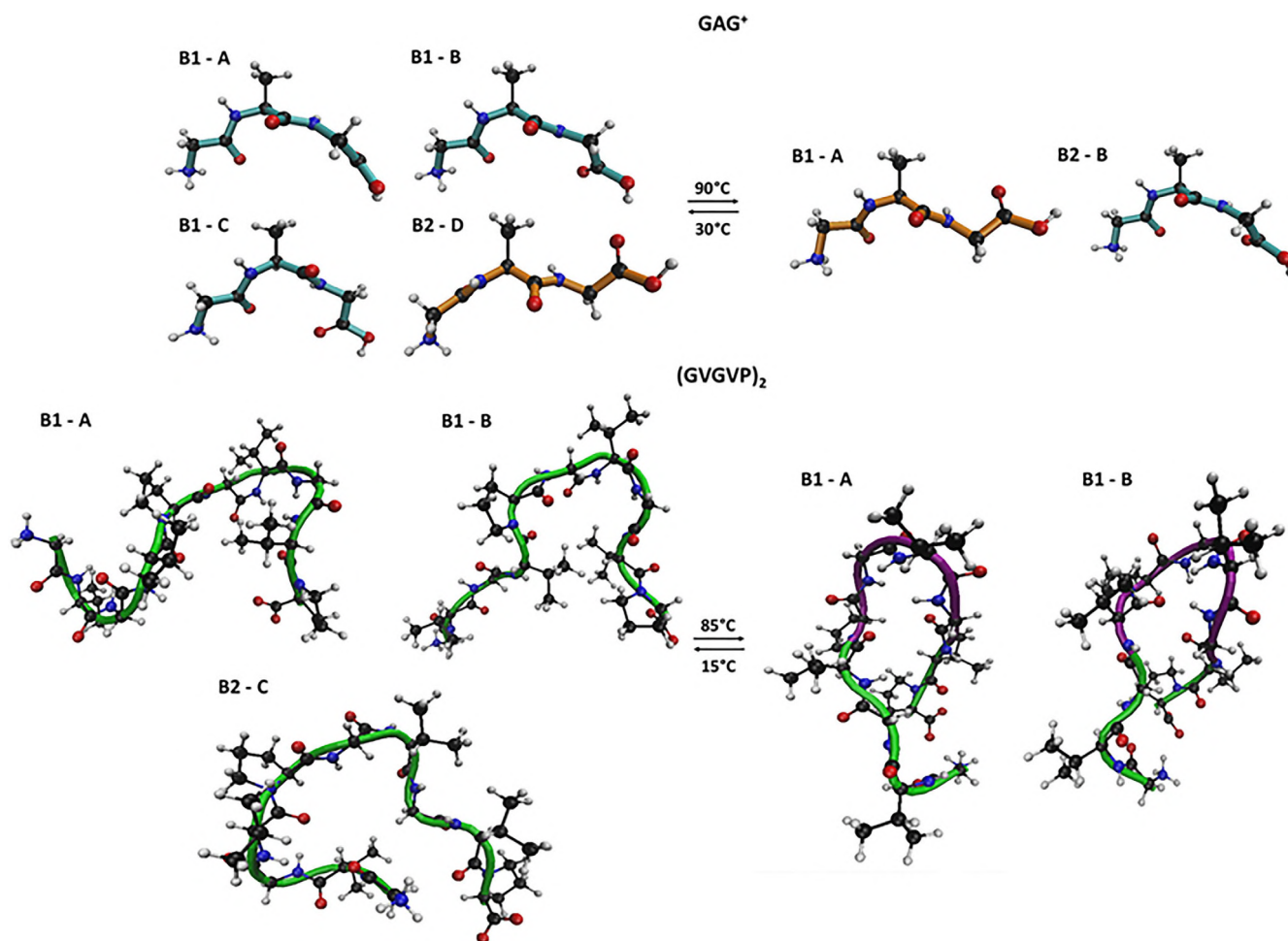
FIGURE 4   Schematic picture of the representative structures extracted by the free energy landscape of Figure 3

presence of two low (free) energy conformational states in relatively fast interconversion.

In line with the experimental data GAG$^+$ shows a relatively high propensity of adopting a pPII state represented by the basin B1 at 30°C. Moreover, confirming the low-entropy character of pPII conformation, upon temperature increase such a motif (B1) becomes comparable in stability with the other structure B2, which more closely resembles the beta-strand one. On the other hand, the conformational space of (GVGVP)$_2$ appears as most sensitive to the temperature hence suggesting a more relevant entropic weight. In particular at higher temperature the conformations identified by our analysis strongly resemble beta-turn structures both characterized by a relatively stable Val4-Gly8 H-bond. On the other hand, at lower temperatures the structures appear as slightly more extended, and almost completely devoid of intramolecular H-bonds. In conclusion: our simulations confirm that the investigated systems exist in two main conformations (i.e., pPII and beta-strand for GAG$^+$; random coil and β-turn for (GVGVP)$_2$) whose ratio depends on the temperature of work.[28,29,31] These systems have been then utilized with their own statistical weight, as obtained from the relative free energies, for the quantum-chemical calculations as described in Section 2.3 and, at the

same, for the subsequent analysis of the solvation shells as described in Section 2.2 and reported in the next paragraph.

3.2 | Constrained-MD and water-peptide clusters conformational states

In the Figure 5, we have reported the spectra of the eigenvalues from the diagonalization of the positional covariance matrix for the constrained-MD trajectories, as obtained from the procedure previously described. Differently from the single peptide (Figure 2) in this case the spectra of the eigenvalues appear as much less steep confirming that when a non-covalent system like the peptide-(solvent)$_N$ cluster is concerned, the overall fluctuation is (not surprisingly) spread over a higher number of internal degrees of freedom.[27]

It follows that in this case the conformational analysis, and hence the identification of the conformational basins should be carried out in an N-dimensional ($N > 2$) hyperspace. This procedure, although possible in principle[46] can represent a very complicated task and, hence, a compromise between quality of the result and computational difficulty is necessary. For this reason, also in this case we projected the
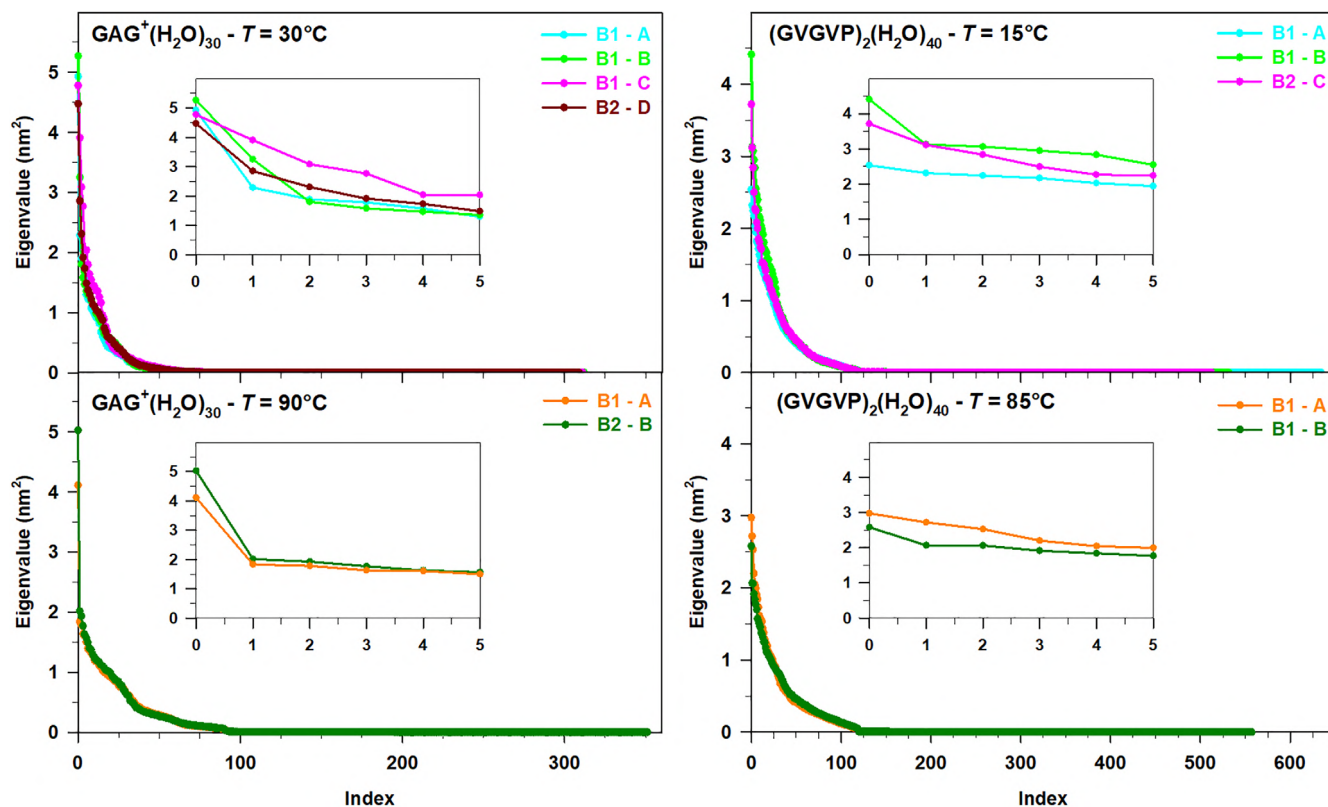
FIGURE 5 Spectrum of the eigenvalues of the covariance matrix of the peptide-(solvent)$_N$ clusters extracted from the constrained-MD simulations. The first six eigenvalues have been highlighted in the inset for all the systems

water trajectory, with respect to the frozen peptide, onto the plane formed by the first two eigenvectors of Figure 5, being well aware of the possible lack of some details for identifying the conformational repertoire of the water molecules in closer interaction with the peptide.

The free-energy landscapes, with the corresponding free-energy basins also in this case obtained using Equation (1), are shown in Figure 6. In this case, a general increase of the Principal Components, if compared to the single peptide results (Figure 3) is observed. Such a result obviously reflects the already commented spectrum of the cluster covariance-matrix eigenvalues and, more precisely, the much higher mobility of the water molecules and the involvement of an increased number of internal degrees of freedom. Additional details on the selected (low free-energy) basins indicated in Figure 6 are collected in Table 2. The representative structures of these basins, whose details can be found in the Supporting Information, were then used—with their own statistical weight and after a constrained optimization—for the calculation of the observables of interest as reported in the next paragraph.

3.3 | UV-CD spectra

The UV-CD spectrum has been calculated for each of the structures representing the basins collected in Table 2, then the overall spectra

have been obtained by summing the individual ones weighted accordingly to the procedure described in previous Section 3.2. Hence, four final UV-CD spectra (two for GAG$^+$, and two for (GVGVP)$_2$) have been compared with the experimental ones available in the literature.[28,29,31]

The general assessment of the spectra reported in the four panels of Figure 7 highlights the overall accuracy of our model, which allows us to coherently simulate the behavior of the selected systems. Indeed, the salient features of the experimental UV-CD spectra are reproduced by the calculated ones, particularly well for the GAG$^+$ system as shown in the left panels of Figure 7. At both temperature values, the experimental minimum around 190 nm has been observed in the corresponding calculated spectra, although some discrepancies in terms of absolute $\Delta\varepsilon$ can be found at 90°C (lower left panel of Figure 7). Moving toward lower energy values the calculated spectra follow nicely the increasing trend, which reaches its maximum around 212 nm. On the other hand, the experimental UV-CD spectrum of (GVGVP)$_2$ in water at 15°C reveals the presence of two minima at 197 and 218 nm, respectively, the first one is found as well in the calculated spectrum at 203 nm, while the calculated counterpart of the second minimum is a shallow inflection at 224 nm. Despite this moderate energy shift, once again the model takes into account the physics of the described system. At higher temperature values, the experimental spectrum is mostly characterized by the minimum at 218 nm, well reproduced by the theory at 211 nm.
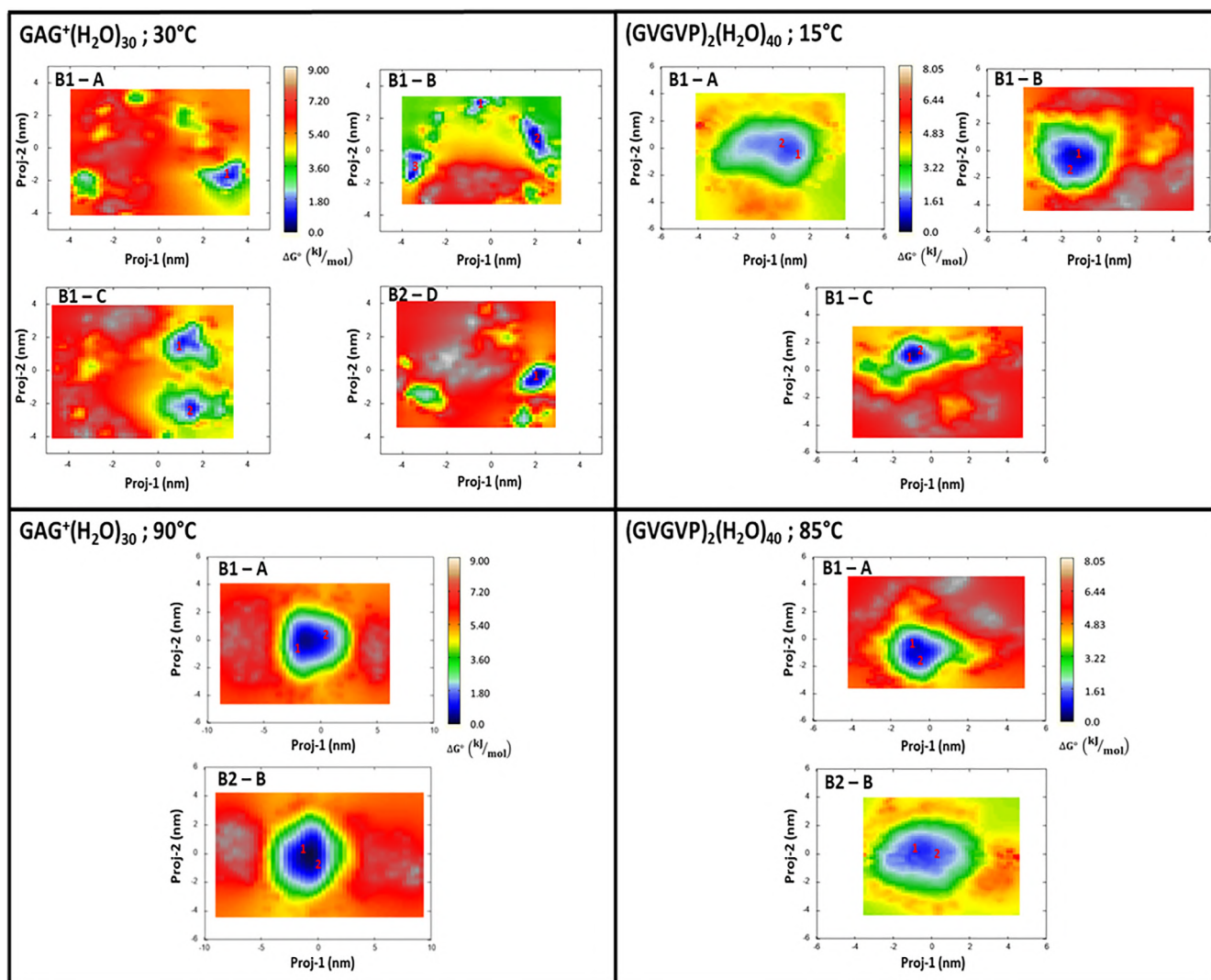
FIGURE 6 Free energy landscape (Equation 1, in kJ/mol) in the essential plane of GAG$^+$ clusters at $T = 30°C$, and $T = 90°C$ (left panels) and (GVGVP)$_2$ cluster at $T = 15°C$, and $T = 85°C$ (right panels). The peptide conformation (see Figure 4) is indicated in the upper left-side of each inset. The scale of the energy values is also reported as a vertical-colored bar

Now we may focus in more detail on the relations among the conformational space, the optical response, and the temperature. Along the ED analysis, we have already pointed out the influence of the temperature on the conformers, while variations of the spectroscopic response related to the temperature have just been mentioned. Therefore, all the parameters can be collected together analyzing the UV-CD spectra in terms of both temperature and conformational dependence, thus verifying if our model considers all the effects in a balanced way. For such an analysis, it is convenient to compare together the experimental CD spectra at different temperatures in order to point out qualitative trends, and then to verify if such trends are reproduced, at least partially, by the calculated spectra.

The experimental UV-CD spectra recorded for GAG$^+$ in water at 30 and 90°C show similar patterns, with the only significant difference related to the intensity. Indeed, both the spectra are characterized by a minimum point around 190 nm followed by an increasing trend at lower energy values. Looking at Figure 8 (lower left panel), it

can be easily noticed that the calculated UV-CD spectra closely resemble such behavior only showing a more marked amplitude in intensity, especially around the minimum region. Nevertheless, the discrepancies in terms of intensity distributions are so modest that do not affect the general qualitative agreement between theory and experiment. Therefore, we can conclude that the calculated spectra well reproduce the decrease of CD that occurs increasing the temperature, which can be interpreted considering the two possible conformations of the tripeptide GAG$^+$. It is known from the literature[28,29] that these short peptides mostly exist in unordered conformations which are in this specific case the pPII and β-strand structure. The populations of the conformers are strongly influenced by the temperature, thus determining the preference for the first conformation at lower T values, and for the latter one at higher T values. Hence, the reduction of the signal at the minimum corresponds to a decrease of the pPII population within the conformational space, also found and supported by our theoretical approach. Indeed, the peptide structures

Peptide	Temperature (°C)	Basin – cluster	Relative free energy (kJ/mol)
GAG$^+$	30	B1 – A1	2.2
GAG$^+$	30	B1 – B1	0.0
GAG$^+$	30	B1 – B2	2.4
GAG$^+$	30	B1 – B3	1.5
GAG$^+$	30	B1 – C1	3.7
GAG$^+$	30	B1 – C2	4.8
GAG$^+$	30	B2 – D1	4.1
GAG$^+$	90	B1 – A1	1.6
GAG$^+$	90	B1 – A2	3.9
GAG$^+$	90	B2 – B1	0.1
GAG$^+$	90	B2 – B2	1.9
(GVGVP)$_2$	15	B1 – A1	1.8
(GVGVP)$_2$	15	B1 – A2	3.3
(GVGVP)$_2$	15	B1 – B1	3.4
(GVGVP)$_2$	15	B1 – B2	3.8
(GVGVP)$_2$	15	B2 – C1	3.0
(GVGVP)$_2$	15	B2 – C2	4.3
(GVGVP)$_2$	85	B1 – A1	0.0
(GVGVP)$_2$	85	B1 – A2	0.5
(GVGVP)$_2$	85	B1 – B1	3.6
(GVGVP)$_2$	85	B1 – B2	4.8

TABLE 2 Relative free energies, at the given temperatures, for GAG$^+$(H$_2$O)$_{30}$ and (GVGVP)$_2$(H$_2$O)$_{40}$ cluster basins extracted from constrained-MD simulations (see Figure 6)
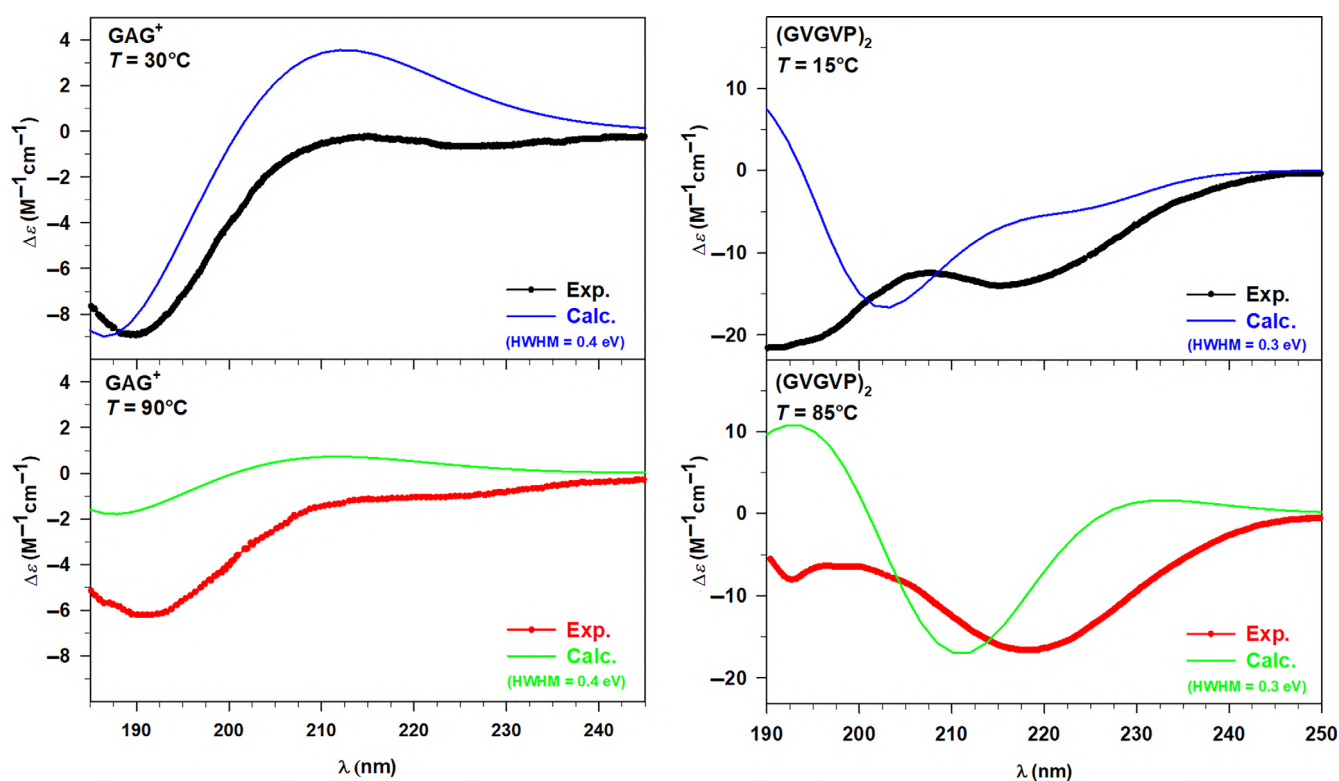
FIGURE 7 Left panels: Experimental (Exp.) and calculated (Calc.) UV-circular dichroism (UV-CD) spectra of GAG$^+$ in water at 30°C (upper panel) and 90°C (lower panel). Right panels: Experimental (Exp.) and calculated (Calc.) UV-CD spectra of (GVGVP)$_2$ in water at 15°C (upper panel) and 85°C (lower panel). The GAG$^+$ and (GVGVP)$_2$ calculated spectra have been convoluted with Gaussian functions with HWHM of 0.4 and 0.3 eV, respectively
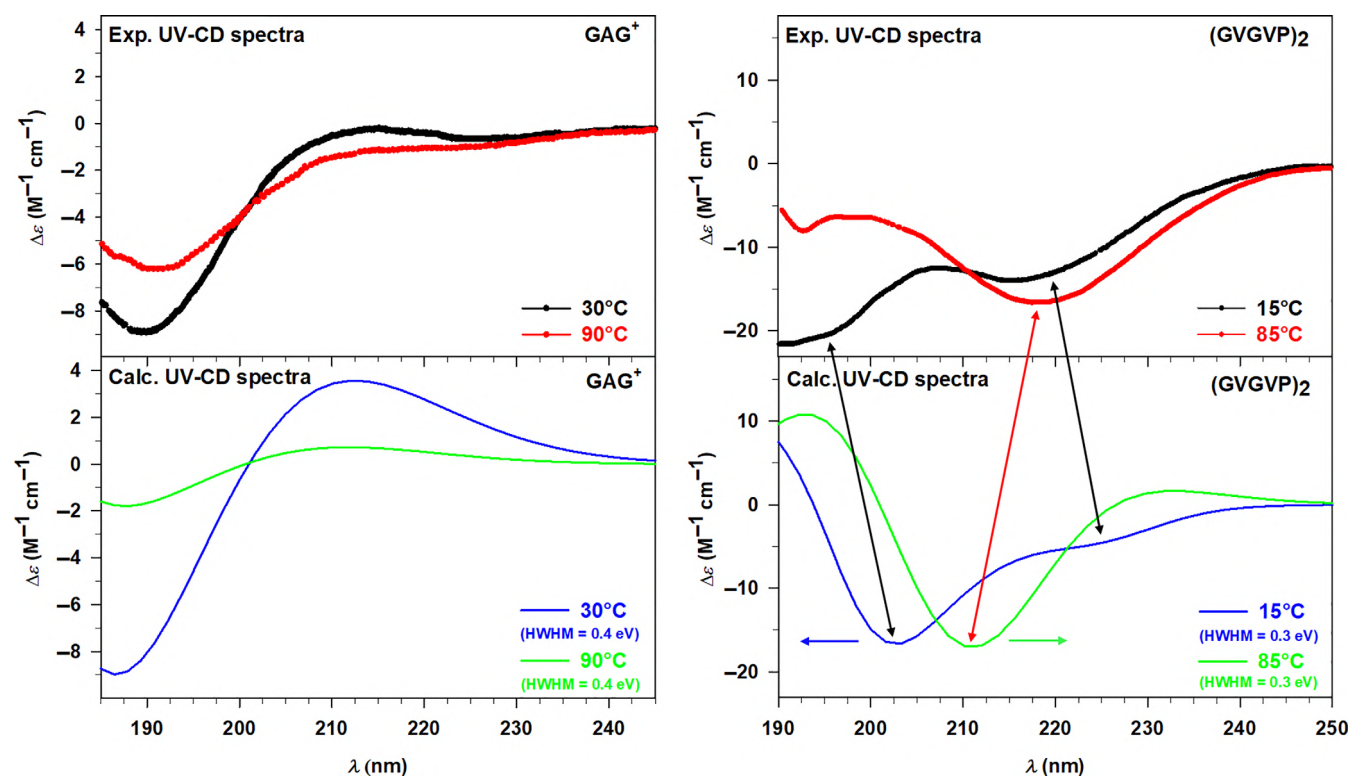
FIGURE 8 Left panels: Experimental (Exp.) UV-circular dichroism (UV-CD) spectra at 30 and 90°C (upper panel) and calculated (Calc.) UV-CD spectra at 30 and 90°C (lower panel) of GAG^+ in water. Right panels: Experimental (Exp.) UV-CD spectra at 15 and 85°C (upper panel) and calculated (Calc.) UV-CD spectra at 15 and 85°C (lower panel) of $(GVGVP)_2$ in water. The GAG^+ and $(GVGVP)_2$ calculated spectra have been convoluted with Gaussian functions with HWHM of 0.4 and 0.3 eV, respectively

(see Figure 4 of the Section 3.1) that we have selected at each temperature from the ED analysis show how the conformations move from having the folded extremities (pPII) toward a more extended structure (β-strand), corroborating the results just discussed.

Despite the evident influence of the conformations on the optical response, the GAG^+ system is characterized by two possible similar structures, both unordered. Therefore, we have chosen the larger system $(GVGVP)_2$ which, as known from the literature,[28,29] exhibits two possible conformations that present much more marked differences. From this perspective, we can discuss the temperature effect on the UV-CD spectra of $(GVGVP)_2$ in correlation with its possible conformations, starting from the experimental data shown in the upper right panel of Figure 8. The two UV-CD spectra of $(GVGVP)_2$ in water present some trend differences, particularly in the higher energy region (190–205 nm) where the spectrum at 15°C is characterized by a minimum, while that at 85°C by a maximum. From the isodichroic point (210 nm) on, the two spectra start showing a similar pattern defined by a minimum around 218 nm followed by an increasing behavior. Looking at the lower right panel of Figure 8, we can point out that a similar behavior has been qualitatively reproduced by our calculations. In addition, the first calculated isodichroic point has been found around 206 nm (slightly blueshifted with respect to the experiment) and the spectra present the minimum (85°) or at least an inflection (15°) at lower energy values (211–227 nm). A second isodichroic point is found in the calculation but is not present in the experiment.

However, as stated above, some discrepancies in terms of energy and intensity distributions are noticed, affecting the accuracy of the results, which can be rationalized considering the predominant random coil structure at low T values and of the ordered β-turn conformation at high T values. Furthermore, it is interesting to notice the presence of a small population of ordered structures also at 15°C, which defines the second less marked minimum point. Therefore, the temperature influence determines the conversion from an unordered conformation toward a folded structure characterized by intramolecular H bonds. Employing our theoretical approach, we have been able to extract the same two conformations (see Figure 4 of the Section 3.1) from the corresponding conformational spaces, thus corroborating the reliability of the method.

3.4 | Analysis of the solvent effect on the optical response

The effect of the solvent on the spectral features of a solvated chromophore is twofold. On the one hand, we can first consider an *indirect* effect consisting in the solvent-driven conformational transitions of the solute; on the other hand, if in the presence of a very polar solvent and a degree of local interactions with the chromophore (H-bond), the solvent might produce a *direct* effect both produced by the bulk effects and by the electronic coupling between the solvent
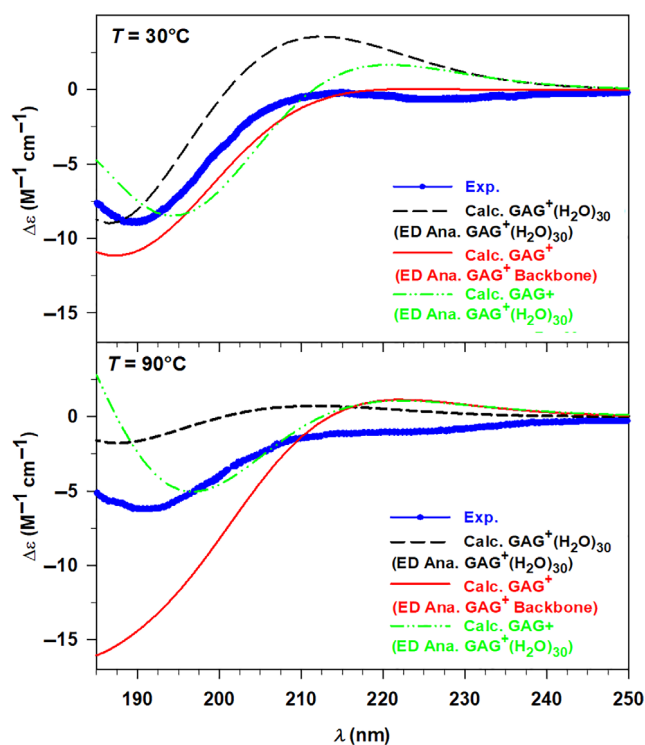
FIGURE 9 Experimental (Exp., blue dots) and calculated (Calc., black dashed line) UV-circular dichroism spectra of GAG$^+$ in water at 30°C (upper panel) and 90°C (lower panel). The two spectra have been compared with those calculated considering only the peptide structures extracted from the ED analysis of the backbone (red solid line) and only the final peptide structures extracted from the ED analysis of the clusters (green dotted line). All the calculated spectra have been convoluted by using Gaussian functions with HWHM of 0.4 eV.

FIGURE 10 Experimental (Exp., blue dots) and calculated (Calc., black dashed line) UV-circular dichroism spectra of (GVGVP)$_2$ in water at 15°C (upper panel) and 85°C (lower panel). The two spectra have been compared with those calculated considering only the final peptide structures extracted from the ED analysis of the clusters (green dotted line). All the calculated spectra have been convoluted by using Gaussian functions with HWHM of 0.3 eV.

molecules directly linked to the chromophore and the chromophore itself which might significantly modify the electronic states as well as the chromophore geometry. In this last part of the study, we decided to address the *direct* solvent effect more explicitly on the optical response of the investigated species.

For this purpose, we have compared the UV-CD spectra of the GAG$^+$ peptide structures extracted from free-MD, reported in Table 1, with their own relative weights with the same structures in the presence of explicit water molecules, reported in Table 2. Obviously, the conformations of the peptides, kept frozen along the second MD simulations, are similar to the structures of the second analysis. However, the geometries employed for the quantum-chemical calculations are not identical. Indeed, the constrained optimizations carried out with or without the solvent bring to slightly different relaxations of the peptide structures, which can be pointed out by the UV-CD spectroscopy, very sensitive to the conformational details. Therefore, such a double analysis allows us to understand how the solvent influences the conformations of GAG$^+$ and the relevance of its inclusion for the calculation of the spectra. All these results have been collected in the following Figure 9.

The results shown in Figure 9 reveal a clear effect of the water at different levels. First, at both the T values of interest, a worse description of the electronic response arises including only the peptide conformers extracted from the first ED analysis (i.e., ED analysis of the GAG$^+$ backbone). The poorer agreement with the experimental spectrum becomes more evident at 90°C, where the minimum around 190 nm displays a much higher CD. This result suggests that the extraction of conformations based only on the free energy analysis of the solute is insufficient for a proper description of the optical response. Hence, the solvent must be included at least in the relaxation of the peptide structure. This hypothesis has been confirmed by considering the conformations extracted and relaxed within the water clusters but without explicit solvent in the calculation of the CD spectra (green lines in Figure 9). Such spectra give already a good agreement with the experimental ones, in particular at 90°C, but the inclusion of the solvent reduces the differences between the experimental and calculated energy distributions. We have applied the same procedure to the second system, the (GVGVP)$_2$, calculating the spectra for the peptide structures previously obtained within the ED analysis of the clusters and optimized in presence of the 40 water molecules. The new results have been compared with the experimental spectra, as well as the overall calculated spectra, and shown in Figure 10.

In this second case, the influence of the solvent on the optical response is more evident than in the previous system and can be justified with the impact of the water molecules on the conformations. Indeed, the weighted UV-CD spectrum calculated for the peptides at 15°C shows an opposite trend with respect to that obtained for the cluster spectrum, and for the experimental one, where the first minimum is the most pronounced. Significant differences can be found for the calculated spectrum at 85°C as well, where the minimum is strongly blueshifted and no qualitative agreement with the experimental result can be pointed out. To conclude our analysis, the small GAG^+ system immediately reveals the effect of the water solvent, not only affecting the solute geometry but also providing a non-negligible electronic effect on the calculated observable, and thus highlights the need of its indirect inclusion for a suitable calculation of the UV-CD spectra. Instead, the direct influence of the solvent on the optical response seems to be more evident when larger systems are considered, for instance the short elastin-like peptide.

4 | CONCLUSIONS

A computational protocol aimed at modeling the UV-CD spectra of solvated species and consisting of quantum-chemical calculations on a series of conformations of a flexible chromophore or on a series of chromophore/solvent clusters extracted from MD simulations, is presented in this study and applied to the aqueous cationic tripeptide GAG^+ and to the aqueous neutral decapeptide $(GVGVP)_2$. The main aim of this study, beyond the description of the method, is to show its physically coherence that is expressed in: (i) its ability in taking into account the conformational repertoire of a solute with reference to the physical conditions of the experiment (temperature, ionic strength, type and density of solvent, etc.); (ii) its ability of directly measuring the actual statistical weight of each conformational state; (iii) its ability in providing us with a reliable quantum-chemical method able of reproducing the observable of interest. The specific feature of the proposed method is its almost zero dependence on the arbitrariness to which it is often necessary to refer for the selection of the different conformational states of the chromophore or involved chromophore/solvent clusters. In the specific case presented in this study, the method satisfactorily reproduces the GAG^+ and $(GVGVP)_2$ UV-CD spectra and, most importantly, their sensitivity to temperature variations. On the other hand the proposed method might intrinsically suffer from: (i) limitations raising from the semi-classical nature of the utilized force field which, moreover, might be sometimes not available (and hence its construction might be necessary) particularly when dealing with species different from the peptides; (ii) a certain degree of arbitrariness in the choice of the number of solvent molecules defining the chromophore(cluster)$_N$ species; (iii) the *possibly flat* spectrum of the CM eigenvalues which would make it very difficult the preliminary analysis on the chromophore conformational repertoire; (iv) the *certainly flat* spectrum of the CM eigenvalues for the chromophore(cluster)$_N$ species which—as already remarked—might produce an incomplete conformational repertoire. However, our

method definitely demonstrates the importance, at least for the systems addressed in the present case, of a physical coherent inclusion of the solvation shells closer to the peptides for the final CD modeling.

DATA AVAILABILITY STATEMENT
Data available in article Supporting Information.

ORCID
Marta Monti https://orcid.org/0000-0003-3191-0122
Mauro Stener https://orcid.org/0000-0003-3700-7903
Massimiliano Aschi https://orcid.org/0000-0003-2959-0158

REFERENCES
[1] N. Berova, L. Di Bari, G. Pescitelli, *Chem. Soc. Rev.* **2007**, *36*, 914.
[2] J. Autschbach, T. Ziegler, *J. Chem. Phys.* **2002**, *116*, 891.
[3] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
[4] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *Phys. Chem.* **1994**, *98*, 11623.
[5] Y. Takeshi, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51.
[6] G. Pescitelli, L. Di Bari, N. Berova, *Chem. Soc. Rev.* **2011**, *40*, 4603.
[7] G. Pescitelli, T. Bruhn, *Chirality* **2016**, *28*, 466.
[8] M. de Wergifosse, J. Seibert, S. Grimme, *J. Chem. Phys.* **2020**, *153*, 084116.
[9] F. Bohle, J. Seibert, S. Grimme, *J. Org. Chem.* **2021**, *86*, 15522.
[10] M. Caricato, *J. Chem. Phys.* **2020**, *153*, 151101.
[11] K. Zhang, T. Balduf, M. Caricato, *Chirality* **2021**, *33*, 303.
[12] T. Balduf, M. Caricato, *J. Phys. Chem. A* **2021**, *125*, 4976.
[13] T. Parsons, T. Balduf, J. R. Cheeseman, M. Caricato, *J. Phys. Chem. A* **2022**, *126*, 1861.
[14] N. Niemeyer, M. Caricato, J. Neugebauer, *J. Chem. Phys.* **2022**, *156*, 154114.
[15] T. Shiraogawa, M. Ehara, S. Jurinovich, L. Cupellini, B. Mennucci, *J. Comput. Chem.* **2018**, *39*, 931.
[16] T. Giovannini, F. Egidi, C. Cappelli, *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864.
[17] S. Del Galdo, M. Fusè, V. Barone, *J. Chem. Theory Comput.* **2020**, *16*, 3294.
[18] G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli, V. Barone, *J. Chem. Phys.* **2020**, *153*, 124110.
[19] M. Srebro-Hooper, J. Autschbach, *Annu. Rev. Phys. Chem.* **2017**, *68*, 399.
[20] J. C. Howard, S. Sowndarya, S. V. I. M. Ansari, T. J. Mach, A. Baranowska-Łączkowska, T. D. Crawford, *J. Phys. Chem. A* **2018**, *122*, 5962.
[21] R. D'Cunha, T. D. Crawford, *J. Phys. Chem. A* **2021**, *125*, 3095.
[22] L. Konecny, M. Kadek, S. Komorovsky, K. Ruud, M. Repisky, *J. Chem. Phys.* **2018**, *149*, 204104.
[23] L. Konecny, M. Repisky, K. Ruud, S. Komorovsky, *J. Chem. Phys.* **2019**, *151*, 194112.
[24] C. G. Chen, M. Giustini, A. Scipioni, A. Amadei, M. D'Abramo, *Comp. Theor. Chem.* **2022**, *1209*, 113591.
[25] A. Amadei, B. M. Linssen, H. J. C. Berendsen, *Proteins* **1993**, *17*, 412.

[26] I. Daidone, A. Amadei, *WIREs Comput. Mol. Sci.* **2012**, *2*, 762.

[27] M. D'Alessando, A. Amadei, M. Stener, M. Aschi, *J. Comput. Chem.* **2015**, *36*, 399.

[28] A. Kumar, R. Schweitzer-Stenner, B. M. Wong, *Chem. Commun.* **2019**, *55*, 5701.

[29] S. E. Toal, D. J. Verbaro, R. Schweitzer-Stenner, *J. Phys. Chem. B* **2014**, *118*, 1309.

[30] A. Kumar, S. E. Toal, D. DiGuiseppi, R. Schweitzer-Stenner, B. M. Wong, *J. Phys. Chem. B* **2020**, *124*, 2579.

[31] H. Nuhn, H.-A. Klok, *Biomacromolecules* **2008**, *9*, 2755.

[32] H. J. C. Berendsen, D. van der Spoel, R. van Drumen, *Phys. Commun.* **1995**, *91*, 43.

[33] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657.

[34] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225.

[35] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, in *Intermolecular Forces* (Ed: B. Pullman), D. Reidel Publishing Company, Dordrecht, the Netherlands **1981**, p. 331.

[36] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.

[37] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463.

[38] T. A. Darden, D. M. York, L. G. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089.

[39] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, UK **1989**.

[40] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931.

[41] N. V. Ilawe, R. Schweitzer-Stenner, D. DiGuiseppi, B. M. Wong, *Phys. Chem. Chem. Phys.* **2018**, *20*, 18158.

[42] T. Bruhn, A. Schaumlöffel, Y. Hemberger, G. Bringmann, *Chirality* **2013**, *25*, 243.

[43] J. A. Schellman, *Chem. Rev.* **1975**, *75*, 323.

[44] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51.

[45] A. Klamt, G. Schüürmann, *J. Chem. Soc. Perkin Trans. 2* **1993**, *5*, 799.

[46] M. Aschi, M. D'Abramo, A. Amadei, *Theor. Chem. Acc.* **2016**, *135*, 132.

SUPPORTING INFORMATION

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

Supporting Information for:

A computational approach for modeling electronic circular dichroism of solvated cromophores

M. Monti[1], M. Stener[1], M. Aschi[2]

[1] Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri, 1, 34127, Trieste, Italy.

[2] Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, Via Vetoio, 67100, L'Aquila, Italy.
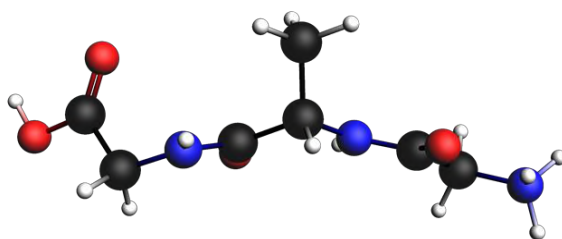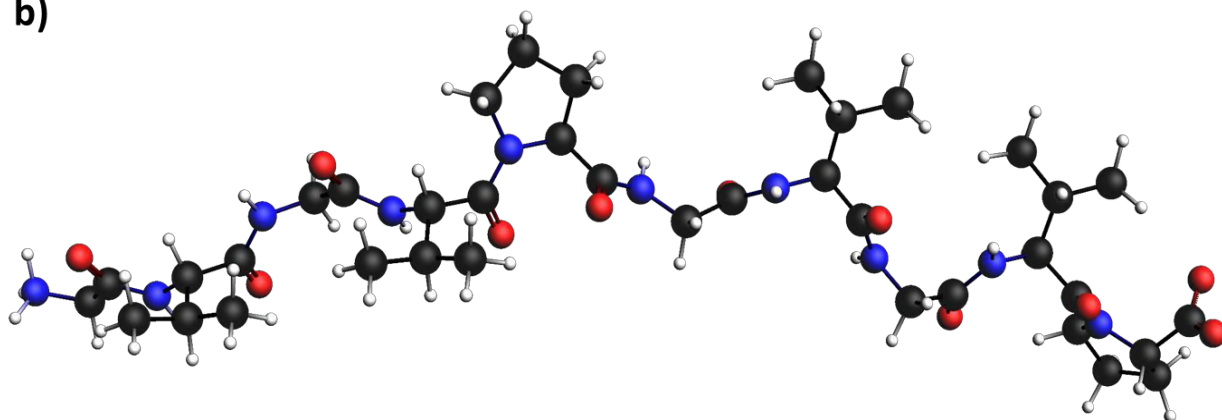
Figure S1. Stick and ball model of a) the cationic GAG tripeptide and b) the GVGVPGVGVP decapeptide. O, C, N, and H atoms are reported as red, black, blue, and white, respectively.

Additional analysis of the conformational spaces

We conducted a further analysis of the conformational spaces of the GAG^+ backbone at the two selected T values, consisting in the projection of both the trajectories along the plane of the first two essential modes at the lower temperature (i.e., 30°C). The same analysis was then repeated for the $(GVGVP)_2$ and all the results are collected in Figure S2.
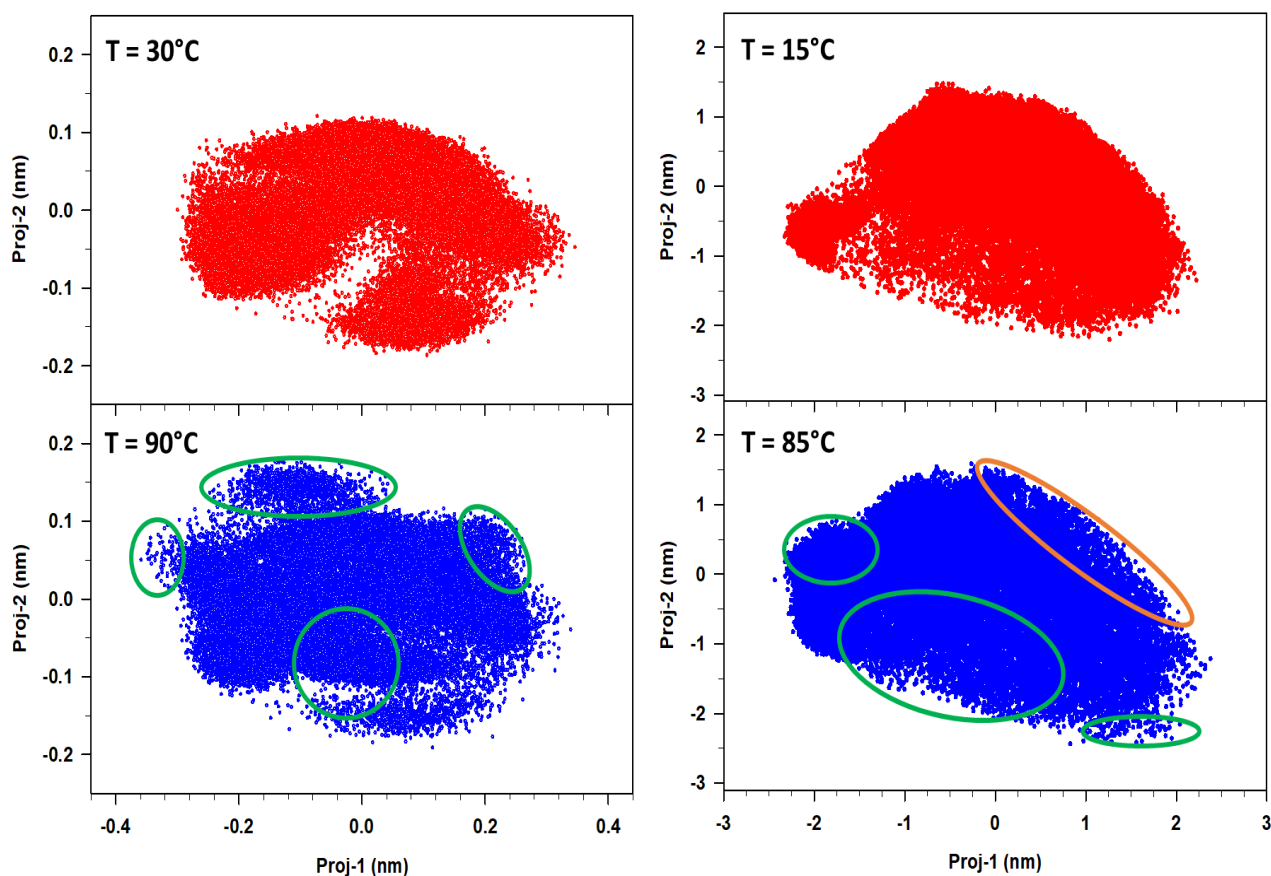
Figure S2. Conformational spaces of GAG^+ (left panels) at 30°C (upper panel) and 90°C (lower panel), and $(GVGVP)_2$ (right panels) at 15°C (upper panel) and 85°C (lower panel). The regions populated at high T values are circled in green, while the regions partially depleted ad high T values are circled in orange.

Figure S2 reveals that the populations which define the two conformational spaces are very similar, with only small regions that become filled increasing the temperature (green regions in the left panel of Figure S2), as expected by the dependence among temperature and accessible motions discussed in the main article. This result is naturally correlated to the small dimensions of the tripeptide which hence can be considered as a good starting point for the validation of the model but limited by its small size. Therefore, we have repeated such analysis of the conformational space for $(GVGVP)_2$, following the procedure explained above for the GAG^+ case. The comparison between the conformational spaces herein shows the increased population of states at higher T values, as well as a slight depletion of the upper right region (green and orange regions, respectively, in the

low right panel of Figure S2), thus suggesting larger conformational differences induced by the temperature.

Individual calculated ECD spectra of the (GVGVP)$_2$(H$_2$O)$_{40}$ clusters at 15°C and 85°C
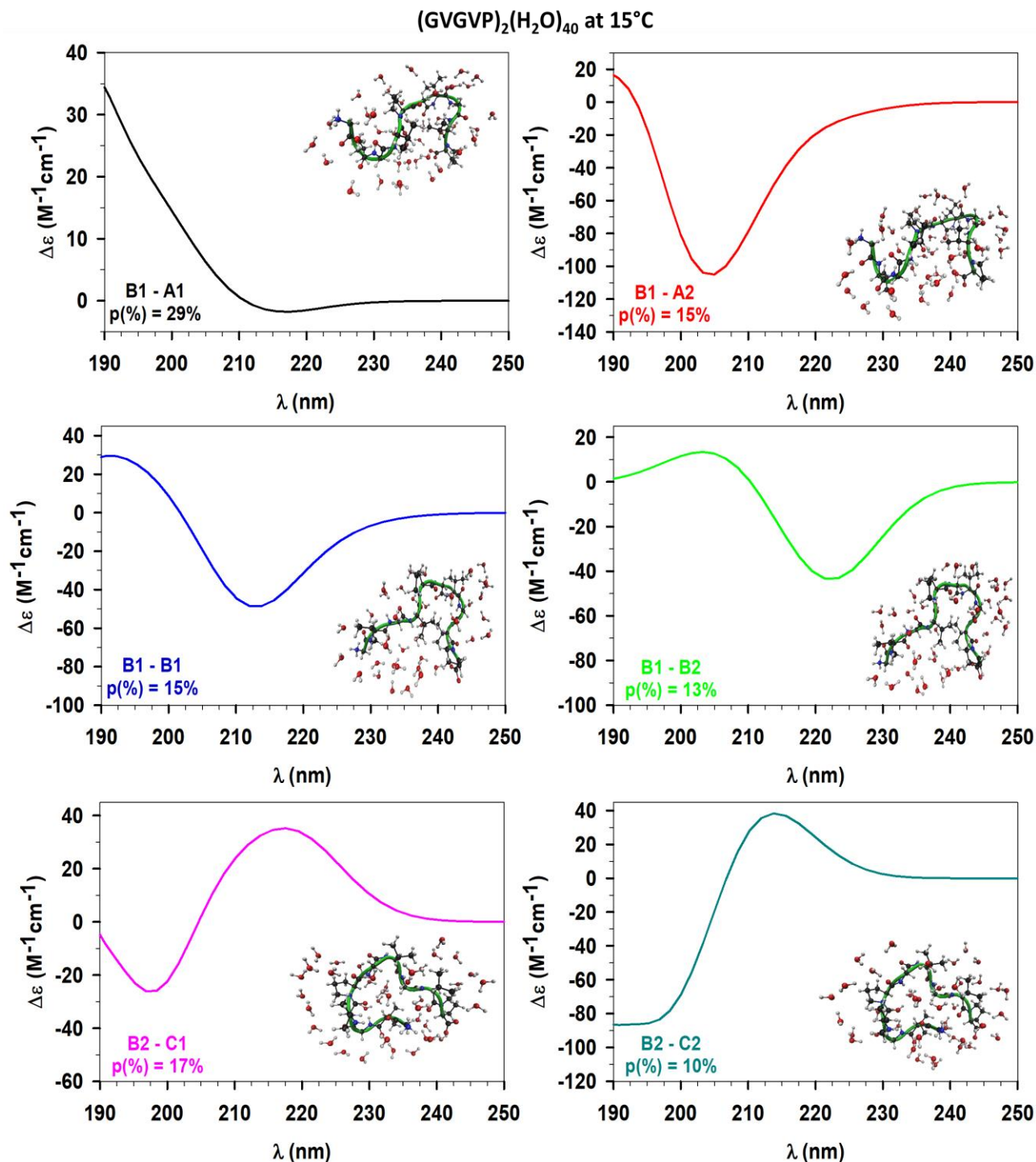
Figure S3. Individual calculated ECD spectra of the (GVGVP)$_2$(H$_2$O)$_{40}$ clusters at 15°C. The labels of the structures are reported in the Figure together with the corresponding statistical weights and the representations of the clusters.
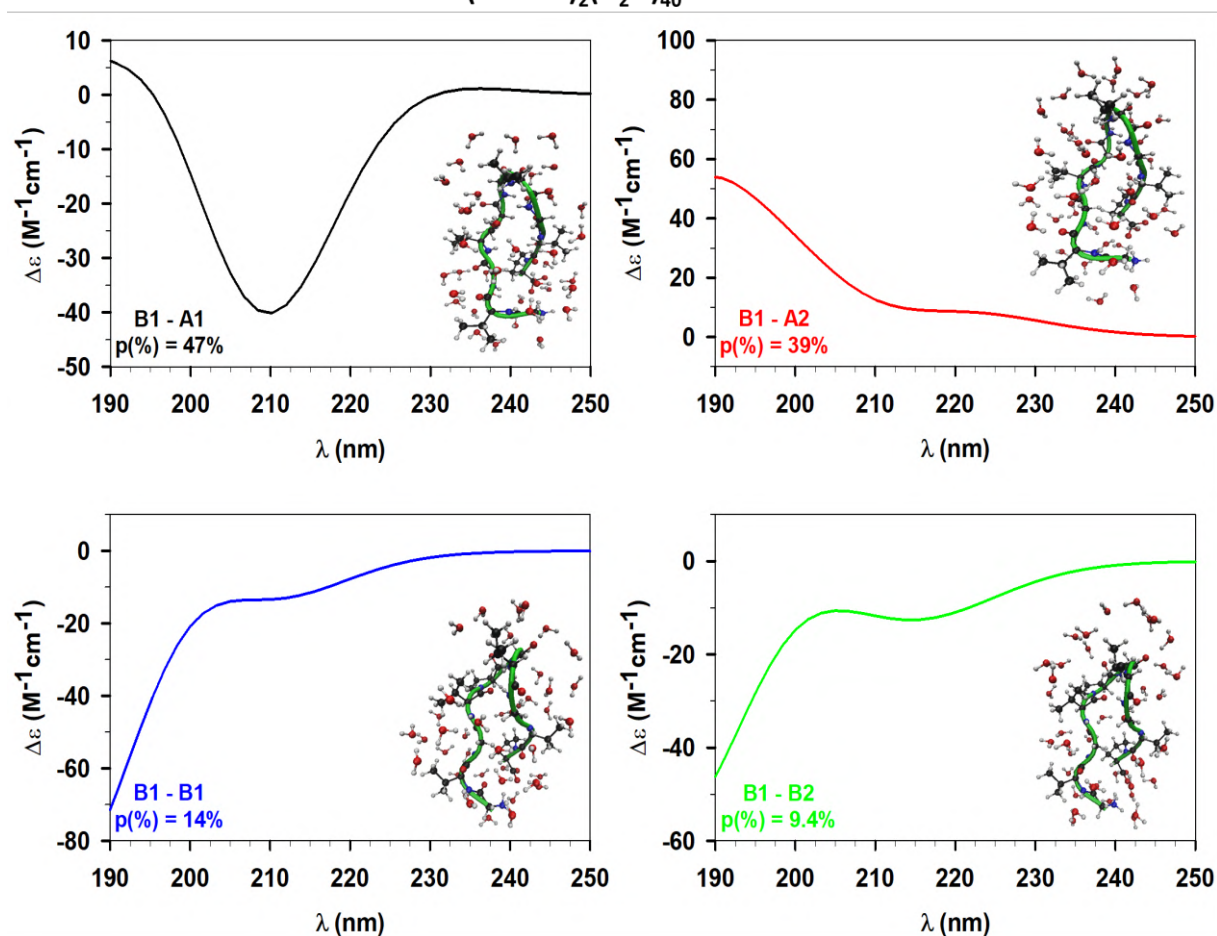
Figure S4. Individual calculated ECD spectra of the (GVGVP)$_2$(H$_2$O)$_{40}$ clusters at 85°C. The labels of the structures are reported in the Figure together with the corresponding statistical weights and the representations of the clusters.

Individual calculated (fitted and discrete) ECD spectra of GAG$^+$(H$_2$O)$_{30}$ and (GVGVP)$_2$(H$_2$O)$_{40}$ at 15°C, 30°C, 85°C and 90°C

GAG$^+$(H$_2$O)$_{30}$ at T = 30°C
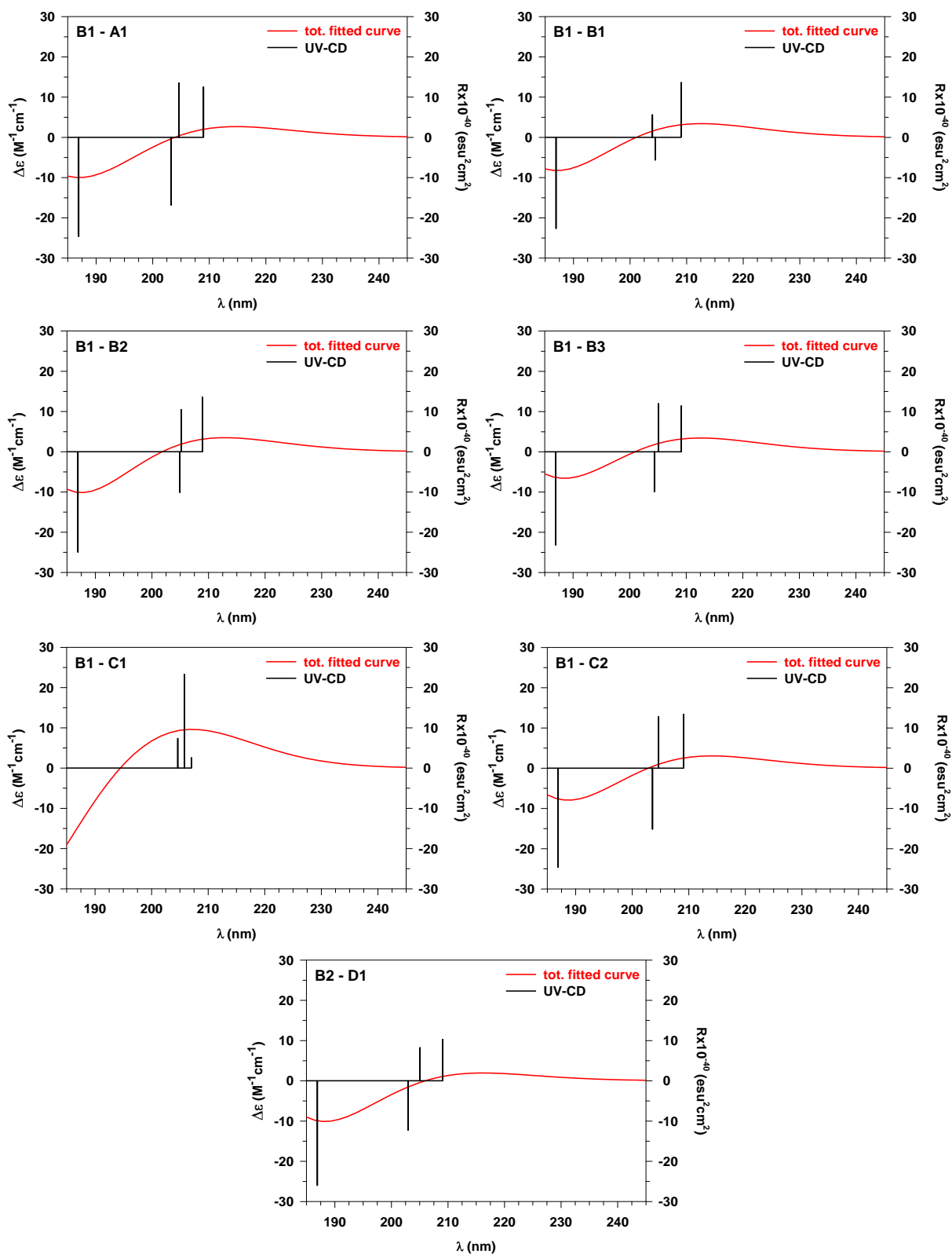
Figure S5. Individual total fitted and discrete ECD spectra of the GAG$^+$(H$_2$O)$_{30}$ clusters at 30°C. The spectra reported herein are not statistically weighted. The labels of the structures are reported in the Figure and consistent with those proposed in Table 2 of the Main Article.

GAG⁺(H₂O)₃₀ at T = 90°C
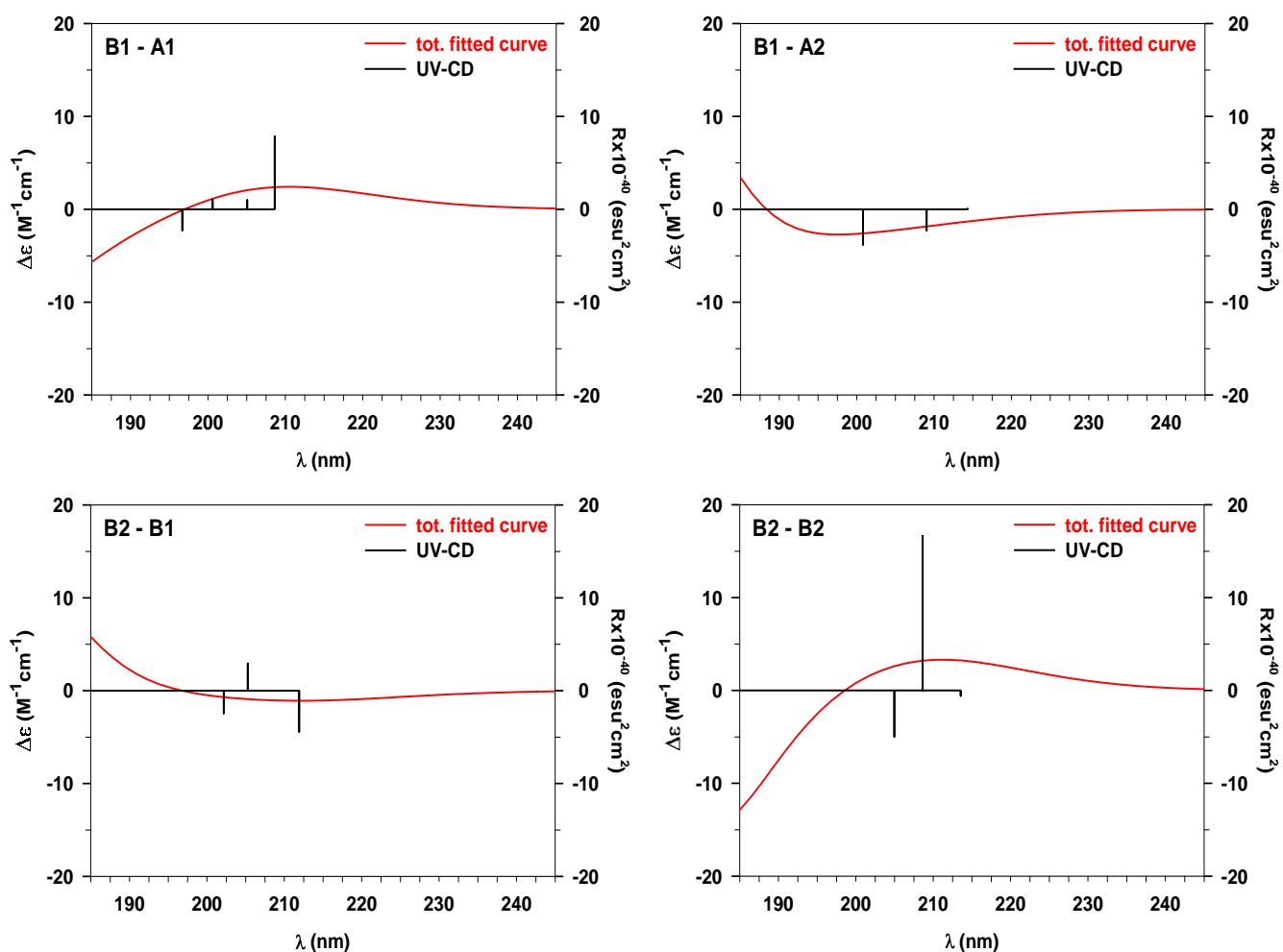
Figure S6. Individual total fitted and discrete ECD spectra of the GAG⁺(H₂O)₃₀ clusters at 90°C. The spectra reported herein are not statistically weighted. The labels of the structures are reported in the Figure and consistent with those proposed in Table 2 of the Main Article.

(GVGVP)$_2$(H$_2$O)$_{40}$ at T = 15°C
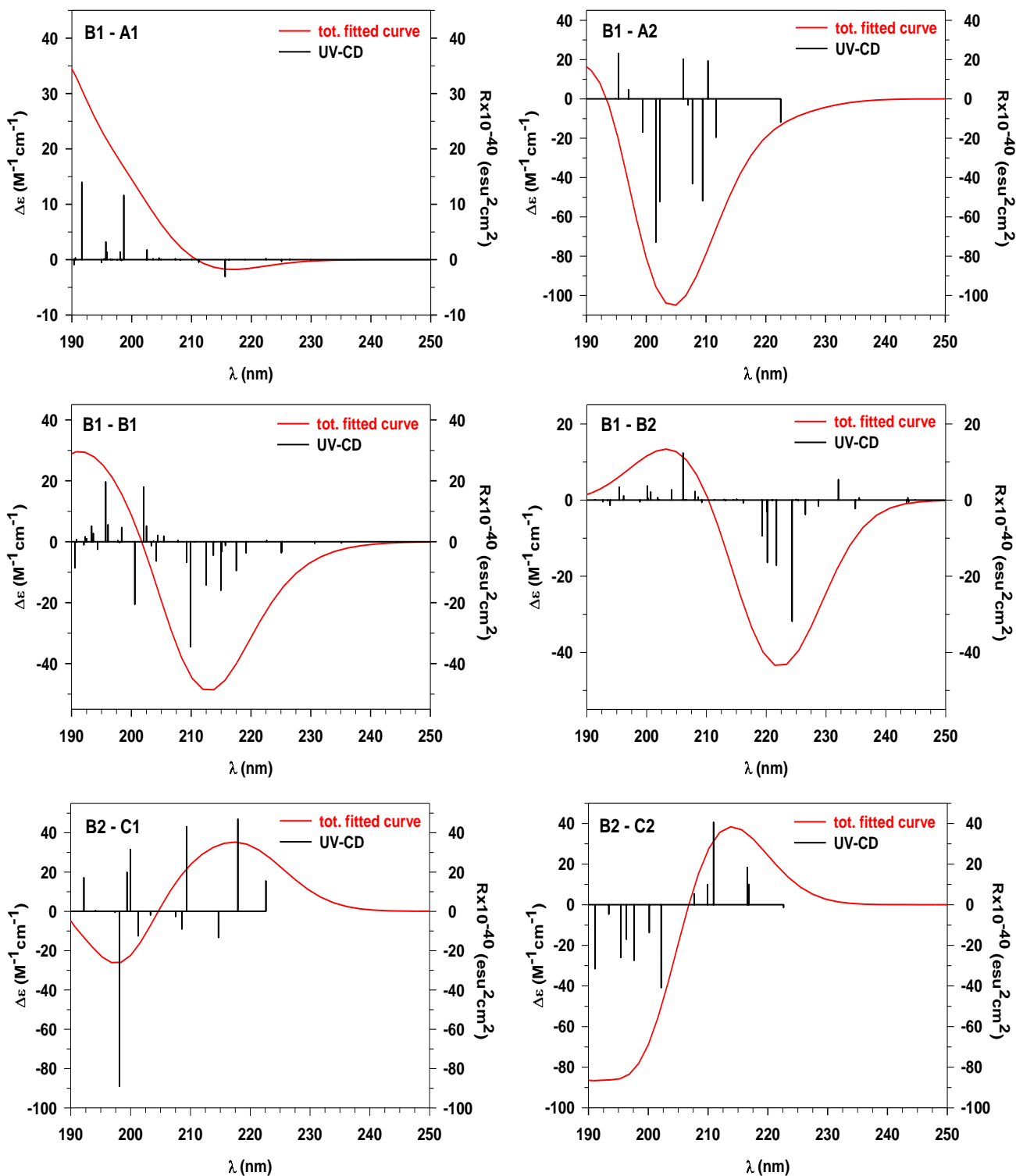
Figure S7. Individual total fitted and discrete ECD spectra of the (GVGVP)$_2$(H$_2$O)$_{40}$ clusters at 15°C. The spectra reported herein are not statistically weighted. The labels of the structures are reported in the Figure and consistent with those proposed in Table 2 of the Main Article.

(GVGVP)$_2$(H$_2$O)$_{40}$ at T = 85°C
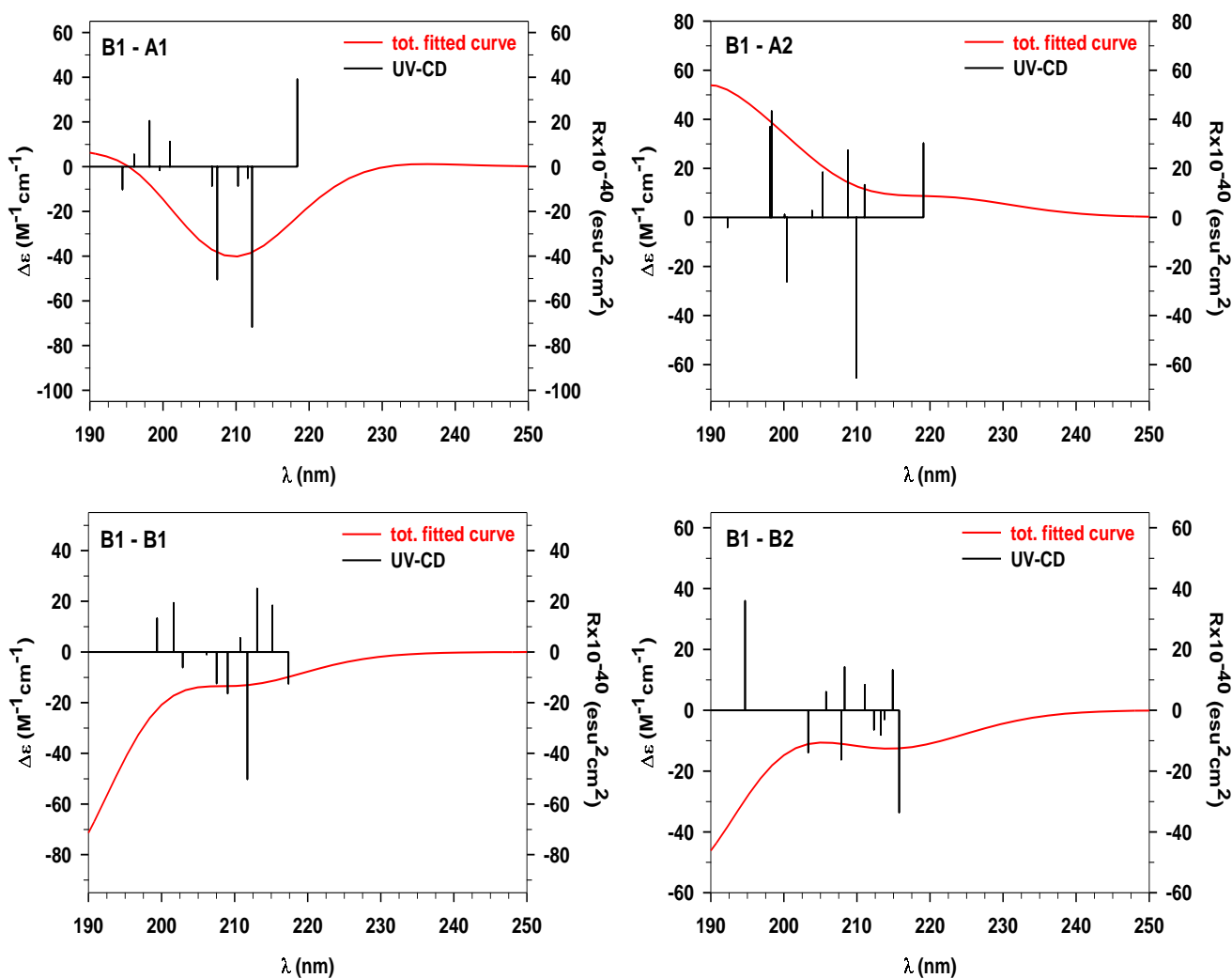
Figure S8. Individual total fitted and discrete ECD spectra of the (GVGVP)$_2$(H$_2$O)$_{40}$ clusters at 85°C. The spectra reported herein are not statistically weighted. The labels of the structures are reported in the Figure and consistent with those proposed in Table 2 of the Main Article.

Effect of the xc functional on the ECD spectra

The choice of the xc functional (wB97X-D) was based on its ability to describe properly intra- and inter-molecular H bonds, as well as charge-transfer excitations[1], which are very relevant to the physics of the systems selected in this work. However, the quality of this decision has been briefly tested employing a second functional, the hybrid CAM-B3LYP, for comparison as explained in the Computational Details section.

The calculations with the CAM-B3LYP were realised only on the most probable conformations at low temperature values (i.e., $GAG^+(H_2O)_{30}$ at 30°C, and $(GVGVP)_2(H_2O)_{40}$ at 15°C) and compared with the ECD spectra obtained for the same structures with the wB97X-D. All the calculated spectra are reported in the following Figure S9 with the experimental ECD as well.
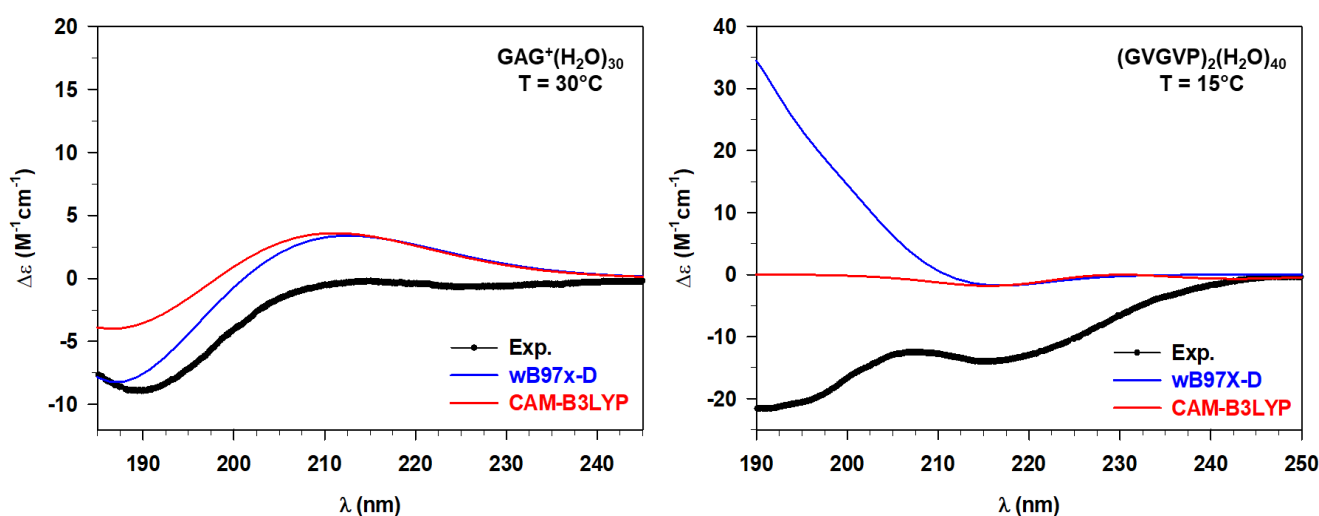
Figure S9. Experimental (Exp.) UV-CD and calculated UV-CD with wB97X-D and CAM-B3LYP xc functional of the most probable conformation of the solvated GAG^+ at 30°C (left panel), and solvated $(GVGVP)_2$ at 15°C (right panel). The spectra reported herein are not statistically weighted.

The left panel of Figure S9 reveals that both the calculated ECD spectra are in good quality agreement with respect to the experimental one, therefore a hybrid functional could be a valid choice as well. However, small improvements of the calculated spectrum can be noticed considering the wB97X-D functional instead of the CAM-B3LYP in terms of intensity scale and position of the minimum peak. The situation becomes more complex for the short elastin-like peptide since both the calculated ECD behaviours differ from the experimental one, as shown in the right panel of Figure S9. Because of the higher complexity of this system, it is clear that a single conformation is not able to reproduce properly all the experimental features (see also Figure S3), but a statistical approach is needed. From a very qualitative perspective, we can just notice that the shape of the experimental minimum around 218 nm is vaguely reproduced by the calculation with the wB97X-D, even though the intensity scale is different from the experimental one, while the dichroism is extremely reduced with the CAM-B3LYP, obtaining an almost null response. However, this is just a preliminary analysis and all the extracted conformers should be treated with both (or more) xc functionals before selecting the one which gives the best agreement with the experiment. Since this would result in an increase of the computational cost, we acknowledge that the calculation with only one xc functional is an approximation and eventual limitation of the approach.

Effect of the solvent mean-field on the ECD spectra

The effect of the solvent on the spectral features has been widely discussed in the article, in particular in Section 3.4 of the Results and Discussion. We focused there on both direct and indirect effects of the water molecules on the treated peptides, such as modification of the electronic states and chromophore geometry. Herein we want to focus on the effect of the mean-field produced by the solvent, thus on the differences obtained using an implicit solvent model instead of treating explicitly the water molecules. A similar discussion can be also found in the literature for the GAG^+ system[2,3].

For this purpose, as briefly explained in the Computational Details section, we calculated the ECD spectra of the bare GAG^+ conformers and of the $GAG^+(H_2O)_{30}$ clusters using the COSMO solvation scheme. The two new resulting statistically weighted ECD spectra have been compared with respect to the experimental one and the calculated spectrum obtained considering only the explicit water molecules. All the data are collected in the following Figure S10.
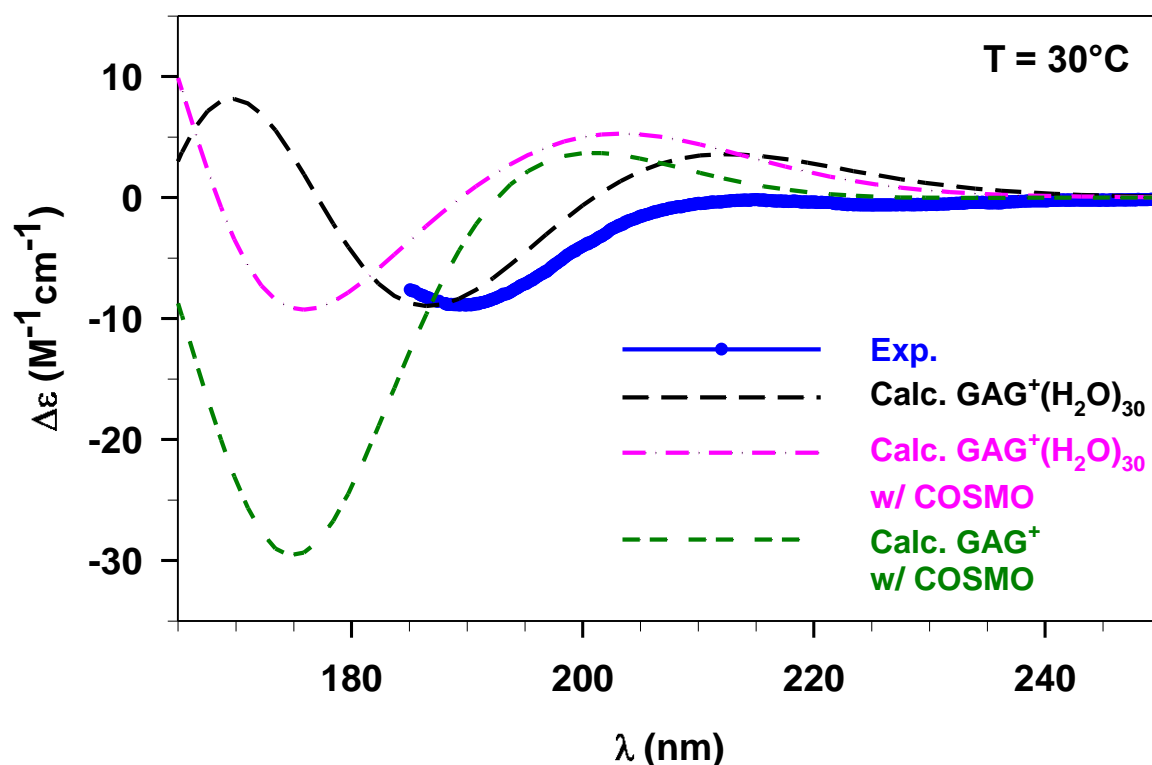
Figure S10. Experimental (Exp.) and statistical calculated (Calc.) ECD spectra of GAG^+ at 30°C. The following conditions were used for the calculated spectra: (i) GAG^+ explicitly solvated with 30 water molecules (black dashed line), (ii) GAG^+ explicitly solvated with 30 water molecules and the COSMO solvation model (pink dashed line), (iii) GAG^+ solvated with the COSMO solvation model (green dashed line).

The results obtained suggest a decrease of the quality agreement between the experimental and calculated ECD when the dielectric model is included. Indeed, in both cases with COSMO reported in Figure S10 (i.e., bare tripeptide, solvated tripeptide) we observe a not negligible blueshift of the main minimum point, as well as large differences in intensity considering only the tripeptide in the dielectric medium. Therefore, a change in the optical response occurs when the systems is surrounded by the dielectric medium, while a good quality agreement is obtained even considering a small number of explicit water molecules (i.e., 30). The deterioration of the results

obtained with only the implicit COSMO approach is not surprising, since in present case the role of solvent hydrogen bonds is crucial for a proper description of the spectrum. On the other hand, also the hybrid implicit COSMO in presence of explicit water molecules does not perform better than the pure explicit solvent treatment. This is probably due to artifacts generated at the boundary between explicit solvent and polarizable continuum, as already pointed out by Brancato et al.[4]

UV- Photoabsorption Spectra

The UV-Photoabsorption spectra have been also considered and compared with those available in the literature[2], first discussing the results for GAG^+ obtained at 30°C in Figure S11.
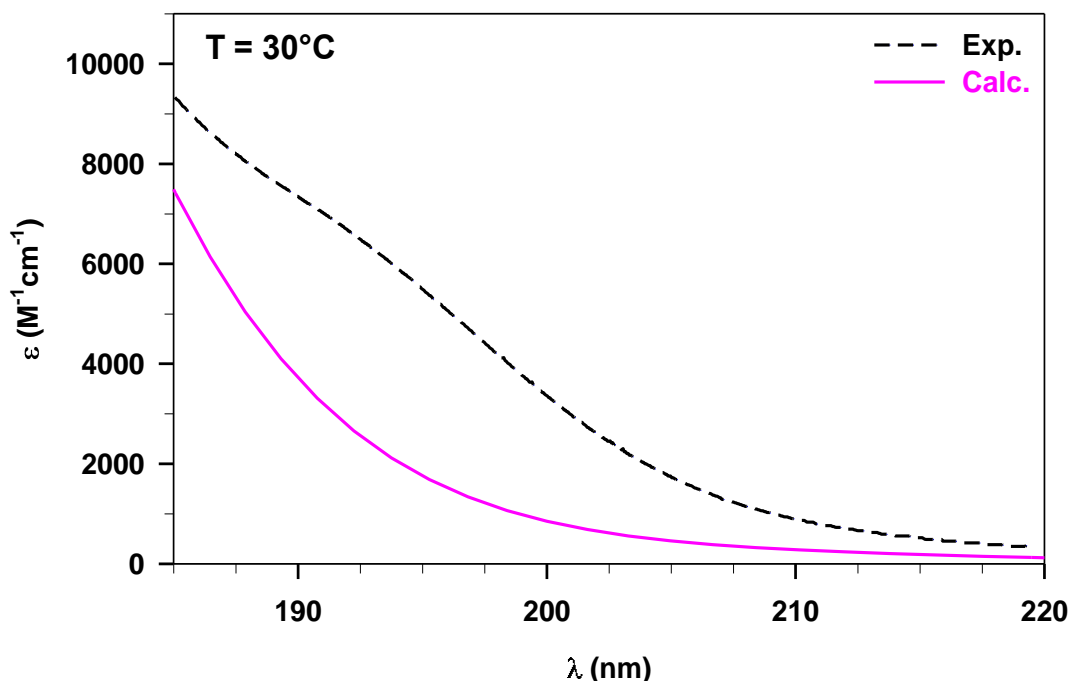
Figure S11. Experimental (Exp. – black dashed line) and calculated (Calc. – pink solid line) UV-Photoabsorption spectra of GAG^+ in water at 30°C. Details of both the convolution and the intensity unit are reported in the Computational Details.

The comparison between the experimental and calculated results points out that a global agreement is found, due to the structureless behaviour of the photoabsorption, although some discrepancies must be taken in account. The intensity difference is particularly significant in the (185-205) nm region, then reduced moving towards lower energy values. Such discrepancy is strongly influenced by the different decreasing trend that can be observed for the two spectra. Indeed, it is possible to notice that the calculated UV-Photoabsorption spectrum shows a much steeper descent behaviour with respect to that experimentally measured. Furthermore, around 193 nm the experimental spectrum displays a weak shoulder which is not present in the calculated profile.

As previously done, it is possible to include within the analysis the variation of the spectral features as an effect of the temperature. According to our knowledge, the UV-Photoabsorption spectrum of this system has not been measured at 90°C. However, a general correlation among the temperature and the spectral features can be extracted comparing our two calculated UV-Photoabsorption spectra (i.e., at 30°C and 90°C) with those experimentally recorded at several T values (i.e., 10°C, 30°C and 70°C).
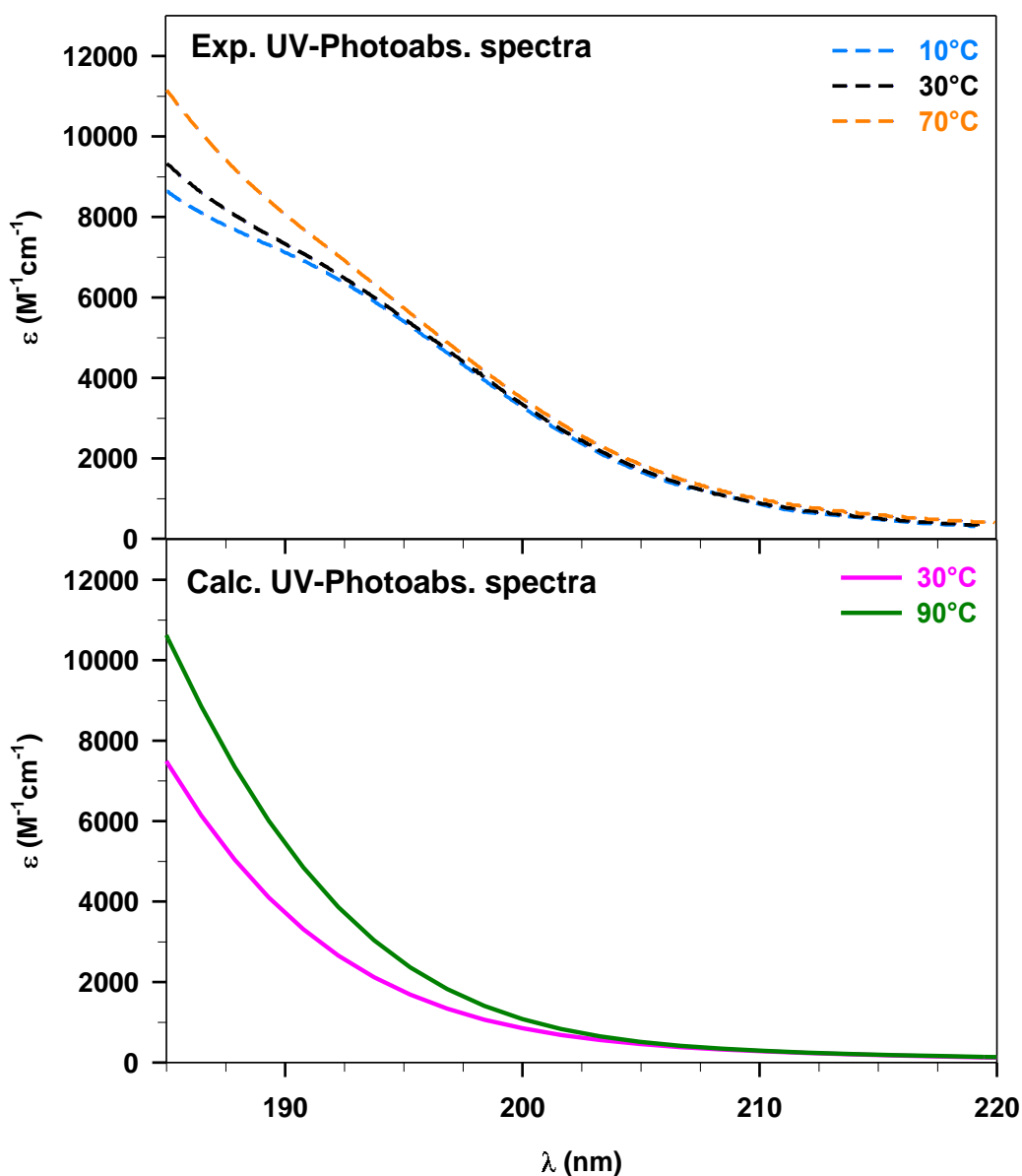
Figure S12. Upper panel: Experimental (Exp.) UV-Photoabsorption (Photoabs.) spectra of GAG$^+$ in water taken at 10°C (blue dashed line), 30°C (black dashed line), and 70°C (orange dashed line), respectively. Lower panel: Calculated (Calc.) UV-Photoabsorption (Photoabs.) spectra of GAG$^+$ in water taken at 30°C (pink solid line) and 90°C (green solid line), respectively. Details about both the convolution and the intensity unit are reported in the Computational Details.

A qualitative analysis of the two calculated spectra (lower panel of Figure S12) highlights no significant changes increasing the temperature, except for a uniform intensity increase. Actually, this difference is particularly marked at higher energy values (around 185 – 190 nm), while it tends to zero towards low energy values. Furthermore, the separation between the different experimental patterns is rapidly reduced from 195 nm onward, bringing to an overlay of the spectra. The same coincidence is achieved also for the calculated results but in a much slighter way (i.e., from 205 nm onward). Beside the limited discrepancies between the experimental and theoretical results, the latter ones describe the spectral changes occurring after the temperature increase.

N.B: In the original version of the SI, we reported all the coordinates of the different conformers that were selected and employed to calculate the ECD spectra.
They are available at: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.27001

References

[1] N. V. Ilawe, R. Schweitzer-Stenner, D. DiGuiseppi, B. M. Wong, *Phys. Chem. Chem. Phys.* **2018**, *20*, 18158.

[2] A. Kumar, S. E. Toal, D. DiGuiseppi, R. Schweitzer-Stenner, B. M. Wong, *J. Phys. Chem. B* **2020**, *124*, 2579.

[3] N. V. Ilawe, A. E. Raeber, R. Schweitzer-Stenner, S. E. Toal, B. M. Wong, *Phys. Chem. Chem. Phys.* **2015**, *17*, 24917.

[4] G. Brancato, N. Rega and V. Barone, *J. Chem. Phys.* **2008**, *128*, 144501.

4.1.2 The Conformational Dynamics of the Ligands Determines the Electronic Circular Dichroism of the Chiral Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ Cluster

This publication represents the first extension of the proposed computational approach to a much more complex case, that is, thiolate-protected gold nanoclusters. These nanosystems, which are very appealing for their properties and applications, show significant larger sizes and more flexibility compared to the previously treated biomolecules. In particular, the organic layer that covers the metallic structure is composed of several ligands which assume different conformations. Furthermore, their size makes the standard Casida's scheme unsuitable for the calculation of the optical properties. In order to study these systems, it is quite common to adopt a model reducing, for instance, the -SR$_n$ chain to just -SCH$_3$ or to use the experimental X-ray structure (when available) as representative of the whole conformational ensemble.

Therefore, this work aimed to show a possibility to go beyond the current treatments by combining MD simulations, ED for the extraction of the significant conformers, and ECD calculations by means of the first-principle polTDDFT approach. The procedure was tested on a well-studied case, that is the Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ cluster in toluene. Here, the MD simulations were performed with the GolP-OPLS-AA force field which contains the parameters to describe these systems. The access to this FF was granted by Prof. Stefano Corni and Dr. Giorgia Brancolini, co-authors of this paper. Performing the ED analysis on this system, a set of 12 final conformers was extracted and used for the ECD calculations (polTDDFT, LB94 xc functional). The final statistically averaged ECD was then compared with the spectra calculated on the experimental X-ray structure, adopting two different xc functionals (LB94, B3LYP). Such comparison pointed out the sensibility of ECD towards different conformations, showing that while the statistical spectrum properly captures the optical features, calculations on a single structure do not. The main discrepancy between the experimental and statistical calculated ECD is observed in the energy region where the gold-associated transitions arise, but it may be ascribed to the LB94 functional. Further improvements could be achieved by using polTDDFT with more accurate functionals, although the computational cost can become unaffordable if several conformers of this size need to be treated. Affordable schemes that combine hybrid functionals with polTDDFT are currently under development in our group.

Letter

The Conformational Dynamics of the Ligands Determines the Electronic Circular Dichroism of the Chiral Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ Cluster

M. Monti, G. Brancolini, E. Coccia, D. Toffoli, A. Fortunelli,* S. Corni, M. Aschi,* and M. Stener*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Effects of the conformational dynamics of 2-PET protective ligands on the electronic circular dichroism (ECD) of the chiral Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ cluster are investigated. We adopt a computational protocol in which ECD spectra are calculated via the first principle polTDDFT approach on a series of conformations extracted from MD simulations by using Essential Dynamics (ED) analysis, and then properly weighted to predict the final spectrum. We find that the experimental spectral features are well reproduced, whereas significant discrepancies arise when the spectrum is calculated using the experimental X-ray structure. This result unambiguously demonstrates the need to account for the conformational effects in the ECD modeling of chiral nanoclusters. The present procedure proved to be able of capturing the essential conformational features of the dynamic Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ system, opening the possibility to model the ECD of soluble chiral nanoclusters in a realistic way.

Chiral thiolate-protected gold nanoclusters (RS-AuNCs) have attracted much attention in recent years because of their several potential applications in different fields, such as chiral catalysis,[1,2] sensing and recognition,[3,4] as well as chiral separation for biological molecules.[5] In general, the correlation between the cluster structure and its optical properties is a fundamental issue to allow rational design for potential applications.[6] Of particular interest in these systems is the electronic circular dichroism (ECD) response beyond the spectral region due to the ligands themselves that generally lies in the UV range, i.e., the metal-based region of optical absorption.[7] An important question concerns the origin of the chiroptical properties of RS-AuNCs, and several mechanisms have been debated over the years to explain and rationalize it. In detail, either the metal core has an asymmetric arrangement which determines an intrinsic chirality,[8] or the chiral response can be induced by the presence of the protective ligands. In the latter case, we can consider: (i) achiral ligands that form chiral adsorption patterns (footprint model)[9] and/or (ii) chiral ligands that induce the chirality in the metal core by trapping its electrons in a dissymmetric electric field.[10]

For instance, the chirality of the Au$_{38}$(SC$_2$H$_4$Ph)$_{24}$ cluster (SC$_2$H$_4$Ph abbreviated as 2-PET from now on), which is the objective of this study, is induced by the passivation of achiral 2-PET ligands on the surface of the metallic core. This nanocluster has attracted research interest not only on fundamental studies of chirality,[11] but also on stability,[12] doping,[13] and catalysis,[14] especially after 2010 when its total crystal structure was resolved by Qian et al.[15] Moreover, the

first characterization of the 8 kDa cluster by Schaaff et al.[16] in 1997, well before the determination of its atomistic structure, inspired many theoretical works to examine structural and electronic properties of the Au$_{38}$(SR)$_{24}$ nanocluster.[17−19] After a decade of density-functional theory (DFT) predictions, both experiment[15] and theory[20,21] confirmed that the Au$_{38}$(2-PET)$_{24}$ shape is prolate with a face-fused bi-icosahedral Au$_{23}$ core protected by 3 short Au(SR)$_2$ and 6 long Au$_2$(SR)$_3$ staples. The Au$_{23}$ core, although slightly distorted, can be associated with a D_{3h} symmetry that is lowered to a D_3 one by the Au−S staple motifs, which assume a chiral arrangement. Indeed, the 6 long staples are divided in 2 triblade fans with a staggered configuration that rotate clockwise or anticlockwise depending on the enantiomer considered (see Figure 1). The 3 short staples instead slightly tilt with respect to the 3-fold axis and follow the handedness of the Au$_2$(SR)$_3$ staples.

Despite many studies on the structure and chirality of Au$_{38}$(2-PET)$_{24}$, the measurement of the ECD was revealed to be very challenging because of the inability to separate the two enantiomers in solution. Such limitation was finally overcome in 2012 when Dolamic et al.[22] performed the separation of the

Side view along C₃ axis   **Top view along C₃ axis**
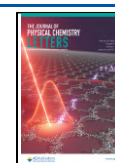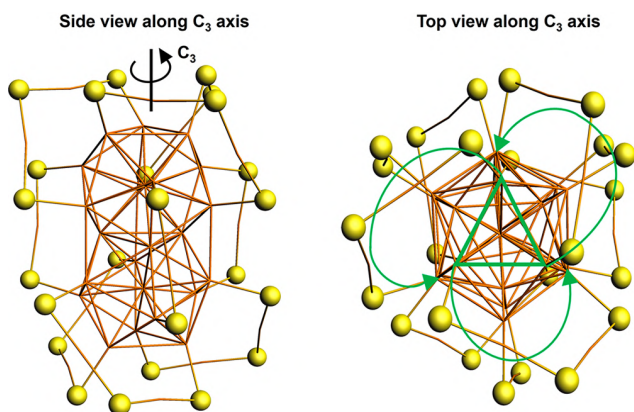
Figure 1. Side (left) and top (right) view along the C_3 axis of the gold−sulfur architecture. For clarity the organic ligands were removed, while the Au and S atoms are reported as sticks in orange and balls in yellow, respectively. The (left-)handedness of the cluster studied in this work is modeled by the green arrows that connect the core Au atoms (green triangle) with the inner Au atoms of the *star-shaped* staples.

racemic mixture employing chiral high-performance liquid chromatography. The ECD of the left-handed enantiomer was proposed again in 2014 by Barrabés and co-workers[23] together with the spectra of $Pd_2Au_{36}(2\text{-PET})_{24}$ clusters to investigate the doping effects on the chiroptical properties. Independently, in the same year Xu et al.[7] proposed an enantioselective synthesis of several $Au_{38}(SR)_{24}$ nanoclusters by using chiral ligands such as 2-phenylpropane-1-thiol. The ECD spectra recorded strongly resemble those measured by Dolamic and co-workers[22] for the $Au_{38}(2\text{-PET})_{24}$ enantiomers.

From a theoretical perspective, the availability of experimental ECD spectra together with the resolution of several X-ray structures represented a huge step forward for the investigation of the RS-AuNCs chirality and the assignment of their absolute configurations. Focusing on Au_{38} NCs, ECD calculations have been already performed for instance by Lopez-Acevedo et al.[21] or by Baseggio and co-workers[24] using time-dependent density functional theory (TDDFT) and different thiolate protective ligands. However, theoretical approaches for modeling the ECD of nanoclusters have so far mostly neglected a fundamental aspect of the physics of this system, i.e., the mobility of the ligands and thus its influence on the chiroptical response. To the best of our knowledge only one previous explorative study has highlighted the role of the different conformations of the ligands on the ECD of the RS-AuNCs.[25] Indeed, the ECD calculation is typically performed on the experimental X-ray structure in a vacuum to overcome the computational costs and difficulties of conformational and response predictions. While the metal and Au−S interface structures are expected to be preserved in solution, significant changes can instead occur in the ligands, so that the neglect of these conformational effects can produce discrepancies between experiment and theory.[24] In recent years, the dynamic effects of RS-AuNCs have been investigated in several works of Bürgi and collaborators[26−28] combining vibrational spectroscopy and theoretical calculations (i.e., DFT and molecular dynamics (MD) simulations). Furthermore, in a recent work of Pyo et al.[29] the interplay between ligand orientations and pK_a of Au_{25} and Au_{102} NCs has been assessed by combining MD simulations and ED analysis.

To solve this issue, herein we propose an affordable computational procedure which includes conformational effects into a proper modeling of the ECD. This provides the possibility of a more realistic calculation of the ECD, not only for this $Au_{38}(SR)_{24}$ cluster, but also of a large variety of soluble chiral gold NCs. The method here adopted is an extension of a protocol recently published by some of these authors[30] where the chiroptical properties of solvated peptides were investigated. The conformers produce an excellent qualitative description of the experimental chiral response[23] of the 2-PET conformational transitions, thus demonstrating the validity of the proposed approach.

MD simulations of the solvated clockwise-$Au_{38}(2\text{-PET})_{24}$ nanocluster were performed with the Gromacs package[31] version 5.1.2. The OPLSA-AA topology was produced with the TPPMKTOP topology generator starting from the experimental X-ray geometry,[15,24] while the Lennard-Jones parameters for Au and S atoms were taken from the GolP-OPLS-AA[32] Force Field (FF), which contains the parameters to describe thiolate-protected gold nanoclusters compatibly with the OPLS-AA FF.[33] The RESP charges and the parametrization procedure have been previously assessed for similar Au_{25} and Au_{144} nanoclusters.[34−36] The OPLS-AA parameters of the toluene solvent were generated using LibParGen.[37] All the MD simulations were performed in the NVT ensemble using the velocity rescaling algorithm[38] to keep the temperature constant. All the bond lengths were constrained by using the LINCS algorithm.[39] The Particle Mesh Ewald method[40] was used with 34 wave vectors in each direction, a cubic interpolation of the fourth order and a 1.0 nm cutoff in order to compute the long-range electrostatic interactions.

The experimental X-ray $Au_{38}(2\text{-PET})_{24}$ structure[15,24] (see Figure S1 of the Supporting Information, SI) was inserted in a cubic box in the presence of toluene at a density resembling the reference experimental conditions,[23] i.e., 303 K and 1.0 bar. For this purpose, since our simulations were carried out in the canonical ensemble (see below), the dimension of the solute−solvent box was then adjusted in order to achieve the average pressure of a box containing only toluene, with the same number of molecules (653), previously simulated in the NVT ensemble at the experimental temperature (303 K) and pure toluene density[41] (857.55 kg/m³). After an initial slow thermal equilibration, we produced a trajectory of 150 ns constraining the gold atoms to preserve the experimental X-ray *staple* (i.e., Au − S) motifs. The importance and effect of the experimental metal architecture on the spectral features have been already discussed elsewhere.[24,42−44] However, for the sake of completeness, we also performed an additional 50 ns all-atoms MD simulation to analyze the contribution of the gold atoms to the principal conformational transitions (essential for the modeling of the final spectrum) and evaluate the quality of Au-constrained MD. The results, discussed more in detail in the Supporting Information (Figures S2 and S3), show that the contribution of Au atoms to the most relevant transitions is almost null. Therefore, the significant conformations of the whole system can be extracted by analyzing only the conformational states of the flexible 2-PET ligands (see next sections).

The MD simulation described above was analyzed in terms of ED[45−47] to sample the conformational landscape of the whole ensemble of the 2-PET ligands present on the nanocluster. The procedure adopted in this work has been

previously explained in detail,[30] and it is here only briefly summarized. The covariance matrix of all the 2-PET moieties atomic coordinates was constructed and diagonalized producing a set of eigenvectors with corresponding eigenvalues representing the mean square fluctuations. The eigenvectors showing the largest eigenvalues represent the directions along which the system undergoes the largest amplitude motions, i.e., the motions associated with the conformational transitions. Consequently, the projection of the Cartesian coordinates along the MD simulation onto such a subspace—i.e., the Principal Components Analysis—allows one to reduce the dimension and the complexity of the conformational space. Moreover, as also shown in the present case, it is possible to consider only the two eigenvectors associated with the two highest eigenvalues, hence allowing a straightforward conformational analysis on a bidimensional (2D) conformational landscape hereafter termed as essential space (ES). In this respect, each region of the ES showing a high density of projected points defines an ith conformational basin, and the number of projected points, divided by the total number of frames, allows one to evaluate the conformational probability $P(i)$. Considering each $P(i)$ and defining one of the above basins as the reference basin with a probability P_{ref} we can calculate the free energy differences between the ith basin and the reference one employing the standard relation:

$$P(i) = P_{ref}e^{-\Delta G^\circ / RT} \tag{1}$$

where ΔG° corresponds to the standard Gibbs free energy difference between the ith and the reference basin on the ES within the following approximations: (i) negligible difference between the partial molar volumes of the $Au_{38}(2\text{-PET})_{24}$ structures falling in the reference and ith basin (note that our simulations are carried out in the NVT ensemble); (ii) negligible differences between the quantum-vibrational molecular partition function of the $Au_{38}(2\text{-PET})_{24}$ structures falling in the reference and ith basin.[48] Therefore, we extracted from all the basins a certain number of $Au_{38}(2\text{-PET})_{24}$ structures— hereafter termed as Representative Conformations (RC)— falling in the (0−2.5) kJ/mol free energy range. Because of the large dimensions of these basins, it was initially necessary to extract more than one RC from each basin. We selected, indeed, a set of 19 structures (see Figures S4 and S5 of the Supporting Information). However, for each basin, the extracted RC were further analyzed in terms of Root Mean Square Deviation (RMSD) to estimate their actual conformational difference. In this respect we assumed as spectroscopically equivalent RC with a RMSD < 1.78 Å following a preliminary analysis based on the outcome of the ECD spectra as reported in the Figure S5 of the Supporting Information. From this latter analysis we then found 10 of the previously extracted RC structures with RMSD values ≥2 Å, whereas the remaining RC were separated in two groups with the RMSD < 1.78 Å. This allowed us to reduce the number of the final RC from 19 to 12 to be considered for the quantum chemical calculations (see below). We calculated the ECD signal on each of the 12 RC, which was then weighted using the normalized probabilities, i.e., eq 1. The sum of all the weighted ECD signals produced the actual spectra reported in the following.

The geometries of the 12 $Au_{38}(2\text{-PET})_{24}$ conformations extracted by the MD simulations were relaxed with specific constraints necessary to maintain the features of the corresponding basins, i.e., the values of the semiclassical coordinates. This was accomplished by relaxing only the ligands stretching and bonding angles and keeping frozen the corresponding proper dihedral angles as well as the metallic atoms coordinates. All the optimizations were realized at the DFT level,[49] employing the standard GGA PBE exchange-correlation (xc) functional[50] combined with the GRIMME-D3 dispersion terms.[51] Furthermore, it was used a basis set of Slater-type orbitals (STO) of triple-ζ plus polarization quality (TZP), and the scalar relativistic effects were treated within the Zero Order Regular Approximation[52] (ZORA).

The ECD spectra of the 12 conformations were calculated with the complex polarizability time-dependent DFT (polTDDFT) algorithm,[53] where the rotatory strength R is defined as

$$R = \frac{3\omega\varepsilon}{2c}Im[\overline{\beta}] \tag{2}$$

In eq 2, $Im[\overline{\beta}]$ represents the average over all the possible orientations of the imaginary part of the rotatory strength tensor, ω is the photon energy, ε corresponds to the imaginary part of the photon energy, and c is the speed of light. In this work, ε was taken as equal to 0.15 eV, allowing a direct comparison with a Lorentzian broadening with the same half width half-maximum. In addition, we used a cutoff of 3 eV above the excitation energy. All the 12 polTDDFT calculations were performed in the gas-phase as a compromise between computational cost and accuracy, moreover considering that solvent effects on the ECD are generally very pronounced for polar molecules in polar solvents.[54−56] Therefore, since the present cluster is a nonpolar molecule as well as the solvent (toluene), we did not consider the solvent effect in the calculation. The LB94[57] xc functional, which includes the correct asymptotic behavior, was employed together with the TZP basis set and the ZORA scheme for the relativistic effects. Each spectrum was multiplied by the corresponding statistical weight (eq 1) and summed up to obtain the final averaged spectrum, then compared with the experimental reference[23] to evaluate the quality of our result.

An additional analysis has been performed calculating the polTDDFT ECD spectrum of the experimental X-ray structure[15,24] with the LB94, and the hybrid B3LYP[58] xc functional which contains a portion of the Hartree−Fock nonlocal exchange. The latter one provides a more accurate description, in particular of the metallic spectral features. Additionally, for the B3LYP calculation we employed the Hybrid Diagonal Approximation[59] (HDA) which allows us to perform polTDDFT calculations with hybrid functionals, although it is still computationally demanding with respect to standard GGA functionals like, for instance, LB94. For this reason, we restricted the HDA-B3LYP calculation to the experimental X-ray diffraction structure only. All the other parameters of the calculation (e.g., basis set, ZORA) were maintained also in these two new calculations.

We report in Figure 2 the spectrum of the covariance matrix eigenvalues (mean square fluctuations) of the 2-PET ($-SC_2H_4Ph$) which may help us in predicting the number of modes needed to describe the collective motions of the $Au_{38}(2\text{-PET})_{24}$ nanocluster, i.e., the size of the ES.

The spectrum of the eigenvalues rapidly decays to 0, with only the first 6 eigenvalues (inset of Figure 2) larger than 0.5 nm^2. However, since the first pair of eigenvalues accounts for the largest fraction of the whole matrix trace (i.e., the whole
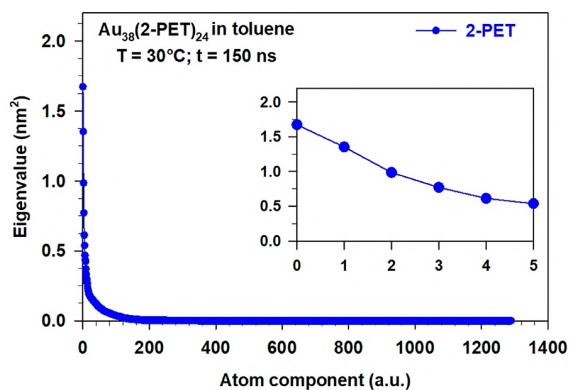
Figure 2. Plot of the eigenvalues resulting from the diagonalization of the 2-PET ($-SC_2H_4Ph$) ligands covariance matrix. The matrix was built with the atomic positions of the thiolate ligands extracted from the MD trajectory of 150 ns at $T = 30$ °C. The first 6 eigenvalues are highlighted in the right inset of the figure.

fluctuation), we can simplify the analysis considering only the corresponding first 2 eigenvectors. In addition, it should be remarked that the inclusion of, at least, the third eigendirection, although possible in principle,[60] might be very expensive in the case of highly complex systems like the present one. Therefore, well aware of the possible incompleteness of the conformational analysis, we decided to adopt a 2D ES for representing the best compromise between accuracy and computational effort. The comparison with experimental data will provide a validation of this approximation.

The conformational landscape, obtained from the procedures described previously, is depicted in Figure 3.
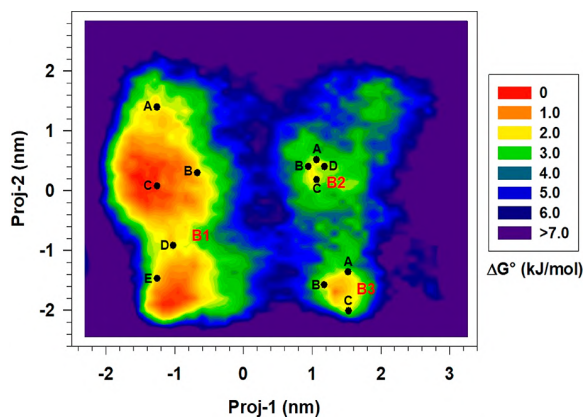
Figure 3. Probability pattern built with the MD trajectory projections (proj-1, and proj-2 in nm) on the first two essential modes. Each square, graphically smoothed in the histogram, represents a conformational basin associated with a specific relative free energy $\Delta G°$ value reported in kJ/mol (right side of the figure). The three low-energy basins are labeled as B1, B2, and B3, respectively, while the positions of the selected 12 low-energy conformations are marked with black dots.

In Figure 3 one observes three low-energy conformational basins (B1, B2, and B3) surrounded by barriers with $\Delta G° \geq 3$ kJ/mol. The three basins differ in terms of dimensions and energy range, indeed while B2 and B3 are quite small and associated with higher $\Delta G°$ values (>2.0 kJ/mol), B1 is very extended and contains the most probable conformers. Such a result shows the possibility of having different, more or less

probable, ligand conformations in different regions of the ES. Therefore, the reduction from 19 to 12 structures thanks to the RMSD analysis does not interfere on the absence of low energy structures in the B2 and B3 basins. As also previously reported, the extension of the three basins, in particular B1, suggests the presence—in each basin—of different conformations which can rapidly mutually interconvert. On the other hand, the B1 ⇆ B2, B1 ⇆ B3, and B2 ⇆ B3 conformational transitions appear as more hindered being characterized by higher energy barriers. This particular feature prompted us to extract a number of RC (12) much larger than the actual number of basins (3) as already described. The energy features of the 12 RC, with their own normalized probabilities, are reported in Table 1. Additional information (i.e., the Cartesian coordinates) can be found in Table S1 of the Supporting Information.

Table 1. Relative Free-Energies and Normalized Probability Values of the 12 $Au_{38}(SC_2H_4Ph)_{24}$ Representative Conformations Selected within the ED Analysis

Basin-Representative Conformation	$\Delta G°$ (kJ/mol)	$P(i)_{norm}$[a]
B1-A	2.0	0.079
B1-B	1.9	0.082
B1-C	0.2	0.165
B1-D	1.8	0.086
B1-E	1.7	0.090
B2-A	2.3	0.071
B2–B	2.4	0.068
B2-C	2.3	0.070
B2-D	2.3	0.070
B3-A	2.3	0.070
B3-B	2.0	0.079
B3-C	2.3	0.071

[a] $\sum_{i=1}^{RC} P(i)_{norm} = 1$

Electronic Circular Dichroism Spectra. The statistically averaged calculated ECD spectrum, using all the selected 12 RC, is compared with the experimental one[23] in Figure 4.

Figure 4 reveals a very good agreement with the experimental pattern in the regions between 250 and 450
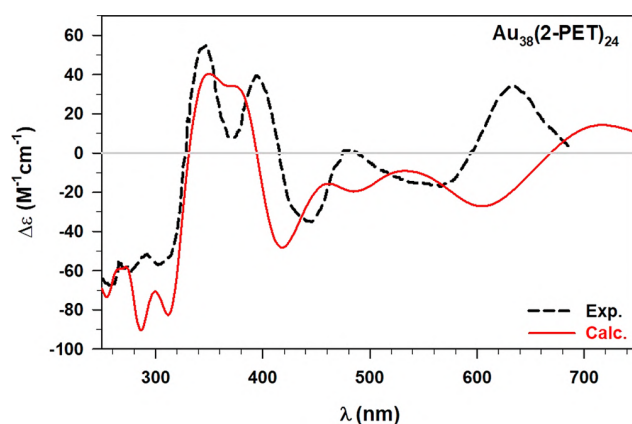
Figure 4. Comparison between the experimental (Exp., black dashed line) and calculated (Calc., red line) ECD spectrum averaged over 12 conformations of the Au_{38}(2-PET)$_{24}$ cluster. All the details of the calculation (i.e., xc functional, basis-set, HWHM) have been reported in the text.

nm, while some energy shifts of the calculated ECD emerge when moving toward higher wavelength values. Indeed, the two calculated peaks at 539 nm (exp. peak at 473 nm) and 712 nm (exp. peak at 627 nm) are systematically found to be red-shifted and also a small maximum peak appears around 460 nm. However, this systematic drawback can be explained by recalling that the metallic response, predominant in this low-energy region, can suffer from the well-known[24,42] attractive character of the LB94 xc functional. On the contrary, the high energy range, where an excellent qualitative agreement between experimental and calculated result is observed, is mostly determined by the spectral features of the flexible thiolate ligands. Therefore, Figure 4 reveals that the combination of a dynamics approach with polTDDFT calculations brings to a qualitatively correct description of the optical response of high-flexible systems, such as the ligands of protected NCs. This result has been corroborated by comparing the statistical ECD with those calculated with different xc functionals but only considering the experimental X-ray structure.[15,24] All the results are reported in Figure 5.

Concerning the high energy range of the ECD spectrum, it can be observed that the calculation on a single structure, even though it is the experimental one obtained by X-ray diffraction, does not reproduce properly the experimental spectral pattern of the 2-PET. Indeed, using the LB94 in the polTDDFT calculation (i.e., middle panel of Figure 5), a negative intense artifact arises around 250 nm, and the two following positive maximum values are blueshifted with respect to experimental peaks. The same blueshift is observed in correspondence of the experimental peak located at 473 nm and even a not negligible difference in intensity arises for this feature. On the contrary, some qualitative similarities between the two ECD spectra can be observed at higher wavelengths, i.e., beyond 600 nm.

In the lower panel of Figure 5, we compare the experimental ECD with the result obtained using the hybrid functional B3LYP (polTDDFT in combination with the HDA). It has been already shown[59] that this approach increases the accuracy of the description of the Au optical response, even though it requires a significant computational effort. Indeed, the calculated spectrum beyond 450 nm is characterized by two maxima (i.e., 452 nm and 585 nm) in reasonable agreement with the experimental values at 473 and 627 nm, respectively. Therefore, despite a minor difference in terms of energy distributions, the hybrid calculation is able to reproduce much better the experimental features in the region of the metallic response. In contrast, in the ligand region of the spectrum, we can notice discrepancies which confirm again the importance of including the ligand conformational effects to model the ECD. The present B3LYP-HDA calculation has been rather demanding, 6 days using 144 cores on the Galileo super-computer of CINECA (Bologna, Italy) consisting of several nodes each with 2 x CPU Intel CascadeLake 8260, with 24 cores each. Instead, for each structure calculated at the LB94 level, the SCF required 5 h at 144 cores on the Galileo supercomputer, while the polTDDFT section took 15 h with 24 CPU on a HP ProLiant DL580 Gen10 server. The latter consists of 4 processors each with 18 cores Intel Xeon Gold 6140 CPU @ 2.30 GHz, in total 72 cores and 728 GB of RAM. We are working on a new implementation of the B3LYP-HDA scheme based on the Resolution of the Identity instead of the presently implemented numerical integration, which is very promising in terms of numerical efficiency.
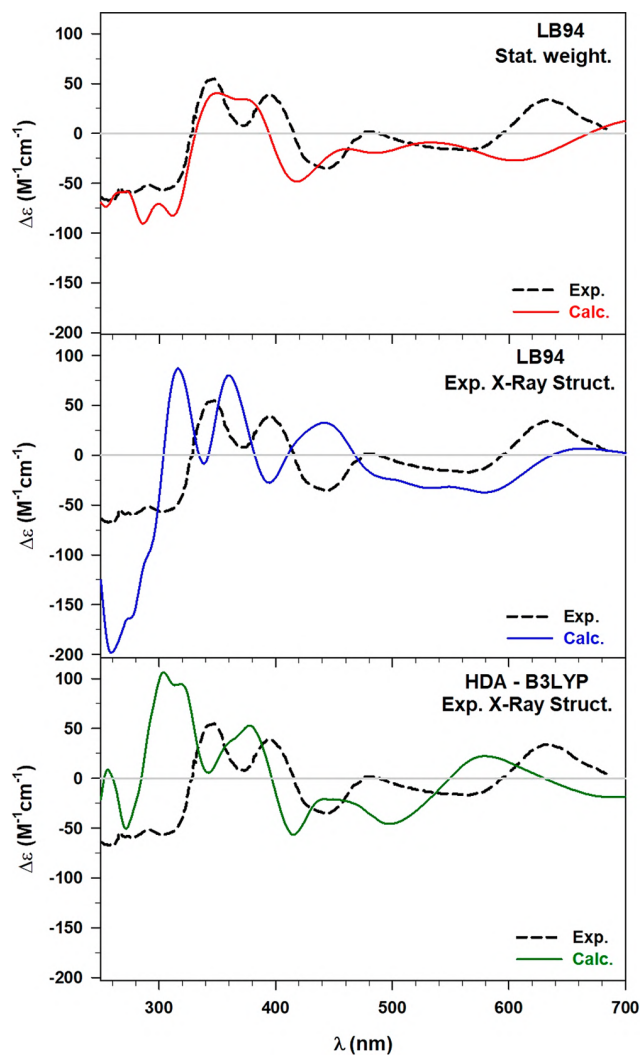
Figure 5. Comparison between the experimental (Exp., black dashed line) and the calculated ECD spectrum. (Top) statistically weighting 12 conformations and by using the LB94 xc functional (upper panel, red line), (middle) using the exp. X-ray structure and the LB94 xc functional (middle panel, blue line), (bottom) using the exp. X-ray structure and the HDA in combination with the B3LYP xc functional (lower panel, green line).

In conclusion, the above data clearly suggest that for a correct ECD modeling of chiral RS-AuNCs, at least the relevant conformations of the protective ligands, must be included in the response predictions. This effect becomes more and more relevant with the increase of the ligand flexibility, but also in the presence of intramolecular interactions as well as solvent effects, especially considering the aqueous environment. The metal core description, which is much less affected by conformational effects, suffers the limitations of the LB94 functional. However, preliminary calculations using the HDA[59]-B3LYP scheme demonstrate the possibility of a better description also of the Au response.

In conclusion, we propose a computational approach that is able to efficiently and accurately include the conformational effects of the dynamic protective ligands into the ECD modeling of chiral RS-AuNCs. Extending a protocol recently proposed for solvated peptides, we investigate these more complex gold NCs and their chiroptical properties focusing on

the well-known $Au_{38}(2\text{-PET})_{24}$ cluster. We start with a MD simulation of this system in toluene using an accurate forcefield, which is then analyzed to sample the conformational space of the ligands via the ED analysis. This statistical tool allows us to consider only the *essential* conformational dynamics, hence to extract properly the most relevant conformations of the system. A reduced number of structures is then considered for the first-principles polTDDFT calculations of the ECD. We obtain an excellent agreement between the experimental and calculated spectral features of the 2-PET ligands both in terms of energy and intensity. Such a result is corroborated by calculating the ECD only on the X-ray crystal structure with two different xc functionals (LB94, B3LYP) and finding a much worse agreement with the experimental response. This demonstrates the need to include conformational effects into the modeling of the ECD of gold nanoclusters, together with the proposal of a reproducible and affordable computational approach to accomplish this goal. Some limitations are found in the description of the metal core response, but we assume that they are related to the chosen xc-functional and not to the procedure here proposed.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.2c03923.

> The experimental X-ray structure, the comparison between all-atoms and Au-constrained MD simulation, the additional analysis on the conformational landscape for the extraction of the final set of 12 RC, and the atomic coordinates of the conformers (PDF)
>
> Transparent Peer Review report available (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

M. Stener − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy;* ⓞ orcid.org/0000-0003-3700-7903; Email: stener@units.it

A. Fortunelli − *CNR-ICCOM, Consiglio Nazionale delle Ricerche, 56124 Pisa, Italy;* ⓞ orcid.org/0000-0001-5337-4450; Email: alessandro.fortunelli@cnr.it

M. Aschi − *Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, 67100 l'Aquila, Italy;* ⓞ orcid.org/0000-0003-2959-0158; Email: massimiliano.aschi@univaq.it

Authors

M. Monti − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy*

G. Brancolini − *Istituto Nanoscienze, CNR-NANO, Center S3, 41100 Modena, Italy*

E. Coccia − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy;* ⓞ orcid.org/0000-0003-3389-0989

D. Toffoli − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy;* ⓞ orcid.org/0000-0002-8225-6119

S. Corni − *Istituto Nanoscienze, CNR-NANO, Center S3, 41100 Modena, Italy; Dipartimento di Scienze Chimiche, Università di Padova, 35131 Padova, Italy;* ⓞ orcid.org/0000-0001-6707-108X

Complete contact information is available at:

https://pubs.acs.org/10.1021/acs.jpclett.2c03923

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Gautier, C.; Taras, R.; Gladiali, S.; Bürgi, T. Chiral 1,1′-binaphthyl-2,2′-dithiol-stabilized gold clusters: Size separation and optical activity in the UV−vis. *Chirality* **2008**, *20*, 486−493.

(2) Zhu, Y.; Qian, H.; Jin, R. Catalysis opportunities of atomically precise gold nanoclusters. *J. Mater. Chem.* **2011**, *21*, 6793−6799.

(3) Lim, I-I. S.; Mott, D.; Engelhard, M. H.; Pan, Y.; Kamodia, S.; Luo, J.; Njoki, P. N.; Zhou, S.; Wang, L.; Zhong, C. J. Interparticle chiral recognition of enantiomers: a nanoparticle-based regulation strategy. *Anal. Chem.* **2009**, *81*, 689−698.

(4) Kang, Y. J.; Oh, J. W.; Kim, Y. R.; Kim, J. S.; Kim, H. Chiral gold nanoparticle-based electrochemical sensor for enantioselective recognition of 3,4-dihydroxyphenylalanine. *Chem. Commun.* **2010**, *46*, 5665−5667.

(5) Shukla, N.; Bartel, M. A.; Gellman, A. J. Enantioselective Separation on Chiral Au Nanoparticles. *J. Am. Chem. Soc.* **2010**, *132*, 8575−8580.

(6) Zhu, M.; Aikens, C. M.; Hollander, F. J.; Schatz, G. C.; Jin, R. Correlating the Crystal Structure of a Thiol-Protected Au_{25} Cluster and Optical Properties. *J. Am. Chem. Soc.* **2008**, *130*, 5883−5885.

(7) Xu, Q.; Kumar, S.; Jin, S.; Qian, H.; Zhu, M.; Jin, R. Chiral 38-Gold-Atom Nanoclusters: Synthesis and Chiroptical Properties. *Small* **2014**, *10*, 1008−1014.

(8) Garzón, I. L.; Reyes-Nava, J. A.; Rodríguez-Hernández, J.; Sigal, I.; Beltrán, M.; Michaelian, K. Chirality in bare and passivated gold nanoclusters. *Phys. Rev. B* **2002**, *66*, 073403.

(9) Gautier, C.; Bürgi, T. Chiral N-Isobutyryl-cysteine Protected Gold Nanoparticles: Preparation, Size Selection, and Optical Activity in the UV−vis and Infrared. *J. Am. Chem. Soc.* **2006**, *128*, 11079−11087.

(10) Goldsmith, M. R.; George, C. B.; Zuber, G.; Naaman, R.; Waldeck, D. H.; Wipf, P.; Beratan, D. N. The chiroptical signature of achiral metal clusters induced by dissymmetric adsorbates. *Phys. Chem. Chem. Phys.* **2006**, *8*, 63−67.

(11) Knoppe, S.; Bürgi, T. Chirality in thiolate-protected gold clusters. *Acc. Chem. Res.* **2014**, *47*, 1318−1326.

(12) Yamazoe, S.; Takano, S.; Kurashige, W.; Yokoyama, T.; Nitta, K.; Negishi, Y.; Tsukuda, T. Hierarchy of bond stiffnesses within icosahedral-based gold clusters protected by thiolates. *Nat. Commun.* **2016**, *7*, 10414.

(13) Negishi, Y.; Igarashi, K.; Munakata, K.; Ohgake, W.; Nobusada, K. Palladium doping of magic gold cluster $Au_{38}(SC_2H_4Ph)_{24}$: formation of $Pd_2Au_{36}(SC_2H_4Ph)_{24}$ with higher stability than $Au_{38}(SC_2H_4Ph)_{24}$. *Chem. Commun.* **2012**, *48*, 660−662.

(14) Nie, X.; Zeng, C.; Ma, X.; Qian, H.; Ge, Q.; Xu, H.; Jin, R. CeO_2-supported $Au_{38}(SR)_{24}$ nanocluster catalysts for CO oxidation: a

comparison of ligand-on and -off catalysts. *Nanoscale* **2013**, *5*, 5912−5918.

(15) Qian, H.; Eckenhoff, W. T.; Zhu, Y.; Pintauer, T.; Jin, R. Total Structure Determination of Thiolate-Protected Au$_{38}$ Nanoparticles. *J. Am. Chem. Soc.* **2010**, *132*, 8280−8281.

(16) Schaaff, T. G.; Shafigullin, M. N.; Khoury, J. T.; Vezmar, I.; Whetten, R. L.; Cullen, W. G.; First, P. N.; GutiéRrez-Wing, C.; Ascensio, J.; Jose-Yacamán, M. J. Isolation of smaller nanocrystal Au molecules: robust quantum effects in optical spectra. *J. Phys. Chem. B* **1997**, *101*, 7885−7891.

(17) Garzon, I. L.; Michaelian, K.; Beltran, M. R.; Posada-Amarillas, A.; Ordejon, P.; Artacho, E.; Sanchez-Portal, D.; Soler, J. M. Lowest energy structures of gold nanoclusters. *Phys. Rev. Lett.* **1998**, *81*, 1600−1603.

(18) Häkkinen, H.; Barnett, R. N.; Landman, U. Electronic Structure of Passivated Au$_{38}$(SCH$_3$)$_{24}$ Nanocrystal. *Phys. Rev. Lett.* **1999**, *82*, 3264−3267.

(19) Garzon, I. L.; Rovira, C.; Michaelian, K.; Beltran, M. R.; Ordejon, P.; Junquera, J.; Sanchez-Portal, D.; Artacho, E.; Soler, J. M. Do Thiols Merely Passivate Gold Nanoclusters? *Phys. Rev. Lett.* **2000**, *85*, 5250−5251.

(20) Pei, Y.; Gao, Y.; Zeng, X. Structural Prediction of Thiolate-Protected Au$_{38}$: A Face-Fused Bi-icosahedral Au Core. *J. Am. Chem. Soc.* **2008**, *130*, 7830−7832.

(21) Lopez-Acevedo, O.; Tsunoyama, H.; Tsukuda, T.; Häkkinen, H.; Aikens, C. M. Chirality and Electronic Structure of the Thiolate-Protected Au$_{38}$ Nanocluster. *J. Am. Chem. Soc.* **2010**, *132*, 8210−8218.

(22) Dolamic, I.; Knoppe, S.; Dass, A.; Bürgi, T. First enantioseparation and circular dichroism spectra of Au$_{38}$ clusters protected by achiral ligands. *Nat. Commun.* **2012**, *3*, 798.

(23) Barrabés, N.; Zhang, B.; Bürgi, T. Racemization of Chiral Pd$_2$Au$_{36}$(SC$_2$H$_4$Ph)$_{24}$: Doping Increases the Flexibility of the Cluster Surface. *J. Am. Chem. Soc.* **2014**, *136*, 14361−14364.

(24) Baseggio, O.; Toffoli, D.; Fronzoni, G.; Stener, M.; Sementa, L.; Fortunelli, A. Extension of the Time-Dependent Density Functional Complex Polarizability Algorithm to Circular Dichroism: Implementation and Applications to Ag$_8$ and Au$_{38}$(SC$_2$H$_4$C$_6$H$_5$)$_{24}$. *J. Phys. Chem. C* **2016**, *120*, 24335−24345.

(25) Toffoli, D.; Baseggio, O.; Fronzoni, G.; Stener, M.; Fortunelli, A.; Sementa, L. Pd doping, conformational, and charge effects on the dichroic response of a monolayer protected Au$_{38}$(SR)$_{24}$ nanocluster. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3585−3596.

(26) Dolamic, I.; Varnholt, B.; Bürgi, T. Chirality transfer from gold nanocluster to adsorbate evidenced by vibrational circular dichroism. *Nat. Commun.* **2015**, *6*, 7117.

(27) Nieto-Ortega, B.; Bürgi, T. Vibrational properties of thiolate-protected gold nanoclusters. *Acc. Chem. Res.* **2018**, *51*, 2811−2819.

(28) Riccardi, L.; De Biasi, F.; De Vivo, M.; Bürgi, T.; Rastrelli, F.; Salassa, G. Dynamic Origin of Chirality Transfer between Chiral Surface and Achiral Ligand in Au$_{38}$ Clusters. *ACS Nano* **2019**, *13*, 7127−7134.

(29) Pyo, K.; Matus, M. F.; Malola, S.; Hulkko, E.; Alaranta, J.; Lahtinen, T.; Häkkinen, H.; Pettersson, M. Tailoring the interaction between a gold nanocluster and a fluorescent dye by cluster size: creating a toolbox of range-adjustable pH sensors. *Nanoscale Adv.* **2022**, *4*, 4579−4588.

(30) Monti, M.; Stener, M.; Aschi, M. A computational approach for modeling electronic circular dichroism of solvated chromophores. *J. Comput. Chem.* **2022**, *43*, 2023−2036.

(31) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.

(32) Brancolini, G.; Toroz, D.; Corni, S. Can small hydrophobic gold nanoparticles inhibit *β*2-microglobulin fibrillation? *Nanoscale* **2014**, *6*, 7903−7911.

(33) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(34) Brancolini, G.; Rotello, V. M.; Corni, S. Role of Ionic Strength in the Formation of Stable Supramolecular Nanoparticle−Protein Conjugates for Biosensing. *Int. J. Mol. Sci.* **2022**, *23*, 2368.

(35) Ray, M.; Brancolini, G.; Luther, D. C.; Jiang, Z.; Cao-Milán, R.; Cuadros, A. M.; Burden, A.; Clark, V.; Rana, S.; Mout, R.; Landis, R. F.; Corni, S.; Rotello, V. M. High affinity protein surface binding through co-engineering of nanoparticles and proteins. *Nanoscale* **2022**, *14*, 2411−2418.

(36) Dutta, S.; Corni, S.; Brancolini, G. Molecular Dynamics Simulations of a Catalytic Multivalent Peptide−Nanoparticle Complex. *Int. J. Mol. Sci.* **2021**, *22*, 3624.

(37) Dodda, L. S.; Cabeza de Vaca, I.; Tirado Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45* (W331), W336.

(38) Bussi, G.; Parrinello, M. Stochastic Thermostats: Comparison of Local and Global Schemes. *Comput. Phys. Commun.* **2008**, *179*, 26.

(39) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(40) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(41) McLinden, M. O.; Splett, J. D. A Liquid Density Standard Over Wide Ranges of Temperature and Pressure Based on Toluene. *Res. Natl. Inst. Stand. Technol.* **2008**, *113*, 29−67.

(42) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Ramankutty Krishnadas, J.; Burgi, T.; Bonacchi, S.; Dainese, T.; Maran, F.; Fortunelli, A.; Stener, M. Predictive optical photo-absorption of Ag$_{24}$Au(DMBT)$_{18}^-$ via efficient TDDFT simulations. *J. Chem. Phys.* **2021**, *155*, 084103.

(43) Knoppe, S.; Michalet, S.; Bürgi, T. Stabilization of Thiolate-Protected Gold Clusters Against Thermal Inversion: Diastereomeric Au$_{38}$(SCH$_2$CH$_2$Ph)$_{24−2x}$(R-BINAS)$_x$. *J. Phys. Chem. C* **2013**, *117*, 15354−15361.

(44) Malola, S.; Häkkinen, H. Chiral Inversion of Thiolate-Protected Gold Nanoclusters via Core Reconstruction without Breaking a Au−S Bond. *J. Am. Chem. Soc.* **2019**, *141*, 6006−6012.

(45) Amadei, A.; Linssen, B. M.; Berendsen, H. J. C. Essential Dynamics of proteins. *Proteins* **1993**, *17*, 412−425.

(46) Daidone, I.; Amadei, A. Essential dynamics: foundation and applications. *WIREs Comput. Mol. Sci.* **2012**, *2*, 762.

(47) D'Alessandro, M.; Amadei, A.; Stener, M.; Aschi, M. Essential dynamics for the study of microstructures in liquids. *J. Comput. Chem.* **2015**, *36*, 399−407.

(48) Amadei, A.; D'Alessandro, M.; Aschi, M. Statistical Mechanical Modeling of Chemical Reactions in Complex Systems: The Reaction Free Energy Surface. *J. Phys. Chem. B* **2004**, *108*, 16250−16254.

(49) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, Oxford, 1989.

(50) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradien Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.

(51) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(52) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597.

(53) Baseggio, O.; Fronzoni, G.; Stener, M. A new time dependent density functional algorithm for large systems and plasmons in metal clusters. *J. Chem. Phys.* **2015**, *143*, 024106.

(54) Pikulska, A.; Hopmann, K. H.; Bloino, J.; Pecul, M. Circular Dichroism and Optical Rotation of Lactamide and 2-Aminopropanol in Aqueous Solution. *J. Phys. Chem. B* **2013**, *117*, 5136−5147.

(55) Fleming, A. M.; Orendt, A. M.; He, Y.; Zhu, J.; Dukor, R. K.; Burrows, C. J. Reconciliation of Chemical, Enzymatic, Spectroscopic and Computational Data To Assign the Absolute Configuration of the DNA Base Lesion Spiroiminodihydantoin. *J. Am. Chem. Soc.* **2013**, *135*, 18191−18204.

(56) Pescitelli, G.; Bruhn, T. Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra. *Chirality* **2016**, *28*, 466−474.

(57) van Leeuwen, R.; Baerends, E. J. Exchange-correlation potential with correct asymptotic behaviour. *Phys. Rev. A* **1994**, *49*, 2421.

(58) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(59) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Fortunelli, A.; Stener, M. An efficient hybrid scheme for time dependent density functional theory. *J. Chem. Phys.* **2020**, *152*, 184104.

(60) Aschi, M.; D'Abramo, M.; Amadei, A. Photoinduced electron transfer in a dichromophoric peptide: a numerical experiment. *Theor. Chem. Acc.* **2016**, *135*, 132−142.

📖 Recommended by ACS

The Conformational Dynamics of the Ligands Determines the Electronic Circular Dichroism of the Chiral $Au_{38}(SC_2H_4Ph)_{24}$ Cluster

M. Monti[1], G. Brancolini[2], E. Coccia[1], D. Toffoli[1], A. Fortunelli[3], S. Corni[2,4], M. Aschi[5], and M. Stener[1]

[1] Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy.

[2] Centro S3, CNR Instituto di Nanoscienza, via Campi 213/A, 41125 Modena, Italy.

[3] CNR-ICCOM, Consiglio Nazionale delle Ricerche, via G. Moruzzi 1, 56124, Pisa, Italy.

[4] Dipartimento di Scienze Chimiche, Università di Padova, Via Francesco Marzolo 1, 35131 Padova, Italy.

[5] Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, Via Vetoio, 67100, l'Aquila, Italy
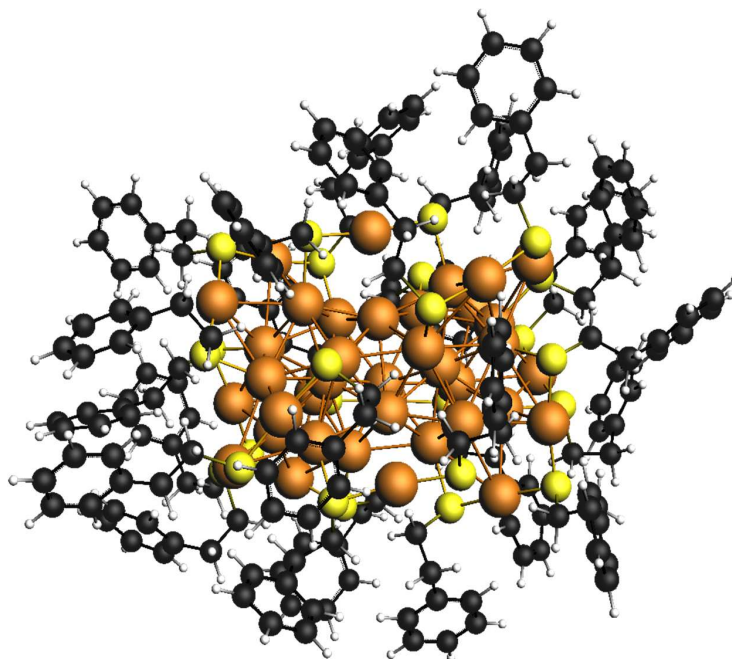
Figure S1. Ball and stick model of the experimental X-ray $Au_{38}(SC_2H_4Ph)_{24}$ structure[1]. Au, S, C, and H atoms are reported in orange, yellow, black, and white, respectively.

All-atoms MD vs Au-constrained MD

It is well known the importance of calculating the ECD spectra on realistic geometries to obtain reliable results and in the specific case of gold nanoclusters, the relevance of a reasonable metal-sulphur architecture has proven to be fundamental for the ECD calculation[1,2]. Therefore, the resolution of several X-ray gold nanostructures has represented a breakthrough for a suitable analysis of the spectral features. Naturally, the X-ray structure is less important when the chirality regards the ligands because of their higher conformational flexibility, which characterizes their ECD in the liquid phase. Hence, the calculation of the ECD on a gas phase nanocluster and its comparison with respect the experimental liquid phase ECD must consider the flexibility of the ligands as well as the necessity of a reliable geometry of the staples (Au-S structures). Starting from these considerations, we constrained the gold atoms to maintain the experimental geometry for the metallic region but without affecting the ligands internal motions. However, we wanted to verify the validity of our choice estimating the effective contribution of the gold atoms to the conformational transitions and to understand if their exclusion in the conformational analysis is justified. Such an estimate was

performed running a 50 ns all-atoms MD simulation with no constraints and calculating the first two total eigenvectors components of the covariance matrix for Au and S (see Figure S2).
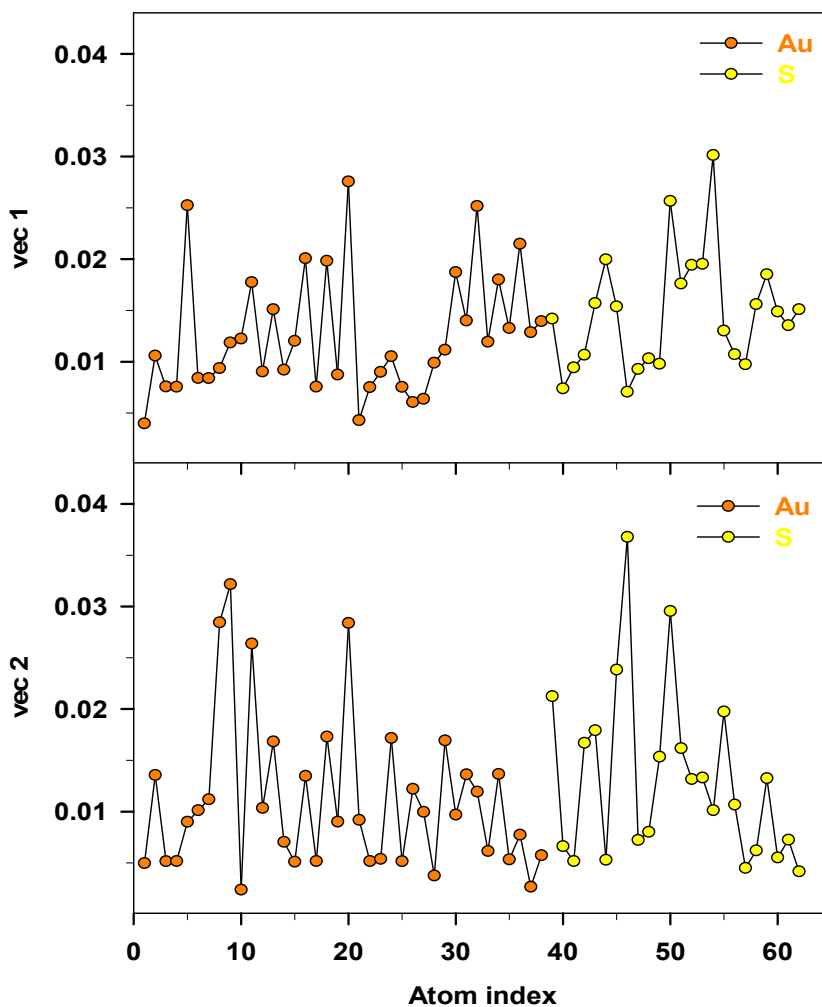
Figure S2. First eigenvector (upper panel) and second eigenvector (lower panel) total components for the gold (orange) and sulphur (yellow) atoms calculated from the 50 ns all atoms free-MD simulation.

Looking at the value scale of the total first pair of eigenvectors, it is worth noting that the contribution of both gold and sulphur atoms to the most relevant conformational transitions is quite small. Such result justifies the constraints on the gold atoms and thus their exclusion from the conformational analysis we carried out to find the most probable structures. The following Figure S3 underlines how small is the contributions of the staple atoms with respect to the ligand ones, which start from atom index 63.
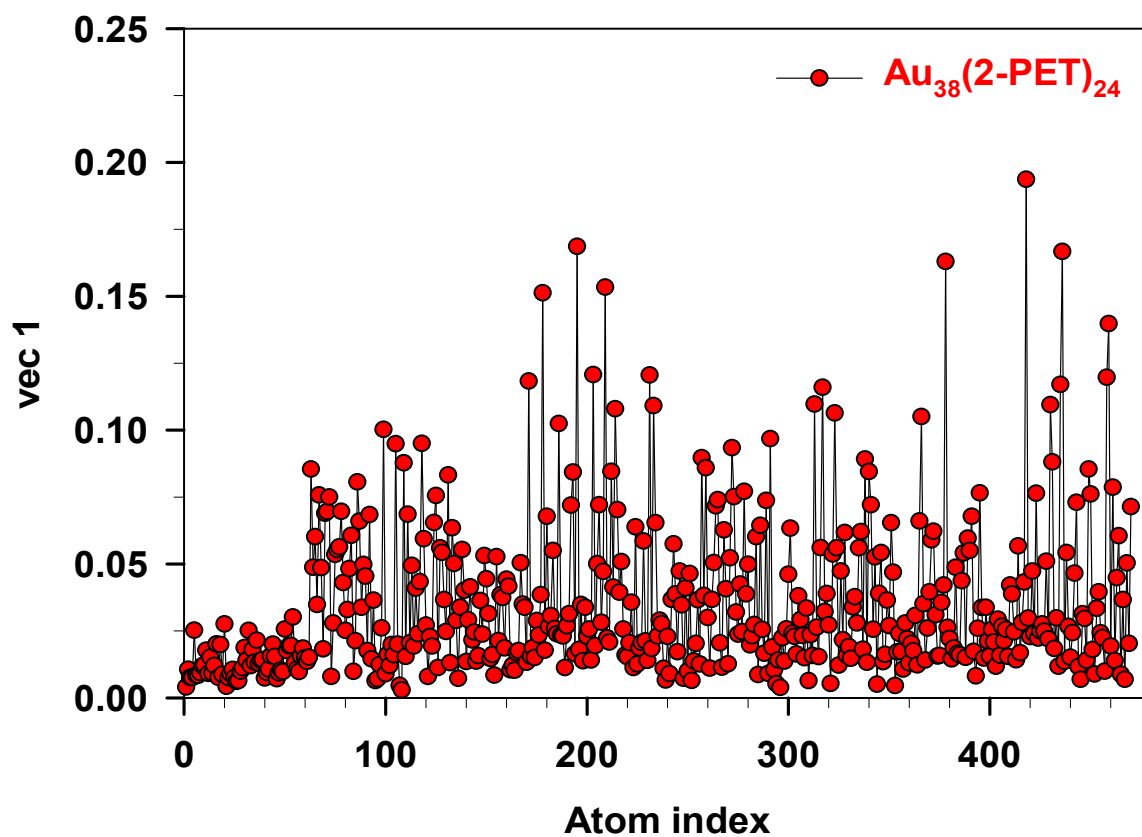
Figure S3. First eigenvector total component for the $Au_{38}(2\text{-PET})_{24}$ atoms calculated from the 50 ns all atoms free-MD simulation.

As we can see, we obtained that the total component of the first eigenvector, which is directly related to the atomic contribution to the most relevant internal transitions, is on average 10 times larger for the (2-PET) atoms with respect to the values found for the Au/S atoms. This result is easily understood comparing the flexibility, and thus the fluctuations, of the metallic and organic components of the system.
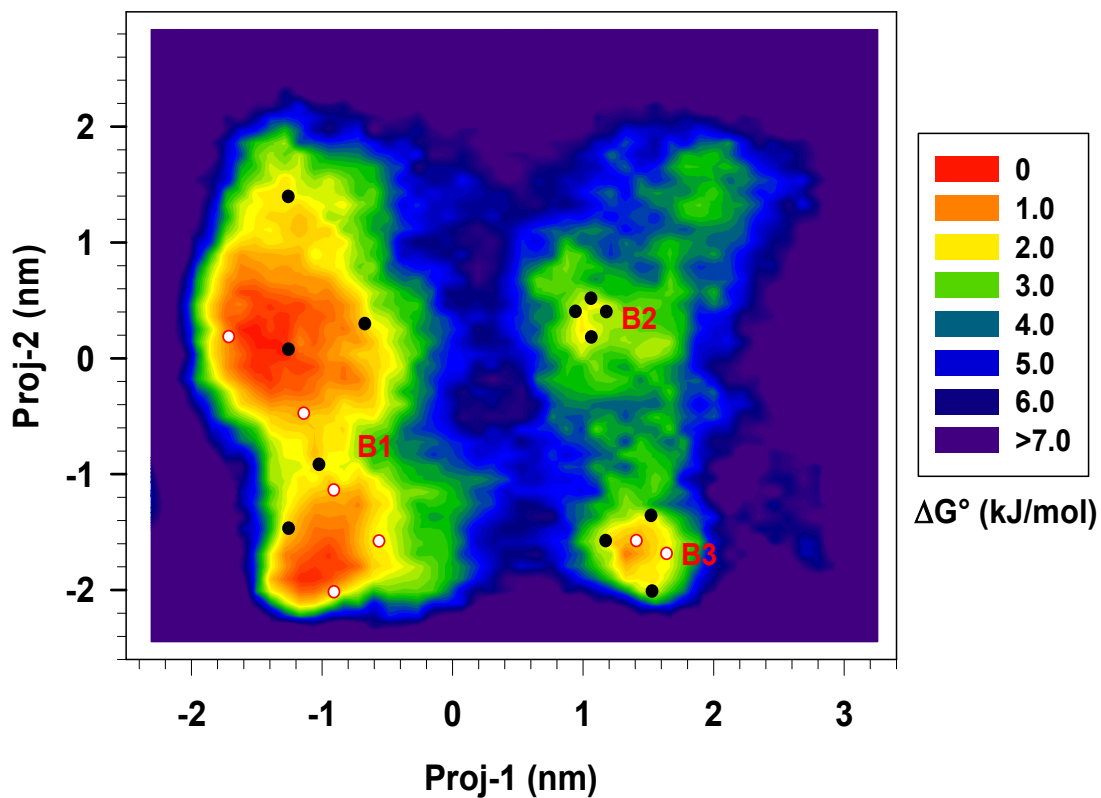
Figure S4. Probability pattern built to represent the ΔG° landscape on the conformational space. The positions of the 19 conformations initially extracted have been marked with dots (i.e., black for the conformations of the final set, white for the conformations discarded in the final analysis).
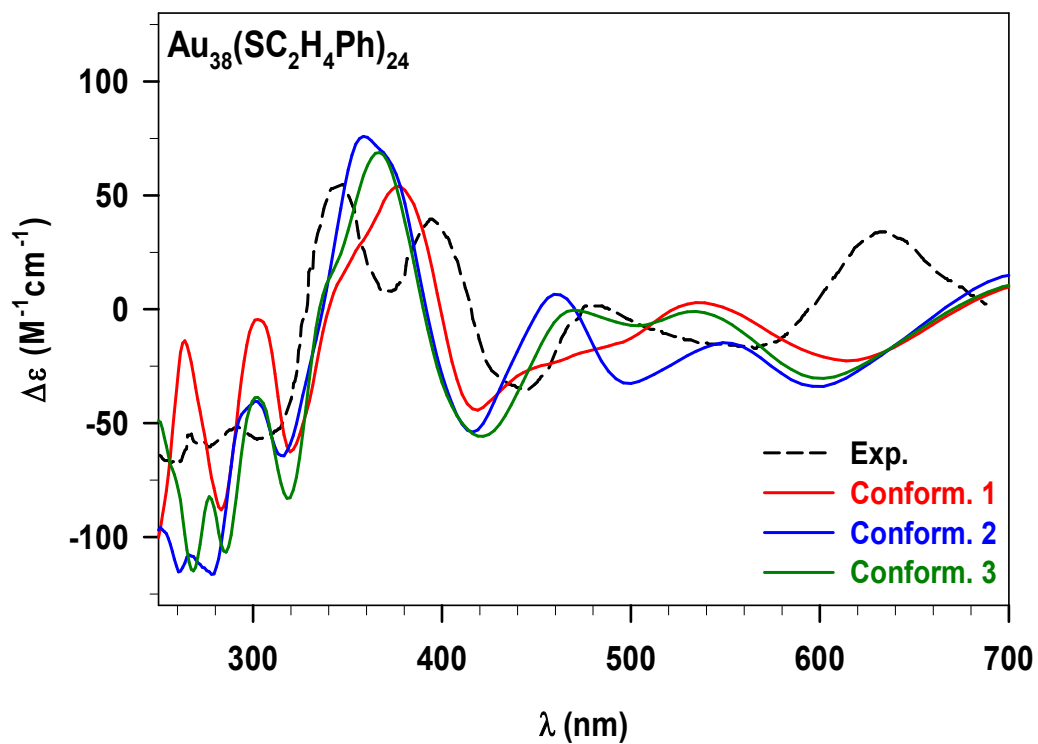
Figure S5. Experimental (Exp., black dashed line) and Calculated ECD spectra of three similar conformations in terms of RMSD values (red, blue, and green solid line). The computational details of the polTDDFT calculations have been reported in the Computational Section.

N.B.: In the original version of the SI, we reported all the coordinates of the different conformers that were selected and employed to calculate the ECD spectra.
They are available at: https://pubs.acs.org/doi/full/10.1021/acs.jpclett.2c03923

References

[1] Baseggio, O.; Toffoli, D.; Fronzoni, G.; Stener, M.; Sementa, L.; Fortunelli, A. Extension of the Time-Dependent Density Functional Complex Polarizability Algorithm to Circular Dichroism: Implementation and Applications to Ag_8 and $Au_{38}(SC_2H_4C_6H_5)_{24}$. *J. Phys. Chem. C* **2016**, *120*, 24335-24345.

[2] Baseggio, O.; de Vetta, M.; Fronzoni, G.; Toffoli, D.; Stener, M.; Sementa, L.; Fortunelli, A. Time-dependent density-functional study of the photoabsorption spectrum of Au25(SC2H4C6H5)18 anion: Validation of the computational protocol. *Int. J. Quantum Chem.* **2018**, *118*, e25769.

4.1.3 What Contributes to the Measured Chiral Optical Response of the Glutathione-Protected Au$_{25}$ Nanocluster?

The aim of this work was to study the chiral features of the [Au$_{25}$(GSH)$_{18}$]$^{-1}$ nanocluster in aqueous environment. This system is particularly fascinating because of its solubility in water and its biological applications. Indeed, it can be used for the selective binding of target molecules such as the glutathione-*S*-transferase, a metabolic isozyme which catalyzes the conjugation of GSH to xenobiotic substrates for detoxification purposes. [Au$_{25}$(GSH)$_{18}$]$^{-1}$ is also quite appealing from a theoretical point of view since the ECD spectrum of the whole structure has never been calculated before, despite being one of the first discovered chiral nanoclusters. Indeed, the system size (655 atoms) is quite prohibitive. In addition, the solvent plays a major role here, thus it should be included explicitly.

We adopted the procedure explained above, employing a unified AMBER-Compatible force field for the Au-S interface, and standard AMBER parameters for glutathione. This force field was developed by Prof. Hannu Häkkinen and his group, who were personally involved in this study. Other parameters of the procedure resemble those adopted in the previous work (section 4.1.2). In this case, the final ECD spectrum was obtained by averaging the optical responses of 13 statistically relevant [Au$_{25}$(GSH)$_{18}$]$^{-1}$-(H$_2$O)$_{126}$ clusters. The comparison of the experimental and calculated spectra revealed a good qualitative agreement in the lower energy region, but some discrepancies arise between (3.5-5) eV. Such differences have been ascribed to the complexity of the system, as well as to some aspects of the computational procedure. For instance, the classical MD trajectory could be affected by some limitations that result in an imbalance of the statistical weights in the ED analysis. Such hypothesis was tested by averaging the spectra of different subsets of conformers, bringing to a better agreement with the experimental spectrum.

Despite that, this work revealed its strength in the explicit inclusion of the solvent. This allowed us to investigate how the water molecules affect the chiroptical features of a thiolate-protected gold nanocluster. As a result, we demonstrated that not only the aqueous solvent has an effect on the conformations, due to the inter-molecular interactions, but it also explicitly contributes to the electronic transitions of the nanocluster. Indeed, separating the contribution of the molecular orbitals (MOs) to the different transitions, it was shown that water MOs are directly involved in the spectral features above 3.5 eV. Furthermore, comparing the ECD of the bare nanocluster with that of the solvated one, we observed that the high-energy features are completely suppressed in the absence of water. Finally, it was also revealed that the water-shell itself assumes a chiral arrangement due to the interactions with the nanocluster, thus showing a non-null ECD response.

What Contributes to the Measured Chiral Optical Response of the Glutathione-Protected Au$_{25}$ Nanocluster?

Marta Monti, María Francisca Matus, Sami Malola, Alessandro Fortunelli, Massimiliano Aschi, Mauro Stener,* and Hannu Häkkinen*

ACCESS | Metrics & More | Article Recommendations | Supporting Information
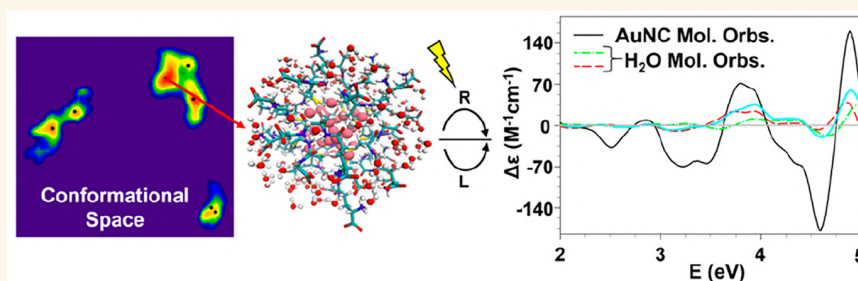
ABSTRACT: The water-soluble glutathione-protected $[\mathrm{Au_{25}(GSH)_{18}}]^{-1}$ nanocluster was investigated by integrating several methodologies such as molecular dynamics simulations, essential dynamics analysis, and state-of-the-art time-dependent density functional theory calculations. Fundamental aspects such as conformational, weak interactions and solvent effects, especially hydrogen-bonds, were included and found to play a fundamental role in assessing the optical response of this system. Our analysis demonstrated not only that the electronic circular dichroism is extremely sensitive to the solvent presence but also that the solvent itself plays an active role in the optical activity of such system, forming a chiral solvation shell around the cluster. Our work demonstrates a successful strategy to investigate in detail chiral interfaces between metal nanoclusters and their environments, applicable, e.g., to chiral electronic interactions between clusters and biomolecules.

KEYWORDS: gold, nanocluster, thiols, chirality, molecular dynamics, density functional theory, essential dynamics

Chirality in thiolate-protected gold nanoclusters (RS-AuNCs) was detected more than 20 years ago by Whetten and collaborators[1,2] who worked on clusters with 20−40 Au atoms protected by the L-glutathione (GSH). Since then, significant progress has been made in the synthesis and total structural characterization as well as on the experimental and theoretical investigations of the electronic and chiroptical properties of RS-AuNCs.[3−14] The current view suggests that the chirality in RS-AuNCs can be classified as intrinsic or induced. The intrinsic one arises when the metal core is chiral itself or can be related to chiral arrangements of the protective achiral ligands, while the second case regards chiral ligands which induce the chirality on the metal architecture. Moreover, several studies have been carried out investigating the peculiar optical properties of these atomically precise systems, which exhibit a discrete electronic structure when the size becomes smaller than *ca.* 2 nm.[15−18] The interest in the discrete structures and properties of RS-AuNCs arises

because of their potential applications in numerous fields such as catalysis,[19] chemical sensing,[20] optical devices,[21] and biomedicine,[22] providing possibility of tuning their properties according to size, shape, and composition.[23,24]

In this work, we focus on one of the first discovered RS-AuNCs, the chiral $[\mathrm{Au_{25}(GSH)_{18}}]^{-1}$, interesting especially in the biomedical field for the selective binding of target molecules such as the glutathione-*S*-transferase.[24] While some initial confusion existed in the literature about its proper chemical composition,[1,25] the chemical formula $[\mathrm{Au_{25}(SR)_{18}}]^{\pm 1}$ was unambiguously identified based on several

high-resolution mass spectroscopy experiments from 2005–2007.[26−29] Over the last two decades, several efforts have also been made to understand the atomic structure of $[Au_{25}(GSH)_{18}]^{-1}$ since its crystal structure is still unresolved. Indeed, despite the $[Au_{25}(GSH)_{18}]^{-1}$ stability demonstrated by Shichibu and co-workers,[28] the bulky and highly flexible GSH ligands make the crystal growth to high-quality single crystals very challenging. However, NMR studies by Wu and co-workers[30] imply that the Au−S architecture is compatible with that found for $[Au_{25}(SCH_2CH_2Ph)_{18}]^{-1,5,12}$ hence with an icosahedral Au_{13} core and six $Au_2(GSH)_3$ motifs. Experimental circular dichroism (CD) spectra in the UV−vis region reported in more recent studies[30,31] agree with the original data from Whetten and collaborators,[1] reconfirming the assignment of the chemical composition.

Theoretical studies on the induced effects of chiral ligands on the electronic structure and chiroptical properties of GSH-protected small Au clusters are scarce. The only exceptions are a density functional theory study[32] where cysteine was used as a small chiral model ligand for $[Au_{25}(SR)_{18}]^{-1}$, excluding solvent effects, and studies of $Au_{18}(GSH)_{14}{}^{33}$ and $[Au_{25}(GSH)_{18}]^{-1\,34}$ using quantum mechanics/molecular mechanics (QM/MM) simulations, which showed that ligand dynamics and solvent effects are important for clusters' structural and electronic properties. Although these results are in line with what is known for small molecules and biomolecules in various solvents,[35−38] no attempts have been made up to date to understand the original CD data measured for $[Au_{25}(GSH)_{18}]^{-1}$ over 20 years ago[2] considering a full theoretical system modeling not only the relevant ligand but also dynamics (conformations) of the ligand shell and solvent−ligand interactions.

Here, we evaluate and analyze the UV−visible CD spectrum of $[Au_{25}(GSH)_{18}]^{-1}$ by combining classical molecular dynamics simulations, essential dynamics analysis, and state-of-the-art time-dependent density functional theory calculations including conformational and solvent effects. We find that the aqueous solvent plays important direct and indirect roles in the chiroptical response of $[Au_{25}(GSH)_{18}]^{-1}$ (see Figure 1) particularly in the UV energy range, where CD peaks can have up to 50% of the intensity originating from transitions involving electronic states of the solvation shell, which by itself has a chiral arrangement around the organic thiol surface. Our work demonstrates the importance to treat environmental factors properly in theoretical studies of RS-AuNCs in order to get a comprehensive understanding of their chiroptical response.

RESULTS

Full-Molecular Dynamics (MD) Simulation and $[Au_{25}(GSH)_{18}]^{-1}$ Conformational States. Figure 2 shows the spectrum of the first 30 eigenvalues extracted from the diagonalization of the nanocluster covariance matrix as obtained from the full-MD simulation. The overall picture shows that the eigenvalues, i.e., the mean square fluctuations of the system, rapidly decrease and a large fraction (around 60%) of the trace of the matrix (corresponding to the whole nanocluster fluctuation) arises from the first 5 eigenvalues. Such a relatively high number of essential eigenvectors, i.e., principal directions along which the internal motions happen, is clearly due to the size of the nanocluster, not only in terms of number of atoms but also in terms of internal flexibility. However, the first component accounts for 30% of this
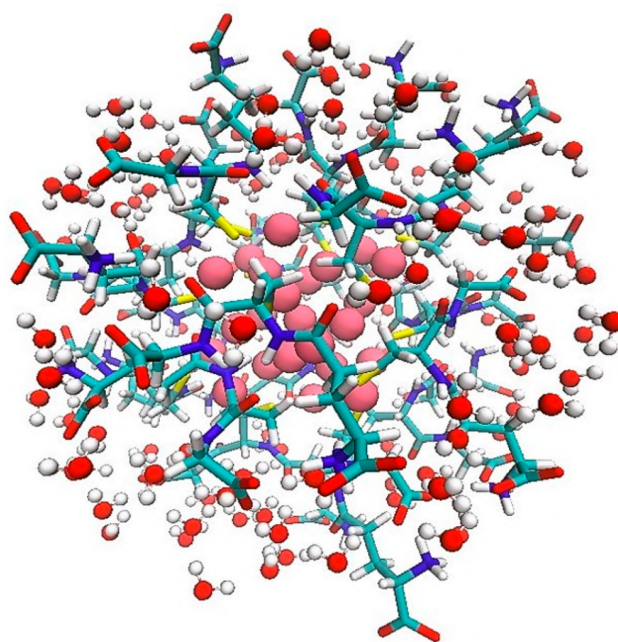

Figure 1. Stick and balls example of the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ cluster conformation extracted with the procedure proposed. Au atoms are reported as pink balls, S, C, N, O, and H atoms of the GSH ligand are reported as sticks in yellow, cyan, blue, red, and white, respectively. A stick and ball model is used for the water molecules with O and H atoms in red and white, respectively.
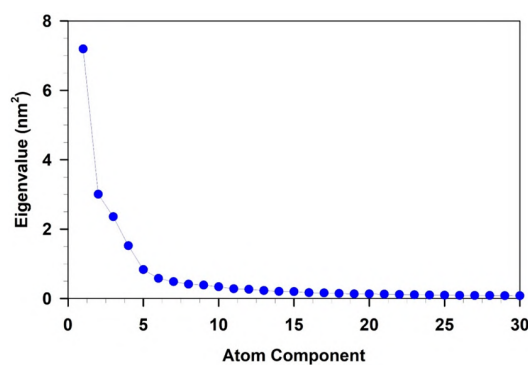
Figure 2. Eigenvalues of the $[Au_{25}(GSH)_{18}]^{-1}$ covariance matrix at RT. Only the first 30 values are shown for the sake of clarity.

fraction, while another 12% results from the second component. Therefore, the further system's conformational analysis was done considering only the corresponding first two eigenvectors to reach a computational affordable description of the physics of this system. Figure 3 shows the projection of the Cartesian coordinates of the nanocluster, at each frame of the MD trajectory, onto the first (proj-1) and second (proj-2) eigendirections of the covariance matrix shown in Figure 2.

The free-energy landscape reveals the presence of several shallow low-energy regions (with ΔG° < 2.5 kJ/mol), suggesting that along the full-MD simulation, the system is very flexible and lies in different stable, short-living, conformational basins. This result is expected considering the size, number, and flexibility of the GSH ligands, which can be involved in various conformational transitions. We can also notice that the low energy states in the conformational space are collected in three groups, with the interconversion being
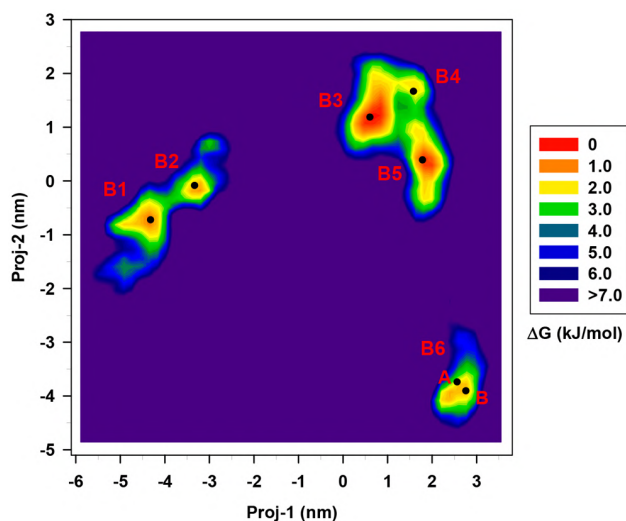
Figure 3. Free energy landscape in the essential plane of $[Au_{25}(GSH)_{18}]^{-1}$ at RT. The energy scale in kJ/mol is reported as a vertical-colored bar on the right. All data with a $\Delta G° > 7.0$ kJ/mol is shown by dark purple. The spatial positions of the 7 extracted most probable conformations are marked with black dots.

fast (barriers below 3 kJ/mol), between B1−B2, and B3−B4−B5, while being slow from one group to another. Such result corroborates again the complexity and high flexibility of the system which can rapidly interconvert from one stable

conformation into another. We extracted one representative conformation per basin except for the region B6 where we found that the root−mean−square deviation (RMSD) of the two structures was higher than our threshold value of 0.17 nm. Energy details of the AuNC conformers are reported in Table S1 of the Supporting Information (SI).

Before proceeding with the investigation of the solvent shell, we extracted additional information on the selected conformers studying the radius of gyration (Rg), the solvent accessible surface area (SASA), and the intramolecular H-bonds of $[Au_{25}(GSH)_{18}]^{-1}$ from the full-MD simulation. All the results are collected in Figure 4.

Fluctuations of the three quantities in Figure 4 (Rg, SASA, and intramolecular H bonds) are all relatively small, suggesting that our system is well equilibrated along the 250 ns MD simulation. This applies particularly to the radius of gyration (Figure 4a), which indicates the system compactness and the preservation of global shape along the MD. Therefore, the initial spherical shape of the nanocluster is conserved through the simulation, and the 7 stable conformations extracted do not display differences in terms of global compactness. A similar result is achieved for the surface area accessible to the solvent (see Figure 4b), where, again, for all the low-energy structures, we found a SASA value around the average one (i.e., 34.6 nm²).

Different results are obtained by analyzing the number of intramolecular H-bonds among the 18 GSH ligands of the nanocluster since some fluctuations are found for the 7 conformers. While for the B2, B4, and B5 conformations, the
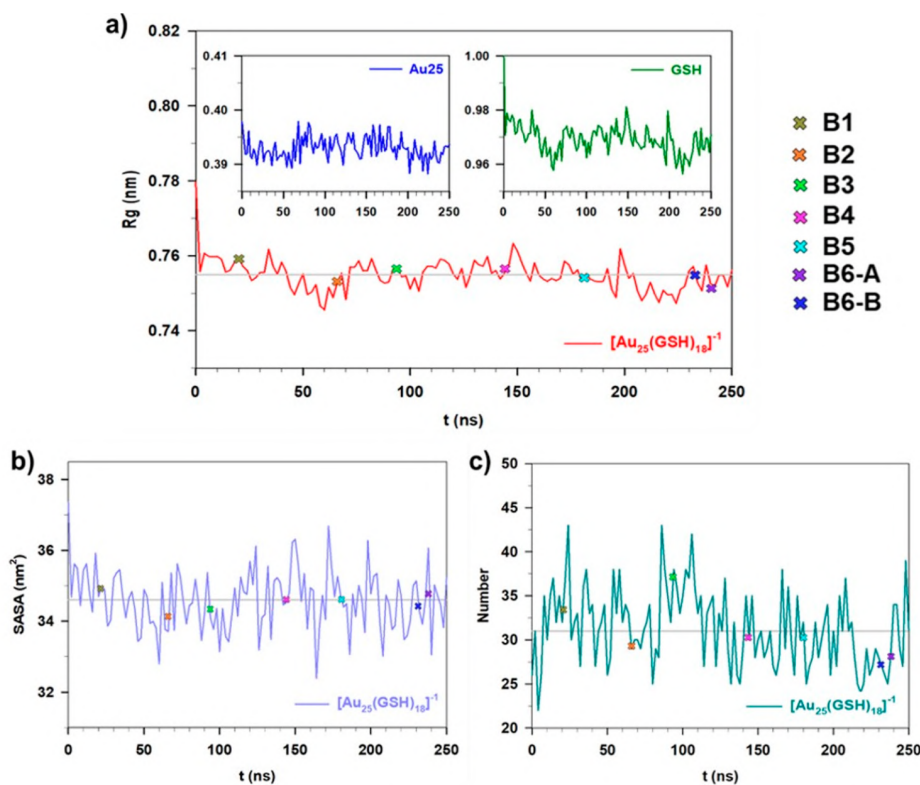
Figure 4. a) Radius of gyration (Rg, in nm), b) solvent accessible surface area (SASA, in nm²), and c) intramolecular H-bonds (a.u.) of $[Au_{25}(GSH)_{18}]^{-1}$ (among the ligands) along the 250 ns full-MD simulation. In the upper panels of a), the Rg of Au_{25} and of the GSH ligands are reported as well. The 7 conformations selected with the previous ED analysis are reported with colored crosses (legend on the upper right region of the figure) in correspondence of their time step in the MD production.

number is around the average value (i.e., 31), a higher number is observed for B1 and B3, in particular for the latter one (i.e., 37), which is also the most stable conformation. An opposite trend is obtained for the structures selected from the B6 region (see Figure 3), where the number of H-bonds is reduced (i.e., 27/28). It is well-known that the presence of H-bonds affects the optical response of the system,[37,38] especially for the highly sensitive ECD technique. Therefore, we expect the intramolecular H-bonds, together with the intermolecular ones (see Figure S1 for a detailed analysis), to have a clear impact on the ECD spectrum.

Constrained-MD Simulations and $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ Clusters Conformational States. The characterization of the solvation shell is, in general, a nontrivial task because of the need to find the actual number of solvent molecules defining the different solvation shells (see the Computational Details and Figures S3 and S4 of the SI) and, most importantly, because of their relatively high mobility. In this study, we have adopted a strategy based on ED which, similarly to what has been described in the previous section, allows us to span the conformational repertoire of a preselected number (126) of solvent molecules with respect to the (frozen) $[Au_{25}(GSH)_{18}]^{-1}$ conformations extracted from Figure 3. The analysis is in principle rigorous, but very difficult in practice, because of the high mobility of the solvent molecules. As a matter of fact, the spectra of the first 30 eigenvalues reported in Figure 5 show this significant increase
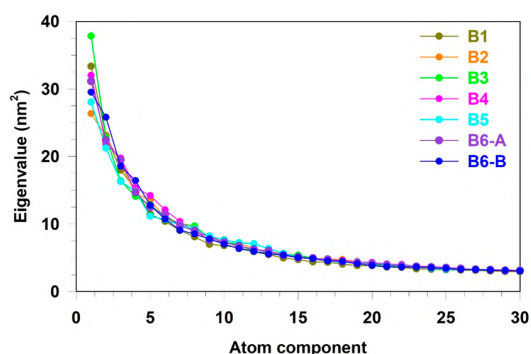
Figure 5. Eigenvalues of the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters covariance matrix obtained from the constrained-MD simulations at RT. Only the first 30 values are shown for the sake of clarity.

in the overall fluctuation (the trace of the covariance matrix) and its spread over a high number of internal degrees of freedom. Despite the reduced weight of the first two eigenvalues, thus some possible limitations, we proceeded with a 2D conformational investigation for computational reasons. The resulting free-energy landscapes, one for each of the 7 $[Au_{25}(GSH)_{18}]^{-1}$ conformations, reported in Figure 6 show a series of important features which deserve some comments.

First, if compared to the analogue conformational space reported in Figure 3, we observe that, as expected, the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ system spans a space much larger than the $[Au_{25}(GSH)_{18}]^{-1}$ one, also showing a higher number of conformational states. These findings obviously reflect the larger number of internal degrees of freedom and the higher mobility of the aqueous solvent. In addition to that, the spots locating the stable $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ appear as very separated onto the conformational space, thus suggesting the

presence of high interconversion barriers. This result is the effect of both the use of a reduced 2D projection subspace and a reflection of the complex interaction between the nanocluster and the solvent. Indeed, several kinds of weak interactions (charge−dipole, dipole−dipole, H-bonds) can define the system, with the water molecules found both in the inner and outer region of $[Au_{25}(GSH)_{18}]^{-1}$. The spatial positions of the selected most probable cluster conformations are marked with black dots.

We then decided to extract one or two conformations among the most probable ones (with $\Delta G^{\circ}_{H_2O} < 1.0$ kJ/mol) from almost opposite regions of each landscape to maximally increase the heterogeneity of our reduced final set of structures. Details of the 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters conformations, which have been subsequently used for the calculation of the UV-CD spectra, are collected in Table 1.

UV-Circular Dichroism (UV-CD) Spectra. The 13 individual UV-CD spectra were calculated, statistically weighted, and summed up to give the final averaged UV-CD spectrum, which was then compared with the experimental data available in the literature[2] and reported in Figure 7.

Figure 7 highlights discrepancies and agreements between the experimental and calculated spectra. In the low energy region, where the signal arises from the ECD of the metal cluster, we observed an excellent agreement in terms of energy and intensity. As discussed in previous studies on this system,[30,34] the achiral metal core shows an optical activity induced by the GSH ligands that disrupt the symmetry. Therefore, the employment of an MD simulation becomes fundamental to capture the metal structure distortion driven by the ligand fluctuations and interactions. In general, the calculated profile matches properly the experimental data up to 3.3 eV; beyond this value, the agreement looks less satisfactory. Indeed, from 4.3 eV onward, the discrepancies between the two spectra become more pronounced. Figure 7 shows that while the experiment presents a maximum positive peak toward 5 eV, the opposite phase is obtained for the calculated ECD spectrum.

In this high energy region, the ECD is more influenced by the solvent, which affects both the geometry and the electronic structure of the chiral species.[38] Hence, some discrepancies can be justified considering that the water effects have been partially included in our work (limited solvation shell) because of the computational cost. However, by investigating the water role with different analyses (see subsection), we were able to point out in detail its strong effect on the AuNC chiral activity, thus confirming the need of its explicit contribution in the ECD calculations.

Some limitations of our ECD spectrum can be also related to an imbalance of the conformational states, thus meaning that some of the structures we extracted are less probable in the experimental sample. This can be ascribed to limitations of the classical MD simulations which result in the inability to reproduce all the thermodynamics properties of the AuNC, thus producing artifacts in the ED analysis. However, upon analyzing the individual ECD spectra (see Figure S1), we noticed that the most probable conformer of each solvation free-energy landscape (Figure 6) shows spectral features in closer agreement with respect to the experimental ones. Moreover, the conformers extracted from the B6 region deviate from the general trends (e.g., H-bonds analysis, see Figures 4c
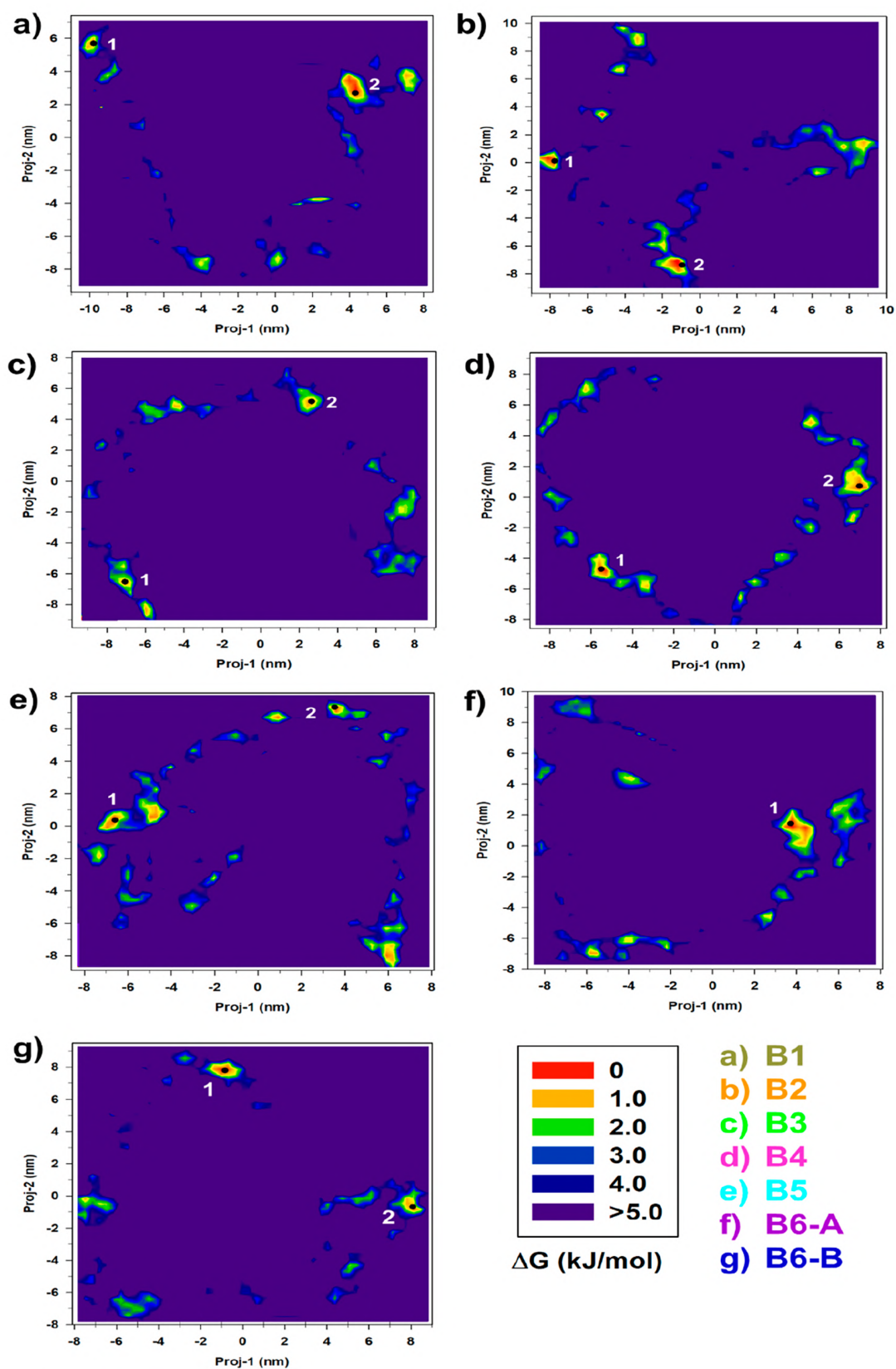
Figure 6. Free energy landscapes in the essential planes of the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters from the constrained-MD simulations at RT. The energy scale in kJ/mol is reported as a vertical-colored bar on the lower right-side of the figure together with the labels of the $[Au_{25}(GSH)_{18}]^{-1}$ conformations used in each constrained-MD (and extracted from the previous full-MD). All data with a $\Delta G° > 5.0$ kJ/mol are shown by dark purple. The spatial positions of the selected most probable cluster conformations are marked with black dots.

Table 1. $\Delta G°$ and Normalized Probability Values of the 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ Conformations Extracted from the ED Analysis of Constrained-MD Simulations

$[Au_{25}GSH_{18}]^{-1}$-$(H_2O)_{126}$ Conformation	$\Delta G°_{TOT}$ (kJ/mol)[a]	$p(j,i)_{norm}$[b]
B1−1	1.61	0.067
B1−2	0.87	0.090
B2−1	2.05	0.057
B2−2	1.45	0.072
B3−1	0.00	0.13
B3−2	0.98	0.087
B4−1	2.21	0.053
B4−2	2.82	0.041
B5−1	0.81	0.093
B5−2	0.34	0.11
B6-A−1	2.03	0.057
B6−B−1	1.40	0.073
B6−B−2	1.51	0.070

$^a\Delta G°_{TOT} = \Delta G°_{[Au25GSH18]^{-1}} + \Delta G°_{H2O}$. $^b\sum p(j,i)_{norm} = 1$.
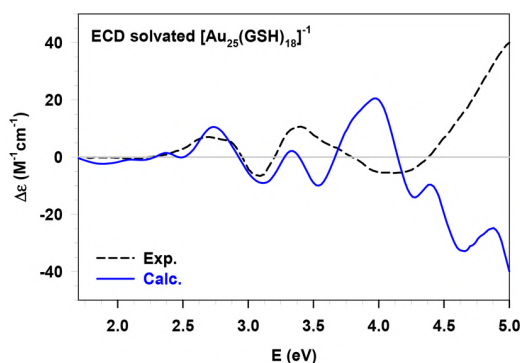
Figure 7. Experimental (Exp., black dashed line, ref 2) and calculated (Calc., solid blue line) ECD spectra of the solvated $[Au_{25}(GSH)_{18}]^{-1}$ nanocluster. The calculated spectrum has been obtained by statistically weighting the 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ cluster conformations extracted with the ED analysis.

Table 2. Renormalized Probabilities Considering Only the Most Probable Conformation from a) Each $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ Conformational Landscape, b) All the Landscapes except for the B6 Regions, c) Only B3, B4, and B5 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ Conformational Spaces

$[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ Conformation	$p(j,i)_{norm}$, Subset a	$p(j,i)_{norm}$, Subset b	$p(j,i)_{norm}$, Subset c
B1−2	0.16	0.20	//
B2−2	0.12	0.16	//
B3−1	0.22	0.28	0.44
B4−1	0.090	0.12	0.18
B5−2	0.19	0.25	0.38
B6-A−1	0.097	//	//
B6−B−1	0.13	//	//
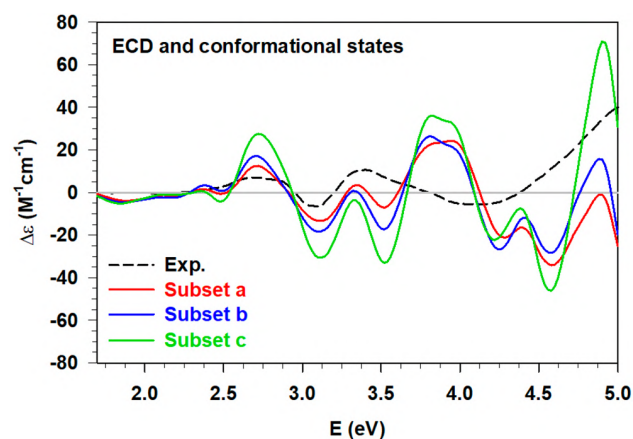
Figure 8. Experimental (Exp., black dashed line, ref 2) ECD of the solvated $[Au_{25}(GSH)_{18}]^{-1}$ in comparison with the calculated statistical ECD obtained considering only the most probable $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ conformations from 1) each conformational landscape (subset a, solid red line), 2) each conformational landscape except for the two B6 spaces (subset b, solid blue line), 3) B3, B4, and B5 conformational landscapes (subset c, solid green line).

and S1) followed by the other stable structures. Starting from these observations, we have built subsets of the original sample of conformers and renormalized their probabilities. Practically, we have considered: a) only the most probable conformer from each $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ free-energy landscape, and b) the same conformers but excluding the B6 regions. For investigating these hypothetical limitations of the MD simulations, we also built a subset c where we focus only on the most probable region of the conformational landscape (basins B3, B4, and B5 in Figure 3) which should be the most reliable of our sampling. All the new probability values reported in Table 2 were used to statistically weight the individual ECD spectra. The three averaged ECD are shown in Figure 8 together with the experimental spectrum.

It is worthy of note that considering a reduced sample of conformations (i.e., subset c), we obtained the correct intensity ratio between the high-energy ECD peaks and a very good agreement with the experimental spectrum. Indeed, while the spectral features of the metal region remain almost unchanged in all the cases, except for the intensity, significant differences are found for the two peaks located around 3.8 and 4.9 eV, respectively. Despite the blueshift of the *middle* maximum value, there is an improvement removing the B6 conformers (subsets b, and c), with the calculated maximum peak and its

shoulder both correctly reproduced. Furthermore, if we consider only the most probable conformers from B3, B4, and B5 (subset c), the metal feature and the high energy maximum peak are in a very good agreement with the experimental ones in terms of energy position and intensity. These results corroborate the hypothesis that along both the full- and constrained-MD simulations there are some imbalances in the statistical weights and certain regions we investigated should be associated to higher free-energy values. However, by adjusting some details of the statistical analysis, we obtained a qualitative agreement between the experimental and calculated ECD, reproducing all the relevant spectral features.

Water Effect on the Chiroptical Properties of $[Au_{25}(GSH)_{18}]^{-1}$. The role of water on the RS-AuNC chirality is clearly shown below comparing the ECD spectra of B3−1, the most probable conformer, in presence and absence of the solvation shell (Figure 9).

Indeed, while the ECD of the solvated B3−1 conformation in Figure 9 (pink line) reproduces quite well the experimental spectrum, several discrepancies are observed when the water molecules are removed (dark green line). Such differences
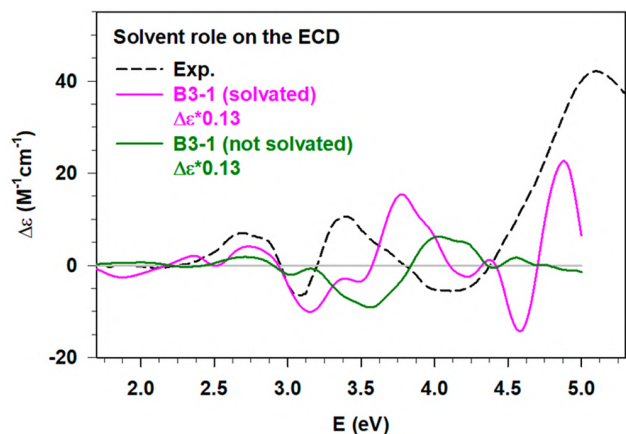
Figure 9. Experimental (Exp., black dashed line, ref 2) ECD in comparison with the ECD calculated for the most probable conformation (B3−1) considering the solvation shell (solid pink line) and removing the water molecules (solid dark green line). The calculated spectra have been multiplied by the statistical weight of B3−1 (0.13) obtained from the ED analysis (see Table 1).

Figure 10. Fragment analysis of the ECD spectrum calculated for the most probable conformer, B3−1. The contributions of the AuNC and H_2O molecular orbitals (initial MOs → final MOs) to the chiroptical properties are pointed out using different colored lines. The H_2O → H_2O transitions are highlighted in the bottom left inset of the figure.

affect even the inner metal region of the AuNC which does not directly interact with the solvent, for instance causing a reduction of intensity of the peak around 2.7 eV but also considering the features below 2.5 eV. In the region between 3 and 4 eV, we find that the experimental positive feature is reproduced much better in the solvated system. The discrepancy becomes much more evident considering the 4−5 eV energy range where the experimental positive peak is reproduced only in the solvated system. These results clearly point out the effect of the polar solvent on the RS-AuNC chiral response, thus the unambiguous need of including, at least, a partial explicit aqueous shell. It should be noted that in Figure 9 the calculated profiles are more structured than the experimental spectrum. This is most likely an effect of our limited sampling of cluster fluctuations corresponding to the experimental temperature. It is known that the experimental absorption spectra are quite sensitive to the temperature.[18]

The role of water has been also assessed by splitting the contributions of the AuNC (fragment 1) and the solvation shell (fragment 2) molecular orbitals (MOs) to the spectral features. Practically, employing a fragment analysis postprocessing tool available in AMS,[39,40] it is possible to understand which MOs are involved in the transitions of the ECD spectrum of the whole system (B3−1 conformer), as shown in Figure 10.

It is worth noting that while in the low energy range the ECD response arises from transitions involving only the AuNC MOs, as expected, the two positive features characterizing the spectrum from *ca.* 3.5 eV onward show significant contributions of the water orbitals. More in detail, the intensity ratios between peaks A/A′ and B/B′ are about 0.5 and 0.4, respectively, thus highlighting the non-negligible contribution of the solvent to the ECD. Furthermore, in the same energy range, we can also observe H_2O → H_2O transitions, although their intensities are much lower compared to those involving the AuNC MOs. This result suggests that the solvation shell itself assumes a chiral arrangement interacting with the nanocluster. The role of the intermolecular interactions (e.g., H-bonds) in relation with the ECD response has been also
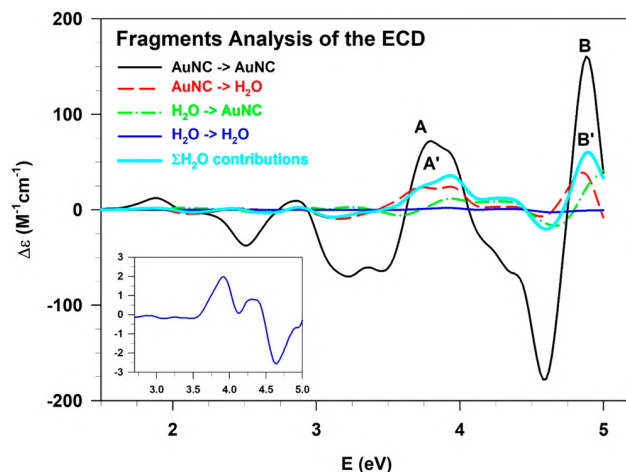
extensively discussed in the SI. In addition to that, the ECD spectrum of only the water shell of the B3−1 conformation has been calculated up to 10 eV, corroborating again the chiroptical features of the solvent itself, particularly strong in the vacuum-UV energy range (see Figure S2). The chirality of $[Au_{25}(GSH)_{18}]^{-1}$, the $(H_2O)_{126}$ shell, and the whole B3−1 system has been also evaluated in terms of Hausdorff chirality measure[41,42] (HCM, see Table S2), obtaining a significant HCM value for both the components of the system.

The solvent effects on the ECD have been already discussed in the literature, as Giovannini et al.[35] highlighted the effects of "chirality transfer" as well as the importance of the formation of hydrogen bonds. Chiroptical properties in solution have been studied also by Del Galdo,[36] while Mancini[43] developed an automated procedure based on evolutionary algorithm.

Therefore, the water shows several effects on the chiroptical properties of RS-AuNCs, such as polarization effects which affect the energy, and intensity of the nanocluster ECD features, but also a direct and significant contribution to the spectral transitions. Furthermore, the achiral solvent shows a low intensity ECD response itself due to the AuNC-H_2O interactions. All these results prove the undeniable need to include explicitly the water molecules for a proper calculation of the RS-AuNCs chiroptical features.

CONCLUSIONS

We have investigated in detail all factors that contribute to the measured ECD spectrum of the chiral, water-soluble $[Au_{25}(GSH)_{18}]^{-1}$ nanocluster combining molecular dynamics simulations, essential dynamics analysis, and state-of-art time-dependent density functional theory calculations. Fundamental aspects such as conformational, weak interactions, and solvent effects have been considered in our approach since they strongly define the physics of this system. Our work showed that the glutathione-protected Au_{25} cluster probes several local basins of different ligand conformations during the molecular dynamics simulations, and statistical sampling of the dynamics becomes an important issue. We also revealed a clear effect of the aqueous environment on the chiroptical properties of RS-

AuNCs, thus the need to treat explicitly the solvent in these calculations. In order to trace out a logical analysis of the relationship between the solvent role and the ECD, we compared the calculated ECD (averaged on the conformations) with the available experimental data, which are in good agreement in the low energy interval (where excitations are dominated by orbitals mostly centered on metal atoms) but show more important deviations in the high energy interval (whose excitations are dominated by orbitals mostly centered on the ligands). In order to understand the origin of such deviation in the high-energy region, we restricted the average only to the most probable conformers, obtaining a definite improvement on the overall agreement between theory and experiment. This clearly indicates that the origin of the disagreement between theory and experiment is due to an imbalanced treatment of the conformational degrees of freedom, and that restricting sampling to the most probable conformers leads to a good agreement with experiment; in particular, it is able to reproduce correctly the ECD maximum at 5 eV. Therefore, we have used this approach of considering only the most probable conformer to focus on the role of the solvent, and compared the spectrum with or without the presence of the solvent explicit molecules. We have thus been able to show that the ECD peak at 5 eV disappears if the solvent is not included, thus unequivocally demonstrating the active role of water molecules to assess the spectral features of the ECD. Analyzing the molecular orbital contributions to the ECD transitions, we were able to corroborate the direct role on the chiroptical features of the solvation shell, which also assumes a chiral arrangement itself around the organic thiol surface. Our work demonstrates a successful strategy to investigate in detail chiral interfaces between metal nanoclusters and their environments, applicable, e.g., to interactions between clusters and biomolecules for chiral sensing.

METHODS

Molecular Dynamics Simulations. The coordinates of $[Au_{25}(GSH)_{18}]^{-1}$ nanocluster were taken from ref 34. MD simulations were carried out with GROMACS software version 2021.4.[44] using a previously published AMBER-compatible force field for nanocluster's Au−S interface[45] and standard parameters[46] for the GSH ligands. The nanocluster was simulated in a cubic box containing TIP3P water[47] and sodium ions to neutralize the system. The volume of the box and the number of solvent molecules were chosen in order to reproduce the concentration used in the experimental reference (i.e., 20 mg/mL).[2] Energy minimizations were performed using the steepest descent algorithm, followed by a 10 ns equilibration in the NVT ensemble at 300 K and 10 ns equilibration in the NPT ensemble at 1 bar. Along these MD simulations, we constrained the internal degrees of freedom of the water molecules with the SETTLE algorithm[48] as well as the position of the heavy atoms of the nanocluster.

Afterward, the position restraints were removed, and a production MD of 250 ns was performed in the NVT ensemble using the velocity-rescale thermostat.[49] The Particle Mesh Ewald method[50] with a cutoff of 1.0 and 0.12 nm grid spacing to treat the long-range electrostatic interactions was used. All the bonds were constrained with the LINCS algorithm[51] to improve the performance. These simulations (here termed full-MD) were subsequently analyzed using the ED and followed by additional MD simulations of 25 ns where we kept frozen the nanocluster in each of the previously extracted conformations. These constrained-MD simulations were performed using the same conditions reported above for the full-MD and analyzed with the ED to map the conformational landscape of the solvation shell closer to each selected nanocluster conformation. Both ED analyses are described in the following section.

Essential Dynamics (ED) Analysis. The features of the ED are widely discussed in the literature[52,53] and the procedure used in this work has been described in detail in ref 38. Therefore, we summarize herein only the relevant aspects of the performed analysis. For the full-MD, we considered the first pair of eigenvectors, obtained by diagonalizing the covariance matrix of the $[Au_{25}(GSH)_{18}]^{-1}$ atomic coordinates, which correspond to the two highest eigenvalues (mean square fluctuations) accounting in this way for most of the nanocluster conformational transitions. In principle, we could consider more than 2 eigendirections,[54] at least the third one, thus conducting the conformational analysis in a $N > 2$ hyperspace. However, such analysis is very demanding and sometimes unnecessary for a still good description of the conformational transitions.[38,55] Therefore, we projected the trajectory along these eigenvectors to obtain the principal components which are the starting point to build a 2D histogram, i.e., to obtain a representation of the 2D conformational landscape of the nanocluster. Each region of this space represents a conformational basin (i) whose probability (P(i)) depends on how frequent is that conformation. Starting from P(i) and P_{ref}, which is the probability of the most probable basin, and assuming as negligible the difference between the volumes (NVT ensemble), we can calculate the standard Gibbs free energy difference ($\Delta G°$) between these basins by using the Boltzmann statistics:

$$\Delta G° = -RT \ln \frac{P(i)}{P_{ref}} \tag{1}$$

The calculation is repeated for each conformational region obtaining the free-energy landscape to investigate. As a result of this analysis, we extracted 7 conformers among the most probable ones ($0 < \Delta G° < 2.5$ kJ/mol) sampling the different low-energy regions of the landscape (see Figure 3) and assuming as identical the conformations with a RMSD < 1.70 Å.[55]

We repeated the ED analysis on the constrained-MD simulations searching for plausible nanoparticle−solvent clusters to be used for the following quantum-chemical calculations. Details on how to perform this analysis are reported in the literature[38,56] and here only summarized. We started constructing the fixed ellipsoid that best describes the nanocluster shape for each of the 7 constrained-MD, followed by the extraction of the N water molecules that show the lower square distances in the just defined ellipsoidal metrics. This step was repeated for each frame of the constrained-MD, obtaining a trajectory of the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_N$ cluster. We selected 126 solvent molecules balancing the number of sites per ligand that can interact directly with the water molecules (i.e., ∼9 sites) and those involved in intramolecular interactions (i.e., ∼2 sites). The latter parameter was estimated considering the average number of intramolecular H-bonds along the full-MD (see Figure 4c). Additional details on the selected water shell are shown in the SI (see Figure S3 and S4). This solvation, even though reduced, allows a compromise between the computational cost of the subsequently quantum-chemical calculations and the accuracy in representing the physics of the real aqueous system. We performed the ED analysis on the 7 obtained $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters to study their conformational landscapes, i.e., the solvation shell conformations. From each space, we extracted one or two cluster conformations associated to a probability p(j,i) and a relative free energy $\Delta G° < 1.0$ kJ/mol calculated by using eq 1. We obtained a final set of 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ cluster conformations (see Figure 1 for an example) with their statistical weights, then used for the quantum-chemical calculations.

Quantum-Chemical Calculations. The geometries of the 13 different conformations of the $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters were optimized at the DFT level[57] by using the Amsterdam Density Functional (ADF) engine of the AMS code.[58] All the optimizations were performed employing the GGA Perdew-Burke-Erzerhof (PBE) exchange-correlation (xc) functional[59] in combination with the GRIMME-D3 dispersion terms,[60] an electron smearing of 0.05 Hartree to help the self-consistent field (SCF) convergence, a basis set of Slater-type orbitals (STO) of triple-ζ plus polarization quality

(TZP), and the Zero Order Regular Approximation (ZORA)[61] for the scalar relativistic effects. Default optimization geometry convergence criteria have been adopted (10^{-5} Hartree for energy, 0.001 Hartree/Å for nuclear gradients and 0.01 Å for the Cartesian step).

The ECD calculations were carried out with the complex polarizability TDDFT (pol-TDDFT) algorithm,[62] also available in the AMS software.[58] The algorithm defines the rotatory strength as

$$R = \frac{3\omega\varepsilon}{2c}\text{Im}[\bar{\beta}] \qquad (2)$$

where $\text{Im}[\bar{\beta}]$ represents the imaginary part of the rotatory strength tensor averaged over all the orientations, ω is the photon energy, c is the speed of the light, and ε is the imaginary part of the photon energy, here taken equals to 0.15 eV. Since the pol-TDDFT calculations are intrinsically broadened by a Lorentzian function, we can directly compare our results with spectra having the half-width half-maximum (HWHM) equal to ε. The ECD calculations were done with the asymptotically corrected LB94 xc functional,[63] the TZP basis-set (appropriately optimized to be used with the pol-TDDFT),[64] and the scalar ZORA approach.[61] We neglected the spin−orbit coupling, since it is important only when isolated transitions are present, like in the lowest part of the spectrum.[65] However, when many transitions are close to each other in energy, like in the present system, the broadening of the spectrum tends to wash out the effects of the spin−orbit coupling.[66] For the most probable conformer, we also performed an ECD calculation with the GGA PBE, comparing the results with both the xc functionals with respect to the experimental spectrum (see Figure S5). So far, we have not tested hybrid functionals (e.g., B3LYP) since they can be included in a pol-TDDFT calculation only employing the hybrid diagonalization approximation (HDA),[67] which is prohibitively expensive for such a large system. Furthermore, for the same most probable structure, we calculated the ECD only on the nanocluster conformation, thus removing the solvent molecules, to evaluate the effect of the explicit inclusion of the solvation shell on the optical response. The solvent impact on this $[Au_{25}(GSH)_{18}]^{-1}$ nanocluster has been already discussed from a different perspective in Rojas-Cervellera et al.'s work[34] and here evaluated in relation with the ECD response (see subsection). The role of the solvent has been also corroborated by performing the fragment analysis available as a postprocessing tool in AMS.[39,40] Considering the NC and the solvation shell as different fragments, it is possible to split the two fragment contributions to the chiro-optical features, thus pointing out if the water orbitals contribute directly to the ECD transitions.

The ECD spectra of all the conformations were finally weighted with their relative $p(j,i)$ values and summed to give the statistically averaged ECD. The overall calculated spectrum was compared with the experimental ECD data[2] to evaluate the quality of the proposed approach.

ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.3c01309.

Free-energy values extracted from ED analysis of $[Au_{25}(GSH)_{18}]^{-1}$ nanocluster, analysis of intermolecular H-bonds of 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters conformations in relation to individual ECD spectra, B3−1 solvation shell ECD spectrum, Hausdorff chirality measure values for B3−1 conformation, additional analysis of selected solvation shell, ECD spectra calculated with different xc functionals of the most probable conformer B3−1 (PDF)

AUTHOR INFORMATION

Corresponding Authors

Mauro Stener − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy;* ⊙ orcid.org/0000-0003-3700-7903; Email: stener@units.it

Hannu Häkkinen − *Department of Physics, Nanoscience Centre, University of Jyväskylä, FI-40014 Jyväskylä, Finland; Department of Chemistry, Nanoscience Centre, University of Jyväskylä, FI-40014 Jyväskylä, Finland;* ⊙ orcid.org/0000-0002-8558-5436; Email: hannu.j.hakkinen@jyu.fi

Authors

Marta Monti − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, 34127 Trieste, Italy*

María Francisca Matus − *Department of Physics, Nanoscience Centre, University of Jyväskylä, FI-40014 Jyväskylä, Finland;* ⊙ orcid.org/0000-0002-4816-531X

Sami Malola − *Department of Physics, Nanoscience Centre, University of Jyväskylä, FI-40014 Jyväskylä, Finland*

Alessandro Fortunelli − *CNR-ICCOM, Consiglio Nazionale delle Ricerche, 56124 Pisa, Italy;* ⊙ orcid.org/0000-0001-5337-4450

Massimiliano Aschi − *Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, 67100 l'Aquila, Italy;* ⊙ orcid.org/0000-0003-2959-0158

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsnano.3c01309

Author Contributions

H.H., M.M., and M.S. conceived the project. M.M. performed the calculations, analyzed the results, and wrote the first draft of the manuscript. M.F.M. and S.M. advised on computational methods, analysis of the results, and on the studied cluster system. M.A. advised on the MD simulations and ED analysis. All authors commented on the final manuscript.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Schaaff, T. G.; Knight, G.; Shafigullin, M. N.; Borkman, R. F.; Whetten, R. L. Isolation and Selected Properties of a 10.4 kDa Gold: Glutathione Cluster Compound. *J. Phys. Chem. B* **1998**, *102*, 10643−10646.

(2) Schaaff, T. G.; Whetten, R. L. Giant Gold-Glutathione Cluster Compounds: Intense Optical Activity in Metal-Based Transitions. *J. Phys. Chem. B* **2000**, *104*, 2630−2641.

(3) Jadzinsky, P. D.; Calero, G.; Ackerson, C. J.; Bushnell, D. A.; Kornberg, R. D. Structure of a Thiol Monolayer-Protected Gold Nanoparticle at 1.1 Å Resolution. *Science* **2007**, *318*, 430−433.

(4) Qian, H.; Eckenhoff, W. T.; Zhu, Y.; Pintauer, T.; Jin, R. Total Structure Determination of Thiolate-Protected Au38 Nanoparticles. *J. Am. Chem. Soc.* **2010**, *132*, 8280−8281.

(5) Heaven, M. W.; Dass, A.; White, P. S.; Holt, K. M.; Murray, R. W. Crystal Structure of the Gold Nanoparticle $[N(C_8H_{17})_4][Au_{25}(SCH_2CH_2Ph)_{18}]$. *J. Am. Chem. Soc.* **2008**, *130*, 3754−3755.

(6) Akola, J.; Walter, M.; Whetten, R. L.; Häkkinen, H.; Grönbeck, H. On the Structure of Thiolate-Protected Au25. *J. Am. Chem. Soc.* **2008**, *130*, 3756−3757.

(7) Zeng, C.; Jin, R. Chiral Gold Nanoclusters: Atomic Level Origins of Chirality. *Chem. Asian J.* **2017**, *12*, 1839−1850.

(8) Zeng, C.; Wu, Z.; Jin, R. Chirality in Gold Nanoclusters. In *Chiral Nanomaterials: Preparation, Properties and Applications*; Tang, Z., Ed.; John Wiley & Sons: Weinheim, 2018; pp 99−119.

(9) Knoppe, S. Chirality in Ligand-Stabilized Metal Clusters. In *Encyclopedia of Interfacial Chemistry: Surface Science and Electrochemistry, Chapter 12982, Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Wandelt, K., Ed.; Elsevier: Amsterdam, 2018; pp 406−416.

(10) Noguez, C.; Garzón, I. L. Optically Active Metal Nanoparticles. *Chem. Soc. Rev.* **2009**, *38*, 757−771.

(11) Gautier, C.; Bürgi, T. Chiral Gold Nanoparticles. *ChemPhysChem* **2009**, *10*, 483−492.

(12) Zhu, M.; Aikens, C. M.; Hollander, F. J.; Schatz, G. C.; Jin, R. Correlating the Crystal Structure of a Thiol-Protected Au25 Cluster and Optical Properties. *J. Am. Chem. Soc.* **2008**, *130*, 5883−5885.

(13) Lopez-Acevedo, O.; Tsunoyama, H.; Tsukuda, T.; Häkkinen, H.; Aikens, C. M. Chirality and Electronic Structure of the Thiolate-Protected Au38 Nanocluster. *J. Am. Chem. Soc.* **2010**, *132*, 8210−8218.

(14) Hidalgo, F.; Noguez, C. How to Control Optical Activity in Organic-Silver Hybrid Nanoparticles. *Nanoscale* **2016**, *8*, 14457−14466.

(15) Shichibu, Y.; Negishi, Y.; Watanabe, T.; Chaki, N. K.; Kawaguchi, H.; Tsukuda, T. Biicosahedral Gold Clusters $[Au_{25}(PPh_3)_{10}(SC_nH_{2n+1})_5Cl_2]^{2+}$ (n = 2−18): A Stepping Stone to Cluster-Assembled Materials. *J. Phys. Chem. C* **2007**, *111*, 7845−7847.

(16) Nobusada, K.; Iwasa, T. Oligomeric Gold Clusters with Vertex-Sharing Bi- and Triicosahedral Structures. *J. Phys. Chem. C* **2007**, *111*, 14279−14282.

(17) Barcaro, G.; Sementa, L.; Fortunelli, A.; Stener, M. Optical Properties of Nanoalloys. *Phys. Chem. Chem. Phys.* **2015**, *17*, 27952−27967.

(18) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Krishnadas, K. R.; Bürgi, T.; Bonacchi, S.; Dainese, T.; Maran, F.; Fortunelli, A.; Stener, M. Predictive Optical Photoabsorption of $Ag_{24}Au(DMBT)_{18}^-$ via Efficient TDDFT Simulations. *J. Chem. Phys.* **2021**, *155*, 084103.

(19) Maye, M. M.; Luo, J.; Han, L.; Kariuki, N. N.; Zhong, C. Synthesis, Processing, Assembly and Activation of Core-Shell Structured Gold Nanoparticle Catalysts. *J. Gold Bull.* **2003**, *36*, 75−82.

(20) Wohltjen, H.; Snow, A. W. Colloidal Metal-Insulator-Metal Ensemble Chemiresistor Sensor. *Anal. Chem.* **1998**, *70*, 2856−2859.

(21) Price, R. C.; Whetten, R. L. Raman Spectroscopy of Benzenethiolates on Nanometer-Scale Gold Clusters. *J. Phys. Chem. B* **2006**, *110*, 22166−22171.

(22) Hainfeld, J. F.; Liu, W. Q.; Barcena, M. Gold−ATP. *J. Struct. Biol.* **1999**, *127*, 120−134.

(23) Gómez-Graña, S.; Fernández-López, C.; Polavarapu, L.; Salmon, J. B.; Leng, J.; Pastoriza-Santos, I.; Pérez-Juste, J. Gold Nanooctahedra with Tunable Size and Microfluidic-Induced 3D Assembly for Highly Uniform SERS-Active Supercrystals. *Chem. Mater.* **2015**, *27*, 8310−8317.

(24) Zheng, M.; Huang, X. Nanoparticles Comprising a Mixed Monolayer for Specific Bindings with Biomolecules. *J. Am. Chem. Soc.* **2004**, *126*, 12047−12054.

(25) Jimenez, V. L.; Georganopoulou, D. G.; White, R. J.; Harper, A. S.; Mills, A. J.; Lee, D.; Murray, R. W. Hexanethiolate Monolayer Protected 38 Gold Atom Cluster. *Langmuir* **2004**, *20*, 6864−6870.

(26) Negishi, Y.; Nobusada, K.; Tsukuda, T. Glutathione-Protected Gold Clusters Revisited: Bridging the Gap Between Gold(I)-Thiolate Complexes and Thiolate-Protected Gold Nanocrystals. *J. Am. Chem. Soc.* **2005**, *127*, 5261−5270.

(27) Tracy, J. B.; Crowe, M. C.; Parker, J. F.; Hampe, O.; Fields-Zinna, C. A.; Dass, A.; Murray, R. W. Electrospray Ionization Mass Spectrometry of Uniform and Mixed Monolayer Nanoparticles: $Au_{25}[S(CH_2)_2Ph]_{18}$ and $Au_{25}[S(CH_2)_2Ph]_{18-x}(SR)_x$. *J. Am. Chem. Soc.* **2007**, *129*, 16209−16215.

(28) Shichibu, Y.; Negishi, Y.; Tsunoyama, H.; Kanehara, M.; Teranishi, T.; Tsukuda, T. Extremely High Stability of Glutathionate Protected Au25 Clusters Against Core Etching. *Small.* **2007**, *3*, 835−839.

(29) Negishi, Y.; Chaki, N. K.; Shichibu, Y.; Whetten, R. L.; Tsukuda, T. Origin of Magic Stability of Thiolated Gold Clusters: A Case Study on $Au_{25}(SC_6H_{13})_{18}$. *J. Am. Chem. Soc.* **2007**, *129*, 11322−11323.

(30) Wu, Z.; Gayathri, C.; Gil, R. R.; Jin, R. Probing the Structure and Charge State of Glutathione-Capped $Au_{25}(SG)_{18}$ Clusters by NMR and Mass Spectrometry. *J. Am. Chem. Soc.* **2009**, *131*, 6535−6542.

(31) Tsukuda, T.; Tsunoyama, H.; Negishi, Y. Systematic Synthesis of Monolayer-Protected Gold Clusters with Well-Defined Chemical Compositions. In *Metal Nanoclusters in Catalysis and Materials Science: The Issue of Size Control*; Corain, B., Schmid, G., Toshima, N., Eds.; Elsevier: Amsterdam, 2008; pp 373−382.

(32) Sánchez-Castillo, A.; Noguez, C.; Garzón, I. L. On the Origin of the Optical Activity Displayed by Chiral-Ligand-Protected Metallic Nanoclusters. *J. Am. Chem. Soc.* **2010**, *132*, 1504−1505.

(33) Chevrier, D. M.; Raich, L.; Rovira, C.; Das, A.; Luo, Z.; Yao, Q.; Chatt, A.; Xie, J.; Jin, R.; Akola, J.; Zhang, P. Molecular-Scale Ligand Effects in Small Gold−Thiolate Nanoclusters. *J. Am. Chem. Soc.* **2018**, *140*, 15430−15436.

(34) Rojas-Cervellera, V.; Rovira, C.; Akola, J. How do Water Solvent and Glutathione Ligands Affect the Structure and Electronic Properties of $Au_{25}(SR)_{18}^-$? *J. Phys. Chem. Lett.* **2015**, *6*, 3859−3865.

(35) Giovannini, T.; Egidi, F.; Cappelli, C. Theory and Algorithms for Chiroptical Properties and Spectroscopies of Aqueous Systems. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864−22879.

(36) Del Galdo, S.; Fusè, M.; Barone, V. The ONIOM/PMM Model for Effective Yet Accurate Simulation of Optical and Chiroptical Spectra in Solution: Camphorquinone in Methanol as a Case Study. *J. Chem. Theory Comput.* **2020**, *16*, 3294−3306.

(37) Kumar, A.; Toal, S. E.; DiGuiseppi, D.; Schweitzer-Stenner, R.; Wong, B. M. Water-Mediated Electronic Structure of Oligopeptides Probed by Their UV Circular Dichroism, Absorption Spectra, and Time-Dependent DFT Calculations. *J. Phys. Chem. B* **2020**, *124*, 2579−2590.

(38) Monti, M.; Stener, M.; Aschi, M. A Computational Approach for Modelling Electronic Circular Dichroism of Solvated Cromophores. *J. Comput. Chem.* **2022**, *43*, 2023−2036.

(39) Sementa, L.; Barcaro, G.; Baseggio, O.; De Vetta, M.; Dass, A.; Aprà, E.; Stener, M.; Fortunelli, A. Ligand-Enhanced Optical Response of Gold Nanomolecules and Its Fragment Projection Analysis: The Case of $Au_{30}(SR)_{18}$. *J. Phys. Chem. C* **2017**, *121*, 10832−10842.

(40) Pelayo, J. J.; Valencia, I.; García, A. P.; Chang, L.; López, M.; Toffoli, D.; Stener, M.; Fortunelli, A.; Garzón, I. L. Chirality in Bare and Ligand-Protected Metal Nanoclusters. *Advances in Physics: X* **2018**, *3*, 1509727.

(41) Buda, A. B.; Mislow, K. A. Hausdorff Chirality Measure. *J. Am. Chem. Soc.* **1992**, *114*, 6006−6012.

(42) Garzón, I. L.; Reyes-Nava, J. A.; Rodríguez-Hernández, J. I.; Sigal, I.; Beltrán, M. R.; Michaelian, K. Chirality in Bare and Passivated Gold Nanoclusters. *Phys. Rev. B* **2002**, *66*, 073403.

(43) Mancini, G.; Fusè, M.; Lazzari, F.; Chandramouli, B.; Barone, V. Unsupervised Search of Low-Lying Conformers with Spectroscopic Accuracy: A Two-Step Algorithm Rooted into the Island Model Evolutionary Algorithm. *J. Chem. Phys.* **2020**, *153*, 124110.

(44) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

(45) Pohjolainen, E.; Chen, X.; Malola, S.; Groenhof, G.; Häkkinen, H. A Unified AMBER-Compatible Molecular Mechanics Force Field for Thiolate-Protected Gold Nanoclusters. *J. Chem. Theory Comput.* **2016**, *12*, 1342−1350.

(46) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(48) Miyamoto, S.; Kollman, P. A. SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithms for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952−962.

(49) Bussi, G.; Parrinello, M. Stochastic Thermostats: Comparison of Local and Global Schemes. *Comput. Phys. Commun.* **2008**, *179*, 26−29.

(50) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(51) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(52) Amadei, A.; Linssen, B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins* **1993**, *17*, 412−425.

(53) Daidone, I.; Amadei, A. Essential Dynamics: Foundation and Applications. *WIREs Comput. Mol. Sci.* **2012**, *2*, 762−770.

(54) Aschi, M.; D'Abramo, M.; Amadei, A. Photoinduced Electron Transfer in a Dichromophoric Peptide: A Numerical Experiment. *Theor. Chem. Acc.* **2016**, *135*, 132−142.

(55) Monti, M.; Brancolini, G.; Coccia, E.; Toffoli, D.; Fortunelli, A.; Corni, S.; Aschi, M.; Stener, M. The Conformational Dynamics of the Ligands Determines the Electronic Circular Dichroism of the Chiral $Au_{38}(SC_2H_4Ph)_{24}$ Cluster. *J. Phys. Chem. Lett.* **2023**, *14*, 1941−1948.

(56) D'Alessandro, M.; Amadei, A.; Stener, M.; Aschi, M. Essential Dynamics for the Study of Microstructures in Liquids. *J. Comput. Chem.* **2015**, *36*, 399−407.

(57) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(58) te Velde, G.; Bickelhaupt, M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931−967.

(59) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(60) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *ab initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(61) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-Component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597−4610.

(62) Baseggio, O.; Fronzoni, G.; Stener, M. A New Time Dependent Density Functional Algorithm for Large Systems and Plasmons in Metal Clusters. *J. Chem. Phys.* **2015**, *143*, 024106.

(63) van Leeuwen, R.; Baerends, E. J. Exchange-Correlation Potential with Correct Asymptotic Behaviour. *Phys. Rev. A* **1994**, *49*, 2421−2431.

(64) Medves, M.; Fronzoni, G.; Stener, M. Optimization of Density Fitting Auxiliary Slater-Type Basis Functions for Time-Dependent Density Functional Theory. *J. Comput. Chem.* **2022**, *43*, 1923−1935.

(65) Jiang, D.; Kühn, M.; Tang, Q.; Weigend, F. Superatomic Orbitals under Spin−Orbit Coupling. *J. Phys. Chem. Lett.* **2014**, *5*, 3286−3289.

(66) Stener, M.; Nardelli, A.; Fronzoni, G. Spin-Orbit Effects in the Photoabsorption of WAu_{12} and $MoAu_{12}$: A Relativistic Time Dependent Density Functional Study. *J. Chem. Phys.* **2008**, *128*, 134307.

(67) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Fortunelli, A.; Stener, M. An Efficient Hybrid Scheme for Time Dependent Density Functional Theory. *J. Chem. Phys.* **2020**, *152*, 184104.

📖 **Recommended by ACS**

Supporting Information for:

What Contributes to the Measured Chiral Optical Response of the Glutathione-Protected Au$_{25}$ Nanocluster?

Marta Monti[1], María Francisca Matus[2], Sami Malola[2], Alessandro Fortunelli[3], Massimiliano Aschi[4], Mauro Stener[1,a,*], Hannu Häkkinen[2,5,b,*]

[1] Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy.

[2] Department of Physics, Nanoscience Centre, University of Jyväskylä, P. O. Box 35, FI-40014, Finland.

[3] CNR-ICCOM, Consiglio Nazionale delle Ricerche, via G. Moruzzi 1, 56124, Pisa, Italy.

[4] Dipartimento di Scienze Fisiche e Chimiche, Università dell'Aquila, Via Vetoio, 67100, l'Aquila, Italy.

[5] Department of Chemistry, Nanoscience Centre, University of Jyväskylä, P. O. Box 35, FI-40014, Finland.

[a]Email: stener@units.it

[b]Email: hannu.j.hakkinen@jyu.fi

Table S1. Projections and $\Delta G°$ values of the 7 $[Au_{25}(GSH)_{18}]^{-1}$ conformations extracted from the ED analysis of the full-MD simulation.

$[Au_{25}(GSH)_{18}]^{-1}$ Conformation	$\Delta G°_{[Au_{25}(GSH)_{18}]^{-1}}$ (kJ/mol)
B1	0.87
B2	1.45
B3	0.00
B4	2.21
B5	0.34
B6-A	2.03
B6-B	1.40

Relation between inter-molecular H-bonds and ECD features

We have investigated the correlation of the ECD spectra of the 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters and the number of H-bonds that each conformation establishes with its corresponding solvent shell. All the results are collected in Figure S1 which confirms how strong the solvent effect is on the nanocluster, as already discussed in this and previous works[Error! Bookmark not defined.]. For instance, we can notice an almost opposite phase between 4.5-5 eV for the B2 (2th row panels), and B4 ECD spectra (4th row panels), even though there is only a difference in the solvation shell conformation. Instead, for the B1 clusters (1st row panels) which have *ca.* the same number of H-bonds, no differences are observed in the spectral features. The B3 conformations (3rd row panels) show a similar number of H-bonds as well, but some differences can be noticed from 4.5 eV onward, mostly due to a blueshift of the B3-2 ECD spectrum. It is worth noting that the average H-bonds value for the B3 constrained-MD simulation is the lowest one (*i.e.,* 117), while they oscillate between 127 (B6-A) and 141 (B2) bonds for the other simulations. This result supports what was found in Section 3.1 (see Figure 4c), *i.e.,* the B3 nanocluster conformation exhibits a higher number of intra-molecular H-bonds with respect to the other selected low-energy conformations. These analyses together with the two B3 ECD spectra suggest that conformations with a larger number of intra-molecular H-bonds produce a better agreement with the experimental spectrum in the high-energy range (maximum peak). The hypothesis is supported also by B2, B4 and B5 conformations where the conformation with fewer inter-molecular H-bonds gives the maximum feature, while the trend is refuted by those extracted from B6-B which show a strong decreasing behavior in this high energy region.

Furthermore, we observed the overlap between the two ECD spectra (lowest panel of Figure S1) even though the solvation shells of B6-B – 1 and – 2 are very different in terms of nanocluster–water interactions. For B6-A – 1 instead, we obtained the expected maximum peak between 4.5-5 eV, although significantly reduced if we consider that the number of inter-molecular H-bonds is quite small (similar to the B3 case). However, we already observed that the nanocluster conformations from B6 deviate from the H-bond trend found for the other low-energy conformations, and similar discrepancies arise now analyzing the inter-molecular interactions.
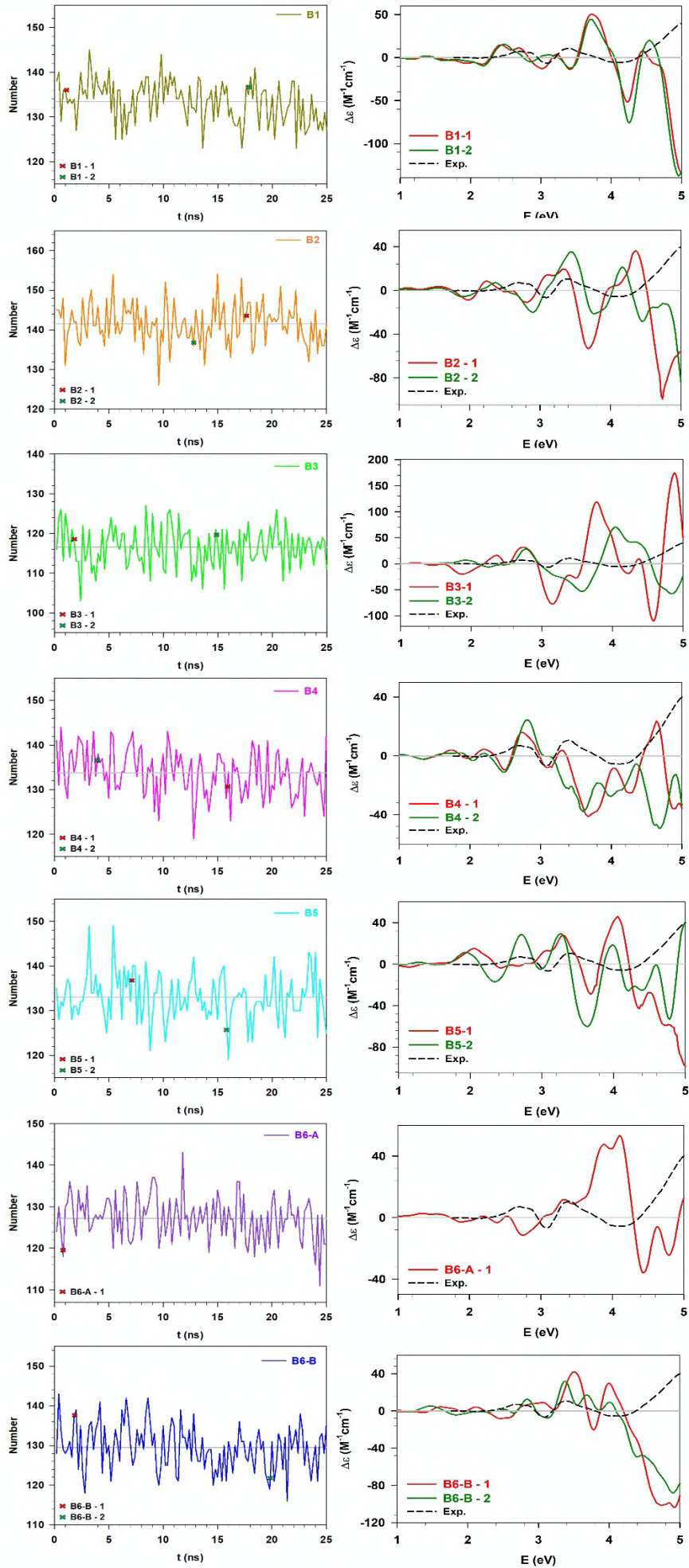
Figure S1. Left panels: inter-molecular H-bonds (a.u.) of $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters along the 25 ns constrained-MD simulations. The clusters conformations extracted (see Table 2) from each constrained-MD are reported with colored crosses in correspondence of their time step in the simulation. The average inter-molecular H-bonds values (<H-bonds>) are reported in the upper left region of each spectrum. Right panels: individual (not weighted) ECD spectra of the 13 conformations.
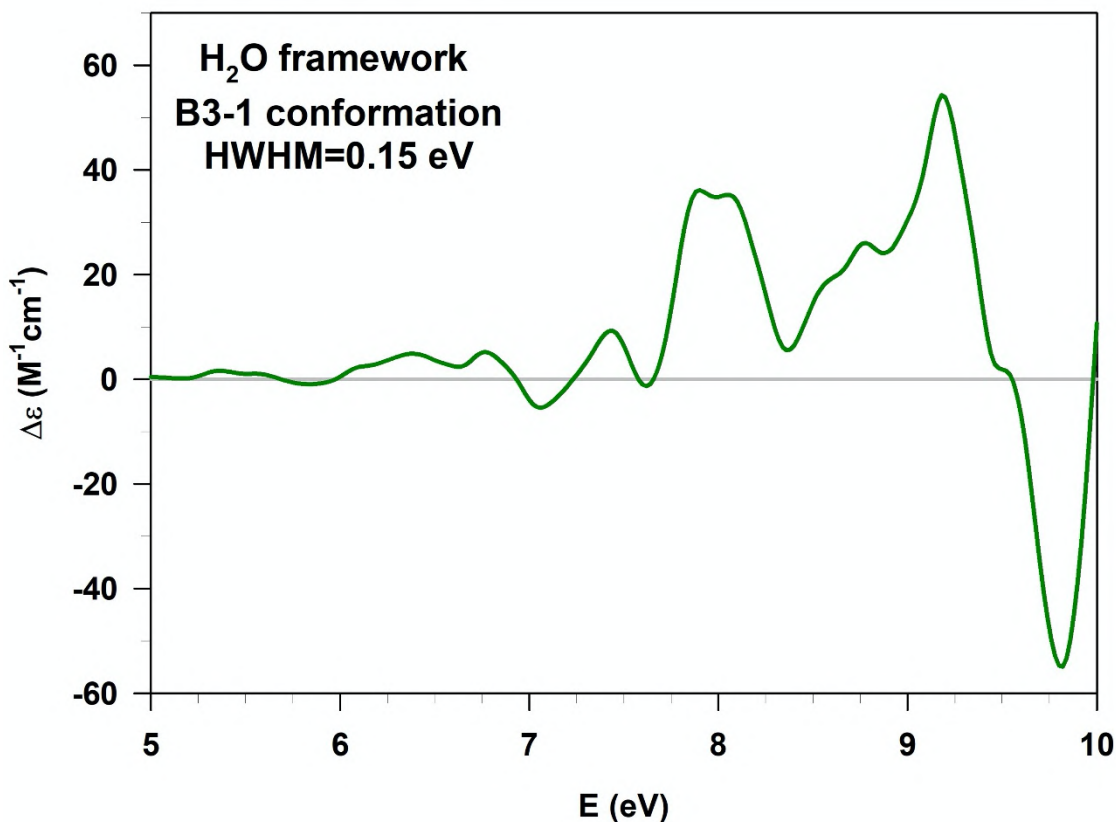
Figure S2. Calculated ECD for the solvation shell of the most probable conformer extracted in this work (B3-1, see Table 1). The spectrum has been calculated employing the LB94 xc functional and a Half Width Half Maximum (HWHM) equals to 0.15 eV.

Table S2. Hausdorff chirality measure (HCM) [1,2] calculated for estimating the chirality of the most probable conformer extracted in this work (B3-1, see Table 1). The HCM values of $[Au_{25}(GSH)_{18}]^{-1}$, the solvation shell $((H_2O)_{126})$, and the whole system $([Au25(GSH)_{18}]^{-1}(H_2O)_{126})$ are reported separately.

$[Au_{25}(GSH)_{18}]^{-1}$	$(H_2O)_{126}$	$[Au_{25}(GSH)_{18}]^{-1}(H_2O)_{126}$
0.127	0.127	0.0868

## Additional details on the selected solvation shell

The glutathione (GSH) ligand shown in Figure S3 presents several sites that can directly interact with the surrounding water molecules. For instance, -CO, -NH, $-NH_3^+$, and $-CO_2^-$ groups can be involved in H-bonds or charge-dipole interactions with the solvent.
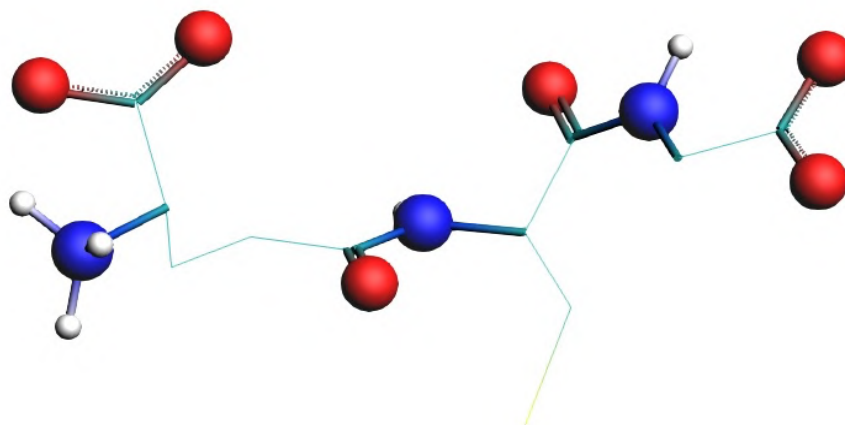
Figure S3. Individual glutathione (GSH) ligand. -CO, -NH, $-NH_3^+$, and $-CO_2^-$ groups are represented as sticks and balls (red for O, cyan for C, blue for N, and white for H), while the remaining C, and S atoms are shown with a wireframe model (H atoms are not shown).

Therefore, we can estimate an average of 9 sites per ligand for possible inter-molecular interactions with the water molecules. Since 18 GSH ligands compose the nanocluster, we must take into account that certain sites will be involved in intra-molecular interactions or will be inaccessible for steric effects. For simplicity, we only considered the first possibility analyzing the full-MD fluctuations of the intra-molecular H-bonds (see Figure 4c) whose average value is around 32, thus an approximatively average of 2 sites per ligand involved in these interactions. Starting from these considerations, we extracted solvation shells of 7x18 water molecules which qualitatively represent the physics of the aqueous system, thus can be used to mimic the real effects of the water solvent on the nanocluster. As we can notice from Figure 1, a qualitative good solvation of the whole ligands is obtained with the water molecules located both in the inner and outer parts of the nanocluster.

We also briefly analyzed the solvation in terms of radial distribution function (g(r)) of the water molecules with respect to the core gold atom (AuC) to estimate the number of shells that must be included for a proper solvation of this system. The results are shown in Figure S4.
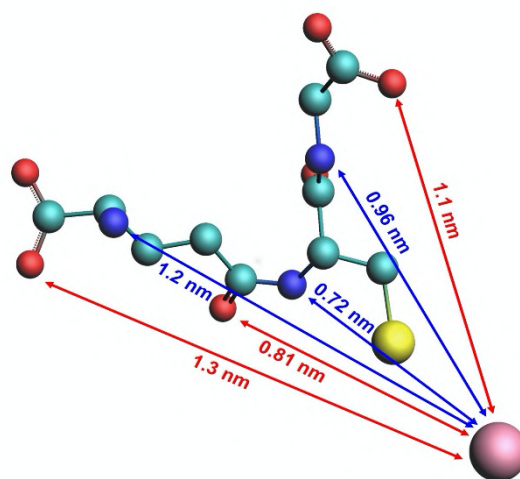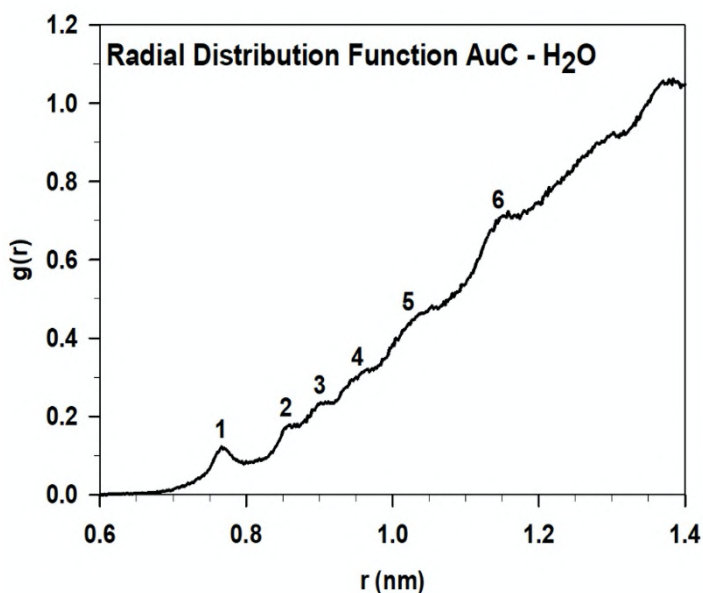
Figure S4. Radial distribution function g(r) of the water molecules with respect to the core gold atom AuC on the left panel. The first 6 solvation shells are highlighted in the figure. Example of the distances of the GSH sites, which can directly interact with the water molecules, with respect to the AuC atom on the right panel. All the distances are reported in nm.

We investigated the g(r) behavior together with the approximate distances of the ligand regions from the core gold atom AuC to evaluate which is the range in nm that should be covered for the solvation. Estimating a distance of ~ 1.2/3 nm with respect to the AuC, at least 6 shells of solvation in terms of g(r) should be considered. It is worthy of note that it is quite common to use only the first shell in order to simulate the solvation of a certain system, which in this case would be a much stronger approximation with respect to the one proposed in this work. Indeed, in that case only the inner part of the ligand would be solvated leaving the chiral C atom of the glutamate residue, which is external, less exposed to the water molecules.
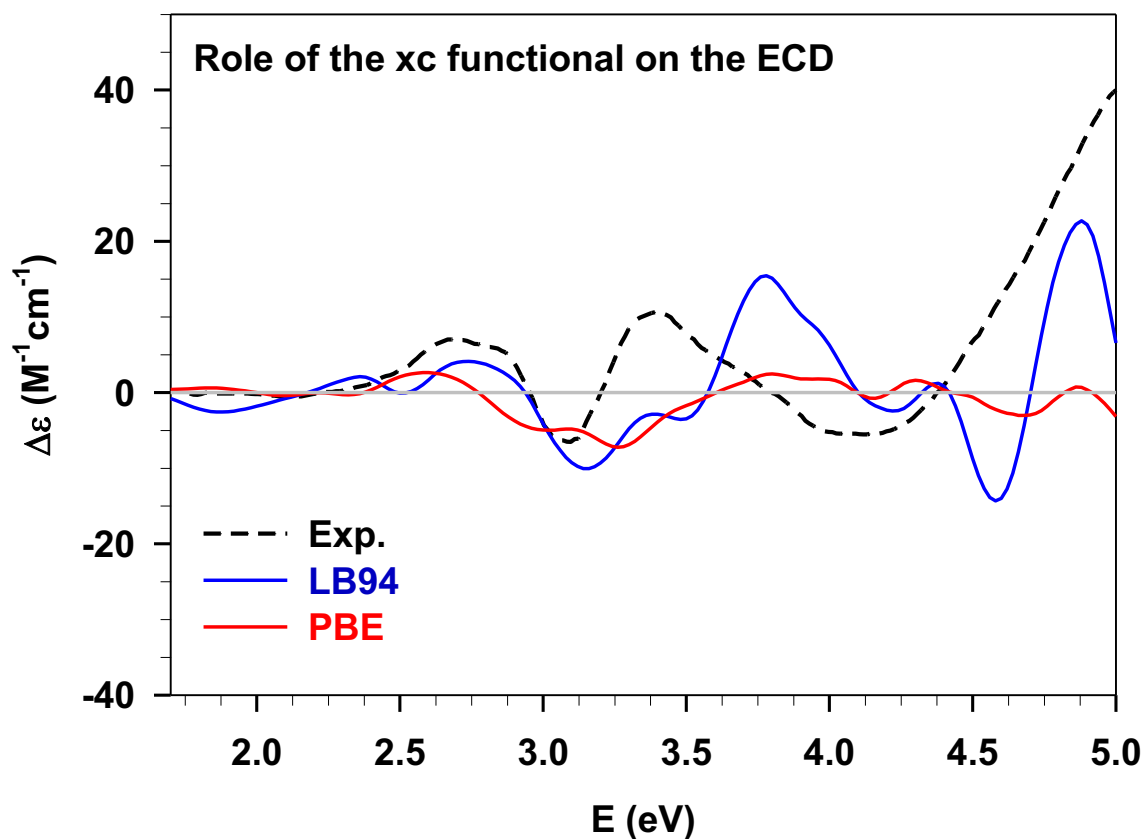
Figure S5. Experimental (Exp., black dashed line) ECD of the solvated $[Au_{25}(GSH)_{18}]^{-1}$ and the ECD calculated for the most probable conformation (B3-1, see Table 2 in the Main Article) employing the LB94 (solid blue line) and PBE (solid red line) xc functionals. The calculated ECD have been multiplied by the statistical weight of this conformation (0.13, see Table 1).

[1] Buda, A. B.; Mislow, K. A Hausdorff Chirality Measure. *J. Am. Chem. Soc.* **1992**, *114*, 6006-6012.

[2] Garzón, I. L.; Reyes-Nava, J. A.; Rodríguez-Hernández, J. I.; Sigal, I.; Beltrán, M. R.; Michaelian, K. Chirality in Bare and Passivated Gold Nanoclusters. *Phys. Rev. B* **2002**, *66*, 073403.

4.1.4 Diverging conformations guide dipeptide self-assembly into crystals or hydrogels

In this combined experimental-theoretical work, the ECD spectra in water were used as a tool to investigate the correlation between the conformational behavior in solution and in the solid-state of heterochiral dipeptides. Working in strict collaboration with Erica Scarel, Prof. Silvia Marchesan, and Prof. Ali Hassanali, we revealed that folded and extended conformers play a key role in the self-assembly of phenylalanine-based dipeptides. More in detail, our computational protocol was applied to several zwitterionic sequences in water containing D-Phe and one aliphatic L-amino acid (*i.e.* valine, norvaline, or isoleucine).

The first pair of dipeptides under investigation was L-Val-D-Phe and D-Phe-L-Val in water at room temperature. The comparison between experimental and calculated ECD revealed a satisfactory qualitative agreement. The only significant discrepancy results in an energy shift of 15 nm for the L-Val-D-Phe calculated spectrum. It was interesting to observe that just switching the position of the two amino acids, different ECD features arise. The differences also result in their conformational landscapes where we found mostly folded conformers for L-Val-D-Phe, while a distribution of folded and extended ones for D-Phe-L-Val. Similar results were obtained for all the other sequences. Therefore, in all the cases where D-Phe is on the C-terminus, plus D-Phe-L-Nva, only folded conformers are found in the lowest energy regions of the free-energy landscape. Remarkably, they strongly resemble the corresponding crystal units of the resolved X-ray structures. This result justifies what was experimentally observed, that is, a very fast crystallization once the system is in solution. Instead, sequences with D-Phe on the N-terminus and bulky aliphatic amino acids (*e.g.* Val, or Ile) are defined by a folded-extended conformational equilibrium where the latter one is predominant. Experiments revealed that these systems self-assemble into gels first, and crystallize over days. Therefore, we can suggest that the extended conformations contribute to the gelation process, while the folded structures, that once again remark the crystal units, push towards the crystallization. Furthermore, it is known from the literature that folded conformations are entalphically favored, whereas extended ones are entropically-driven. Thus, electrostatic, steric, hydrophobic, and thermal effects, as well as stabilization through H-bonds can justify and explain this conformational distribution. These findings were also validated heating the sequences that only crystallize at room temperature to promote the conformational change. Going up to 70 °C, it was noticed the formation of fibrils which are not observable at 25 °C, thus supporting our hypothesis.

ChemComm

COMMUNICATION

Check for updates

Diverging conformations guide dipeptide self-assembly into crystals or hydrogels†

M. Monti, [ID] [a] E. Scarel, [ID] [a] A. Hassanali, [b] M. Stener [ID] *[a] and S. Marchesan [ID] *[a]

The prediction of dipeptide assembly into crystals or gels is challenging. This work reveals the diverging conformational landscape that guides self-organization towards different outcomes. *In silico* and experimental data enabled deciphering of the electronic circular dichroism (ECD) spectra of self-assembling dipeptides to reveal folded or extended conformers as key players.

Self-assembly is a process through which building blocks form a supramolecular structure held together by non-covalent interactions. Several classes of molecules are known to undergo this process. Among them, short peptides have attracted much interest over the past 20 years due to their easy, modular, and low-cost syntheses, and their biocompatibility. Their use is being explored to develop smart, green materials with potential uses that span broadly from medicine to electronics.[1–6]

The most studied self-assembling dipeptide is diphenylalanine (Phe–Phe).[7] This motif was reported to form various nanomorphologies *e.g.*, nanotubes, nanowires, necklaces, and nanovesicles.[7–11] Its success has prompted modifications to derive new functionalities. For instance, the substitution of the N-terminal Phe with an aromatic N-cap,[12,13] or with another aliphatic amino acid (*e.g.*, Leu,[14] or Ile[15,16]) yields self-assembled hydrogels. The chirality of amino acids can have a drastic impact on the dipeptide ability to self-assemble, too.[14,16,17] In the Phe–Val case, only the heterochiral isomers form hydrogels, while the homochiral ones do not. This divergent behaviour was ascribed to the increased hydrophobicity of the heterochiral isomers. However, their ability to give rise to supramolecular water-channels – unlike their L-analogues – could play a role too,[17] as it is a feature shared by other dipeptide gelators, such as heterochiral Phe–Phe[18] and Phe–Ile.[16]

It is clear that several factors affect the ability of peptides to self-assemble. Consequently, the possibility to predict the supramolecular behaviour of an amino-acid sequence using computational approaches has also attracted much attention. Tuttle and co-workers have worked on di- and tri-peptides, screening all the possible combinations for aggregation propensity through coarse-grained molecular dynamics (MD).[19–22] Using this approach, several gelators were discovered. Coarse-grained and all-atomistic MD simulations were also used to model the self-assembly behaviour of classes of peptides and their derivatives.[23–25] Recently, machine-learning techniques were employed as predictive tools for the formation of peptide-based hydrogels.[26,27] To the best of our knowledge, the modelling of the aggregation of heterochiral dipeptides has not been explored thus far.

Over the years, recurrent structural features have been identified for dipeptide hydrogelators, such as hydrophobicity. One of the outstanding unanswered questions pertains to the molecular ingredients that drive self-assembly towards hydrogelation, as opposed to crystallisation. For instance, heterochirality was found to promote hydrogelation of Phe–Val and Phe–Ile.[16,17] However, changing the order of amino acids to Val–Phe and Ile–Phe,[16,17] or substituting Val for its linear isomer norvaline (Nva), led to crystallisation.[28] To shed light on the divergent supramolecular behaviour of such similar structures, here we investigate the conformational landscape of heterochiral (L,D or D,L) dipeptides (Scheme 1) containing D-Phe and one aliphatic L-amino acid, *i.e.* Val, Nva, or Ile.

For this purpose, we adopted a computational approach[29] that allows to calculate accurately the electronic circular dichroism (ECD) spectrum, exploiting its ability to assess which are the main conformers in solution.[29,30] Extraction of the most significant conformations, through a combined MD-Essential Dynamics (ED) approach, enabled the calculation of a statistically averaged ECD spectrum, subsequently used to rationalise the corresponding experimental data.[16,17,28] Details on the procedure are reported in Section S2 of the ESI.† Remarkably, for each dipeptide, the conformers defining the ECD features in

[a] Chem. Pharm. Sc. Dept., University of Trieste, Via L. Giorgieri 1, Trieste 34127, Italy. E-mail: stener@units.it, smarchesan@units.it

[b] The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, Trieste 34151, Italy

† Electronic supplementary information (ESI) available: *In silico* and experimental methods and additional data. See DOI: https://doi.org/10.1039/d3cc02682e
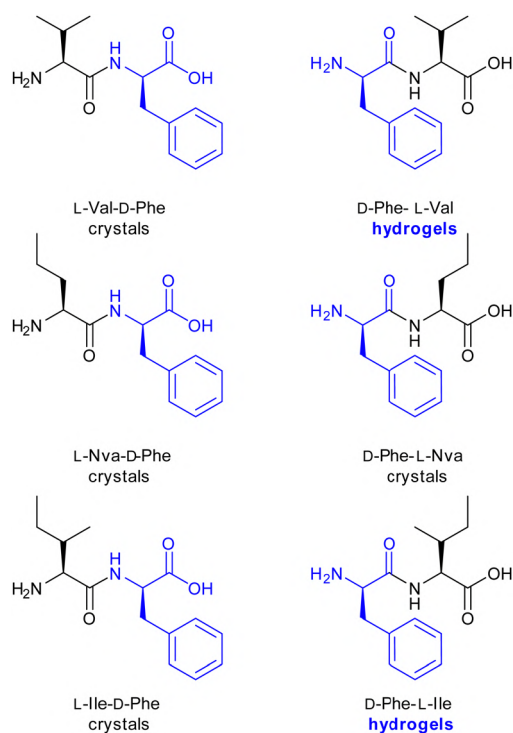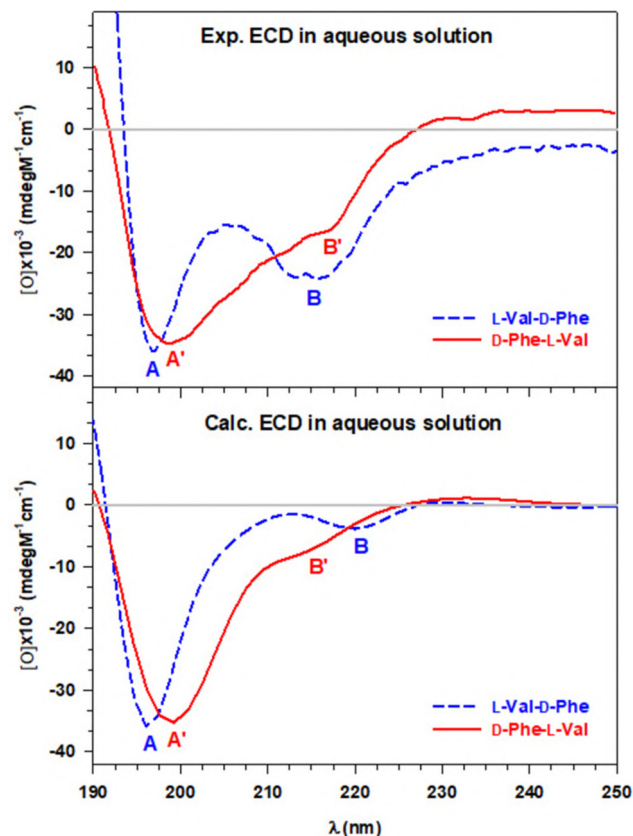
Scheme 1 Heterochiral dipeptide sequences with D-Phe and an L-aliphatic amino acid (*i.e.*, Val, Nva, or Ile) studied in this work to rationalise their supramolecular behaviour towards hydrogels or crystals in phosphate buffer at neutral pH.[16,17,28]

Fig. 1 Comparison between the experimental (Exp., top) and calculated (Calc., bottom) ECD in aqueous solution of L-Val-D-Phe (dashed blue line) and D-Phe-L-Val (solid red line). The Calc. ECD for L-Val-D-Phe and D-Phe-L-Val were shifted by +15 nm, and +0.7 nm, respectively. The intensity scales of the calc. ECD were normalised to match the exp. peak A/A′.

solution can be correlated with those observed in the solid-state (*vide infra*), *i.e.*, in the hydrogels or crystals.[16,17,28]

The first pair of D,L-dipeptides investigated in this work was L-Val-D-Phe and D-Phe-L-Val. The former readily crystallises in phosphate buffer at neutral pH, while the latter gels, then undergoes crystallisation after two days.[17] The comparison between the experimental and calculated ECD response in solution in Fig. 1 shows a satisfactory qualitative agreement, within the limits of the present model and the complexity of the real system. Indeed, the calculations reproduce the two minima (*i.e.*, A/B, and A′/B′) observed for both peptides. Fig. 1 also points out that for L-Val-D-Phe, peaks A (197 nm) and B (216 nm) are well-defined and distinguished. By contrast, for D-Phe-L-Val, A′ (198 nm) and B′ (217 nm) are broader and coalesce together. In addition, peak A is narrower than peak A′ in both the experimental and calculated spectra. The maximum at 205 nm is more defined for L-Val-D-Phe, than for D-Phe-L-Val. Hence, switching the amino acid position, we observe variations in the spectra reflecting the ability of ECD to capture conformational changes, even in such small systems.

In Fig. 2 we report the conformational landscapes obtained from the MD-ED analysis together with some of the most representative conformers ($\Delta G < 1.0$ kJ mol^{-1}). The space of L-Val-D-Phe (Fig. 2a) is defined by two shallow low-energy basins, A and B, which contain only folded conformations. Instead, the landscape of D-Phe-L-Val is more complex, with four stable basins, A–D, that are separated by lower energy barriers, relative to the L-Val-D-Phe case, and a fifth independent low-energy region, E.

The complexity of the space is reflected in the observed conformers, ranging from strongly folded in basin D, towards more or less extended in basins A and E. The same conformational analysis was performed for the other dipeptides in Scheme 1, and their free energy landscapes are collected in Fig. S1 and S2 (ESI†). All the crystallising peptides (*i.e.*, L-Nva-D-Phe, L-Ile-D-Phe, and D-Phe-L-Nva) show two main low-energy regions ($\Delta G \leq 1.0$ kJ mol^{-1}) in Fig. S1 (ESI†), resembling the L-Val-D-Phe conformational space. This is particularly true for L-Nva-D-Phe and L-Ile-D-Phe, while for D-Phe-L-Nva three additional small basins appear in the energy range of 1.5–2.5 kJ mol^{-1}. The conformational landscape of gelling D-Phe-L-Ile is defined by five low-energy regions (Fig. S2, ESI†), in agreement with D-Phe-L-Val.

Remarkably, the low-energy ($\Delta G \leq 2.0$ kJ mol^{-1}) conformations extracted for the crystallising dipeptides (Fig. S3 and Table S1, ESI†) are exclusively folded, closely resembling those of the corresponding crystal-unit structures.[16,17,28] For D-Phe-L-Nva, we found a small number of stretched conformations associated with a ΔG of 2.1–2.5 kJ mol^{-1} (basin E, Fig. S4, ESI†). Notably, this dipeptide does transiently gel in MeCN.[28]

We repeated the same procedure of overlapping the MD and crystal structures for the two hydrogelators[16,17] shown in Fig. S5 (ESI†). Here, both adopt folded conformations, in agreement
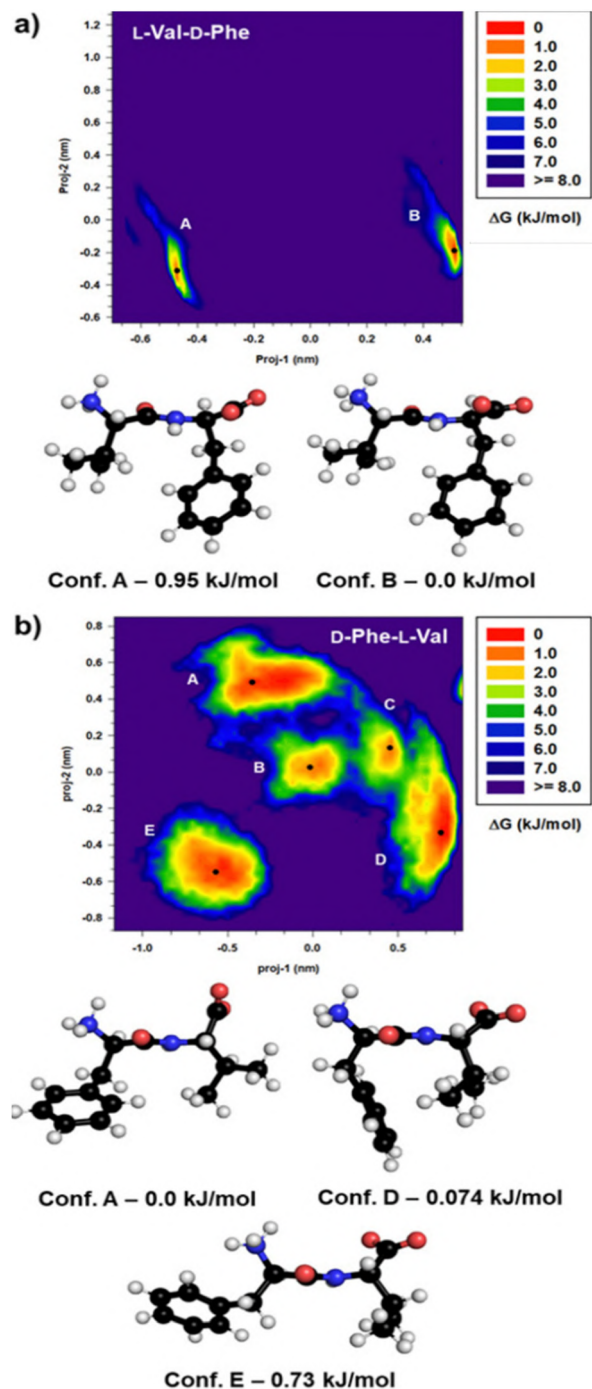
Fig. 2 Relative free energy landscape in the essential plane of L-Val-D-Phe (panel a, top) and D-Phe-L-Val (panel b, bottom) at room temperature (RT). The energy scale, in kJ mol^{-1}, is reported as a vertical colour bar. All of the low-energy regions are labelled with capital letters. The most prevalent conformations with $\Delta G \leq 1.0$ kJ mol^{-1} extracted from these basins are shown in the bottom panel.

with the crystal structure, and extended ones which deviate from them. Interestingly, the conformational equilibrium for the gelators is shifted towards the extended structures.

The clear effect of different conformations on the spectral features is also noteworthy. As shown in Fig. S6 (ESI†) for

D-Phe-L-Val, differences in the energy, intensity, and even sign of the peaks arise because of both the secondary structure, and the solvation shell. ECD confirms to be very sensitive to even small, conformational changes and solvent effects.[29]

This conformational analysis suggests that the systems that crystallise assume folded conformations very quickly in solution similar to those found in the crystal structures. In contrast, hydrogelators are mostly defined by β-strand-like conformations that favour fibrillation.[31] However, the folded conformations, which are still significant, can play a role in proceeding towards the thermodynamically stable phase over days. Therefore, the conformational landscape of a single peptide in solution may indicate its solid-state evolution.

To assess similarities between solution and crystal phases, we also calculated the ECD of models composed of four unit cells for L-Val-D-Phe, D-Phe-L-Val, and L-Nva-D-Phe. All the spectra calculated from the crystals, reported in Fig. S7 and S8 (ESI†), are defined by two minima, thus resembling the spectral features in solution and supporting our analysis above.

Analysis of the conformational distribution is essential to justify and predict why and whether a certain sequence undergoes gelation. Detailed investigations on the conformational preferences of short peptides are available.[32–35] Folded conformers are favoured enthalpically, whereas extended ones are entropically-driven. Thus, electrostatic, steric, hydrophobic, and thermal effects, as well as the stabilisation through intra-, and inter-molecular H-bonds play a role in the folded-extended equilibrium. For the dipeptides in Scheme 1, we suggest that repulsion takes place when Phe is at the C-terminus, contributing to the folding of the aromatic ring, while the $-NH_3^+-\pi$ attractive interaction favours the opening of N-terminal Phe. Cation–π interactions promote peptide gelling.[36] In addition, C-terminal bulky residues with β-branching (e.g., Val, Ile) can destabilise the folded conformer, which is instead favoured with the linear Nva.

To validate these findings, we heated samples to promote a conformational change towards fibrils, as opposed to crystals. We first measured the ECD spectra at 70 °C for L-Nva-D-Phe and L-Val-D-Phe (Fig. S9 and S10, ESI†). Our computational procedure was repeated for the former case, and the calculated ECD is shown in Fig. S9 (ESI†). In both cases, the chiroptical features (peaks B/B′ for L-Nva-D-Phe, and peaks A/A′, and B/B′ for L-Val-D-Phe) are reduced, being indicative of a conformational change, which resulted in a differing solid phase, i.e. fibres (Fig. 3 and Fig. S11, ESI†).

In summary, this work sought a correlation between the conformational behaviour in aqueous solution and the solid-state structures of D,L Phe-based dipeptides. Combining previous[16,17,28] and new experimental data (i.e., ECD spectra, and microscopy data) with a recent computational procedure,[29] we rationalised the chiroptical features that were observed by identifying representative conformers.

The crystallising dipeptides (i.e., L-Val-D-Phe, L-Nva-D-Phe, L-Ile-D-Phe, and D-Phe-L-Nva) are defined by conformational landscapes with two main low-energy regions containing only folded conformers that resemble their crystal units. To assess
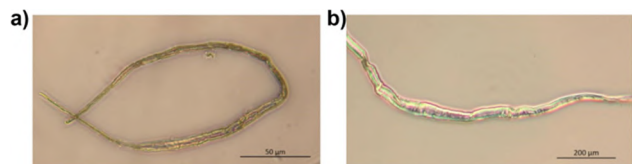
Fig. 3 Optical microscope images of fibres under normal light for L-Nva-D-Phe (panel a, left) and L-Val-D-Phe (panel b, right) after the assembly at 70 °C. Magnification 20× (left) and 10× (right).

the similarities between structures in solution and in the crystal, we calculated ECD spectra on four unit cell models, obtaining an optical response in qualitative agreement with the experimental one in solution, particularly for L-Nva-D-Phe.

In contrast, complex conformational landscapes characterise the peptides that undergo gelation (*i.e.*, D-Phe-L-Val, and D-Phe-L-Ile) followed by crystallisation. In these latter cases, both folded and extended conformers are present, with the latter ones predominating and likely guiding gelation. However, the presence of folded conformers in solution could be responsible for the observed transition from gel to crystal over days, as confirmed by the structural similarity between the calculated folded conformers and those found in the crystal by XRD. Finally, heating drove a conformational change from folded to extended for crystallising dipeptides, as confirmed by ECD and fibrillation (Fig. 3).

The agreement between experimental and *in silico* data validates the utility of gaining key insights about dipeptide assembly in the solid state by investigating the conformational preferences in solution. Extension of this approach to other peptides could offer a tool to predict crystallisation or gelation based simply on their ECD response, potentially enabling more rapid progress in the area of smart biodegradable materials.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 H. Wang, Z. Feng and B. Xu, *Angew. Chem., Int. Ed.*, 2019, **58**, 10423.
2 E. Radvar and H. S. Azevedo, *Macromol. Biosci.*, 2019, **19**, e1800221.
3 G. Uzunalli and M. O. Guler, *Ther. Delivery*, 2020, **11**, 193.
4 M. G. Herrera and V. I. Dodero, *Biophys. Rev.*, 2021, **13**, 1147.
5 D. Giuri, P. Ravarino and C. Tomasini, *Org. Biomol. Chem.*, 2021, **19**, 4622.
6 M. Amit, S. Yuran, E. Gazit, M. Reches and N. Ashkenasy, *Adv. Mater.*, 2018, **30**, 1707083.
7 M. Reches and E. Gazit, *Science*, 2003, **300**, 625.
8 X. Yan, Q. He, K. Wang, L. Duan, Y. Cui and J. Li, *Angew. Chem., Int. Ed.*, 2007, **46**, 2431.
9 J. Ryu and C. B. Park, *Angew. Chem., Int. Ed.*, 2009, **48**, 4820.
10 S. Yuran, Y. Razvag and M. Reches, *ACS Nano*, 2012, **6**, 9559.
11 A. N. Rissanou, E. Georgilis, E. Kasotakis, A. Mitraki and V. Harmandaris, *J. Phys. Chem. B*, 2013, **117**, 3962.
12 A. M. Garcia, R. Lavendomme, S. Kralj, M. Kurbasic, O. Bellotto, M. C. Cringoli, S. Semeraro, A. Bandiera, R. De Zorzi and S. Marchesan, *Chem. – Eur. J.*, 2020, **26**, 1880.
13 S. Roy and A. Banerjee, *Soft Matter*, 2011, **7**, 5300.
14 O. Bellotto, S. Kralj, R. De Zorzi, S. Geremia and S. Marchesan, *Soft Matter*, 2020, **16**, 10151.
15 N. S. de Groot, T. Parella, F. X. Aviles, J. Vendrell and S. Ventura, *Biophys. J.*, 2007, **92**, 1732.
16 O. Bellotto, S. Kralj, M. Melchionna, P. Pengo, M. Kisovec, M. Podobnik, R. De Zorzi and S. Marchesan, *ChemBioChem*, 2022, **23**, e202100518.
17 O. Bellotto, G. Pierri, P. Rozhin, M. Polentarutti, S. Kralj, P. D'Andrea, C. Tedesco and S. Marchesan, *Org. Biomol. Chem.*, 2022, **20**, 6211.
18 S. Kralj, O. Bellotto, E. Parisi, A. M. Garcia, D. Iglesias, S. Semeraro, C. Deganutti, P. D'Andrea, A. V. Vargiu, S. Geremia, R. De Zorzi and S. Marchesan, *ACS Nano*, 2020, **14**, 16951.
19 P. W. J. M. Frederix, R. V. Ulijn, N. T. Hunt and T. Tuttle, *J. Phys. Chem. Lett.*, 2011, **2**, 2380.
20 P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn and T. Tuttle, *Nat. Chem.*, 2015, **7**, 30.
21 A. Lampel, R. V. Ulijn and T. Tuttle, *Chem. Soc. Rev.*, 2018, **47**, 3737.
22 A. van Teijlingen, M. C. Smith and T. Tuttle, *Acc. Chem. Res.*, 2023, **56**, 644.
23 O. S. Lee, S. I. Stupp and G. C. Schatz, *J. Am. Chem. Soc.*, 2011, **133**, 3677.
24 O. S. Lee, V. Cho and G. C. Schatz, *Nano Lett.*, 2012, **12**, 4907.
25 P. Divanach, E. Fanouraki, A. Mitraki, V. Harmandaris and A. N. Rissanou, *J. Phys. Chem. B*, 2023, **127**, 4208.
26 F. Li, J. Han, T. Cao and L. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 11259.
27 R. Batra, T. D. Loeffler, H. Chan, S. Srinivasan, H. Cui, I. V. Korendovych, V. Nanda, L. C. Palmer, L. A. Solomon, H. C. Fry and S. K. R. S. Sankaranarayanan, *Nat. Chem.*, 2023, **14**, 1427.
28 E. Scarel, G. Pierri, P. Rozhin, S. Adorinni, M. Polentarutti, C. Tedesco and S. Marchesan, *Chemistry*, 2022, **4**, 1417.
29 M. Monti, M. Stener and M. Aschi, *J. Comput. Chem.*, 2022, **43**, 1997.
30 N. Berova, L. Di Bari and G. Pescitelli, *Chem. Soc. Rev.*, 2007, **36**, 914.
31 K. Sivanesam and N. Andersen, *Arch. Biochem. Biophys.*, 2019, **664**, 51.
32 W. Chin, F. Piuzzi, I. Dimicoli and M. Mons, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1033.
33 Y. Loquais, E. Gloaguen, S. Habka, V. Vaquero-Vara, V. Brenner, B. Tardivel and M. Mons, *J. Phys. Chem. A*, 2015, **119**, 5932.
34 W. Ji, C. Yuan, P. Chakraborty, S. Gilead, X. Yan and E. Gazit, *Commun. Chem.*, 2019, **2**, 65.
35 T. Vermeyen and C. Merten, *Phys. Chem. Chem. Phys.*, 2020, **22**, 15640.
36 C. C. Cole, M. Misiura, S. A. H. Hulgan, C. M. Peterson, J. W. Williams III, A. B. Kolomeisky and J. D. Hartgerink, *Biomacromolecules*, 2022, **23**, 4645.

Electronic Supplementary Information for:

"Diverging conformations guide dipeptide self-assembly into crystals or hydrogels"

M. Monti[1], E. Scarel[1], A. Hassanali[2], M. Stener[1], S. Marchesan[1]

[1] Chem. Pharm. Sc. Dept., University of Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy

[2] The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy

Table of contents

1. Experimental part

All the compounds investigated in this work were previously published.[1,2,3] Their purity and molecular characterisation were established with different techniques, including LC-MS, ^{1}H NMR and ^{13}C NMR. In particular, LC-MS traces show a single sharp chromatographic peak for each compound and the corresponding mass spectra in positive and negative ion modes are in agreement with the structural elucidation obtained with NMR spectra, corresponding to the desired product.

1.1 Circular dichroism:

L-Val-D-Phe and L-Nva-D-Phe were dissolved in milli-Q-water (resistivity > 18 M Ω cm) at the desired concentration and the pH was corrected with a small amount of NaOH 1 M to reach pH 7.0. All samples were loaded into a 0.1 mm quartz cuvette and sealed with a thin layer of grease to prevent the solvent evaporation during the experiment. All spectra were acquired on a Jasco j-815 spectropolarimeter at 70 °C (Peltier) with the following parameters: data pitch 1 nm, scan speed 100 nm/min, D.I.T 1 s. For every sample were collected 20 spectra in order to evaluate the consistency of the signal, check the thermal equilibrium and the possible leaking of solvent due to evaporation. Then all the spectra were merged to increase the signal/noise ratio.

1.2 Fibres solutions:

L-Val-D-Phe was dissolved in PBS buffer at 80 mM following the procedure reported in literature.[1] The solution was then placed in an oil bath at 70 °C for 15 minutes and then was left cool down at room temperature.

L-Nva-D-Phe formed fibres after the CD experiment inside the cuvette, therefore the pictures were acquired on that sample.

1.3 Optical microscope:

All pictures were acquired on a ZEISS Primovert inverted microscope with Axiocam 208 colour. Few microliters of the fibre solution were placed on a glass slides and then investigated under microscope. Pictures were collected with normal light and with polarized light in order to highlight the chirality of the object. For L-Val-D-Phe some pictures with fibres and crystalline material were acquired. In this case the drop was left dry on the glass slide in order to allow the crystallisation of the remaining peptide in solution.

2. Computational procedure

Here we describe the computational set-up employed to study the conformational landscape of the dipeptide series and to simulate the ECD spectra for L-Val-D-Phe, D-Phe-L-Val, and L-Nva-D-Phe. Further details on this procedure have been discussed in previous works.[4,5]

All the molecular Dynamics (MD) simulations were performed with the GROMACS package[6,7] version 2022.3 and the OPLS-AA force field.[8,9] The water solvent was described with the TIP3P model.[10] All the simulations were carried out in the NVT ensemble, keeping the temperature constant with the velocity-rescale thermostat.[11] The LINCS algorithm[12] was used to constrain all the bonds for improving the performance. The long-range electrostatic interactions were treated with the Particle Mesh Ewald method,[13] using a cut-off of 1.0 nm and a 0.12 nm grid spacing. Each dipeptide has been initially inserted in a cubic box and solvated. The volume and number of water molecules were chosen in order to reproduce the concentrations used for the experimental ECD measurements in solution (*i.e.,* 1-5 mM).[1,2] On these initial boxes, energy minimisations were performed with the steepest descent algorithm. Afterwards, each box was adjusted to reproduce the correct density (in infinite dilution in water) at the temperature of interest (*i.e.,* 25, or 70°C), and a pressure of 1.0 bar. In detail, the solute-solvent volume was modified until its average pressure reached that measured on a box of pure water, with the same number of molecules, previously simulated in the NVT ensemble at the experimental density of pure water at the selected temperature: 997.1 g/L (25°C), 977.6 g/L (70°C). For each system, the equilibration was followed by a production MD (hereafter termed full-MD) of 100 ns.

The full-MD simulations have been then analysed with the essential dynamics (ED), whose features are extensively described in the literature,[14,15] and here only summarised. As a first step, we built and diagonalised the covariance matrix of the dipeptide atomic coordinates, thus obtaining a set of eigenvectors and corresponding eigenvalues. Such eigenvectors correspond to the eigendirections along which the system undergoes its internal motion, while the associated eigenvalues represent the actual values of these mean square fluctuations. As a matter of fact, the first eigenvectors, which correspond to the highest eigenvalues, represent the directions essential to account for most of the conformational transitions of the system. Herein, we limited the number of essential eigenvectors to the first pair as a compromise between accuracy and computational cost. For each dipeptide, we projected the Cartesian coordinates of the MD trajectory onto the two eigendirections, obtaining the so-called Principal Components (PCs) which allow us to reduce and simplify the dimension and complexity of the conformational landscape. In practice, the PCs are visualised building a 2D histogram where each square represents a conformational basin i^{th}, whose probability $P(i)$ depends on how frequently that square is spanned (how probable is that conformation). Considering P_{ref} (=1) as the probability of the most probable region, and a negligible difference between the partial molar volumes (MD in the NVT ensemble), the (standard) Gibbs free energy difference ($\Delta G°$) between these basins may be estimated with the Boltzmann statistics:

$$\Delta G^{\circ} = -RT \ln \frac{P(i)}{P_{ref}} \tag{1}$$

Repeating the calculation for each conformational basin (square), we obtain the relative free energy landscapes shown in Fig. S1-S2, and in the main text (Fig. 2). The analysis of these conformational spaces is focused only on the regions associated to a $\Delta G^{\circ} < k_B T$, where the most significant conformers are found. Therefore, we extracted one or two structures from each low-energy basin as representative of the corresponding region.

We then performed a second set of 15 ns MD simulations (hereafter termed constrained-MD), with the same box and protocol discussed above, keeping frozen the conformations just extracted in order to investigate the conformational role of the solvent. All the constrained-MD trajectories were analysed again with the ED analysis[16] for locating the most plausible peptide-solvent clusters, then used for further analyses and the quantum-chemical calculations. Such clusters, containing 50 water molecules, were built employing a procedure well described in previous works[4,16] and here only briefly summarised. For each constrained-MD, we first constructed the fixed ellipsoid which best describes the frozen dipeptide geometrical shape. Afterwards, we selected the N (=50) water molecules showing the lowest square distances in the so-defined ellipsoidal metrics. This step is repeated at each frame of the constrained-MD, thus obtaining a trajectory for the selected peptide-$(H_2O)_{50}$ cluster. All the trajectories were investigated with ED again, providing us with the 2D conformational landscapes of the peptide-solvent clusters. Once again, we focused our attention on the most probable regions of the space, extracting one or two cluster conformations associated to the higher probability $p(j,i)$ values (lower ΔG°, Eq. 1). Therefore, multiplying $P(i)$ and $p(j,i)$, we obtain a final probability $P_{TOT}(k)$ for each k peptide-$(H_2O)_{50}$ cluster. The P_{TOT} values of the extracted clusters were then normalised. Therefore, considering for each dipeptide case, a certain number N of selected conformations, the total sum (P_{ALL}) of the corresponding P_{TOT} values will be:

$$P_{ALL} = \sum_{k=1}^{N} P_{TOT}(k) = 1 \tag{2}$$

The full-, and constrained-MD, together with the ED analysis, were performed for all the heterochiral dipeptides studied in this work (*i.e.,* L-Val-D-Phe, D-Phe-L-Val, L-Nva-D-Phe, D-Phe-L-Nva, L-Ile-D-Phe, and D-Phe-L-Ile).

For L-Val-D-Phe (25°C), D-Phe-L-Val (25°C), and L-Nva-D-Phe (25°C and 70°C) we also carried out the quantum-chemical calculations. More in detail, we first optimised the geometry of the selected N cluster conformations at the Density Functional Theory (DFT)[17] level by using the Amsterdam Density Functional (ADF) engine of the AMS code.[18] A set of 4, 8, 5, and 4 dipeptide-$(H_2O)_{50}$ clusters were considered for L-Val-D-Phe (25°C), D-Phe-L-Val (25°C), L-Nva-D-Phe (25°C),

and L-Nva-D-Phe (70°C), respectively. In order to maintain the features of the extracted conformations, we constrained the internal degrees of freedom corresponding to the semi-classical motions (torsional angles). For all the optimisations, a Slater-type orbital set of triple-zeta quality was used. In order to describe properly the physics of the system, *i.e.,* intra-, inter-molecular H-bonds and charge-transfer excitations, we employed the range-separated exchange-correlation (xc) functional wB97X-D which contains the dispersion and non-local exchange asymptotical corrections.[19]

Finally, we calculated the ECD spectra for the selected conformations at the Time-Dependent DFT (TDDFT) level by using the Casida approach[20] (Davidson algorithm). The same basis set and xc functional mentioned above were used. The lowest 100 excited-states were analysed to cover the experimental energy range.[1,2] Furthermore, the calculated spectra were reported in $[\theta] \times 10^{-3}$ (mdegM^{-1}cm^{-1}) units for an easy comparison with the experimental measurements. Since the calculated spectra are obtained as discrete lines in terms of rotatory strength (R_{0k}, in cgs units), we used the following expressions to convert them in $\Delta\varepsilon$ units[21] first, and in $[\theta] \times 10^{-3}$ then:

$$\Delta\varepsilon(E) = \frac{1}{2.297 \times 10^{-39}} \frac{1}{\sigma\sqrt{\pi}} \sum_k R_{0k} E_{0k} e^{-\left(\frac{E-E_{0k}}{\sigma}\right)^2} \tag{3}$$

$$[\theta] \times 10^{-3} = \Delta\varepsilon \times 3.298 \tag{4}$$

where σ is related to the Half Width at Half Maximum (HWHM) of the Gaussian functions used to convolute the calculated spectra:

$$\sigma = \frac{HWHM}{\sqrt{\ln 2}} \tag{5}$$

In this work, we used HWHM = 0.15 eV to suitably match the experimental features. Each spectrum was weighted by the corresponding normalised $P_{TOT}(k)$ (k=1,N, see Eq. 2) and summed up to give the final statistically averaged ECD. They were then compared with the experimental references[1,2] for a quality evaluation.

To qualitatively assess the differences between solution and crystal phases, we decided to compare the spectra in solution with those calculated on crystal models (see Fig. S7-S8). Starting from the X-ray crystal structures available for these systems,[1,2] we built four-units models as a compromise between a suitable representation of the crystal molecular orientations and the limitations of the Casida scheme.[22] Although different orientations of the units can be considered, we focused only on the models where all the four units are directly involved in the formation of the water channels that are observed for these crystals. The parameters of the TDDFT calculations are those mentioned above for the calculation in solution.

All the molecular visualisations were realised using the PyMOL software.[23]

3. Figures and Tables
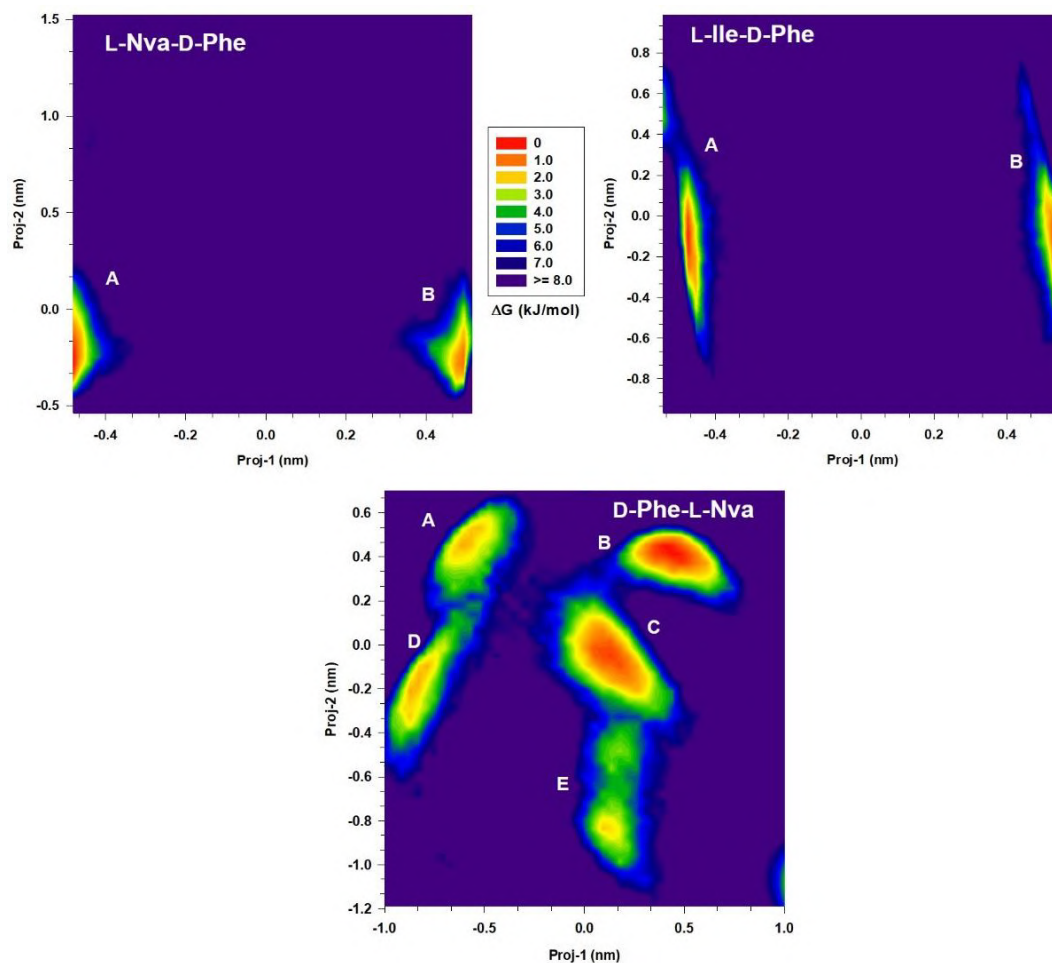
Figure S1. Relative free energy landscapes in the essential plane of L-Nva-D-Phe (top left), L-Ile-D-Phe (top right), and D-Phe-L-Nva (bottom) at RT. The energy scale, in kJ/mol, is reported as a vertical-coloured bar on the top left region. All the low-energy regions, within which the most probable conformations lie, have been labelled with capital letters.
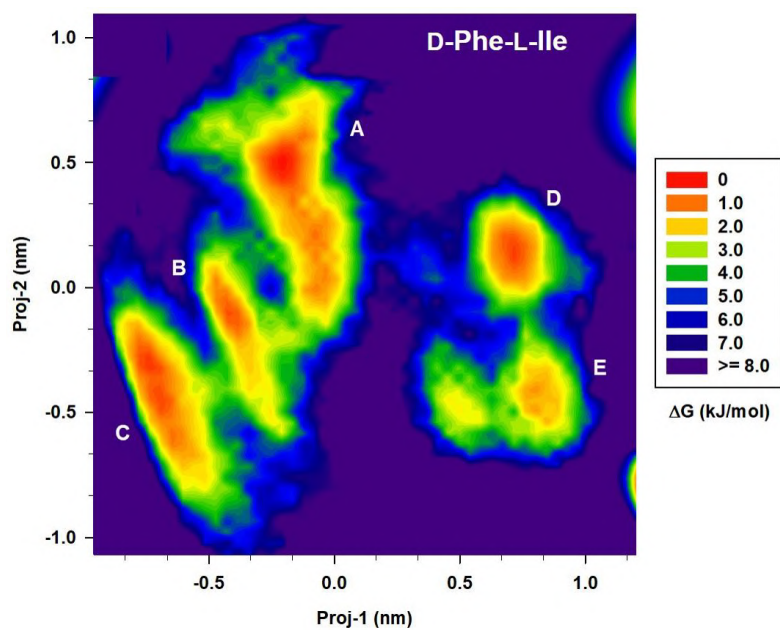
Figure S2. Relative free energy landscape in the essential plane of D-Phe-L-Ile at RT. The energy scale, in kJ/mol, is reported as a vertical-coloured bar on the left. All the low-energy regions, within which the most probable conformations lie, have been labelled with capital letters.
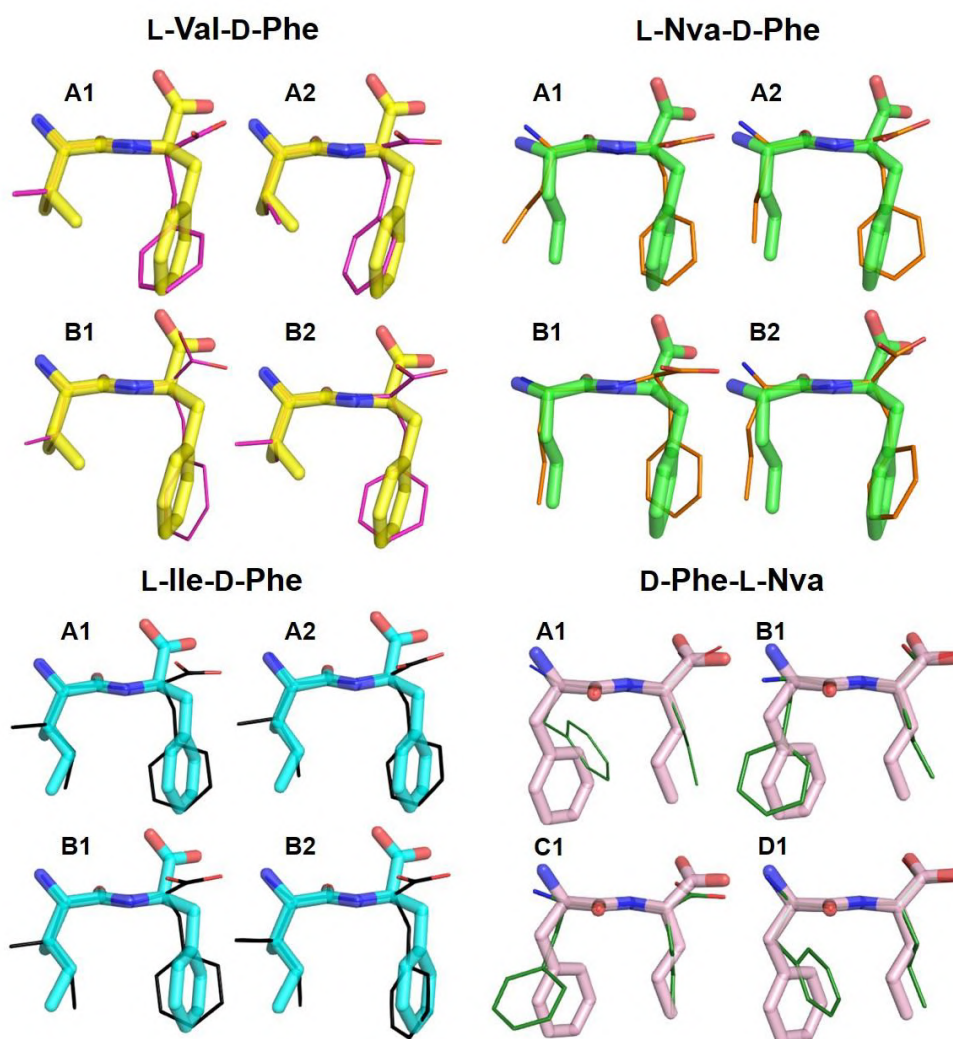
Figure S3. Representative significant (ΔG ≤ 2.0 kJ/mol) conformers extracted from the corresponding 2D conformational landscape (see Fig. S1) of the L-Val-D-Phe (top left), L-Nva-D-Phe (top right), L-Ile-D-Phe (bottom left), and D-Phe-L-Nva (bottom right) dipeptide. All the conformations (shown as coloured wire lines) have been overlapped with the corresponding experimental crystal units (shown as coloured sticks).[1,2,3] The relative free-energy values associated to these conformations are reported in Table S1.

E1 – ΔG = 2.13 kJ/mol   **E2 – ΔG = 2.49 kJ/mol**
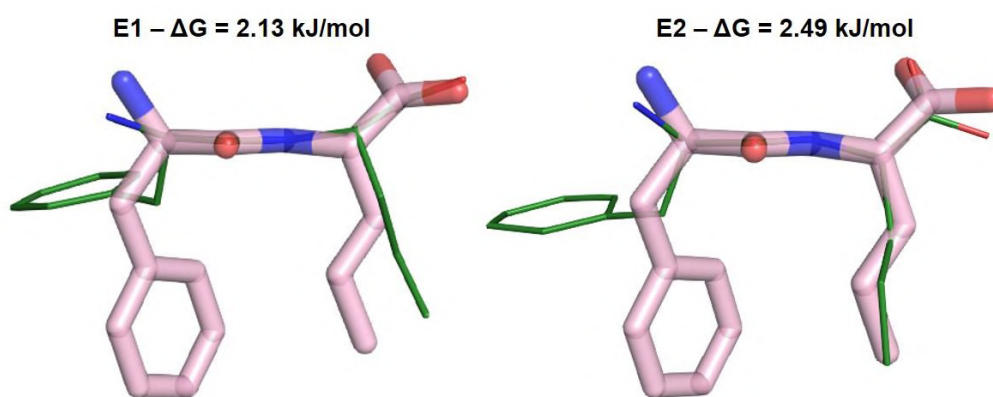
**Figure S4.** Representative conformers with a ΔG around (2.1-2.5) kJ/mol extracted from the 2D conformational landscape (see Fig. S1) of D-Phe-L-Nva. The two conformations (shown as pink wire lines) have been overlapped with the corresponding experimental crystal unit (shown as pink sticks).[2]

a) L-Val-D-Phe

Label	Relative Free Energy (kJ/mol)
A1	0.95
A2	2.04
B1	1.90
B2	0.0

b) L-Nva-D-Phe

Label	Relative Free Energy (kJ/mol)
A1	1.18
A2	0.0
B1	1.66
B2	0.96

c) L-Ile-D-Phe

Label	Relative Free Energy (kJ/mol)
A1	0.0
A2	1.96
B1	1.52
B2	0.56

d) D-Phe-L-Nva

Label	Relative Free Energy (kJ/mol)
A1	1.68
B1	0.0
C1	0.64
D1	1.71

Table S1. Relative Free Energy values of the significant conformers (ΔG ≤ 2.0 kJ/mol) extracted for: a) L-Val-D-Phe, b) L-Nva-D-Phe, c) L-Ile-D-Phe, and d) D-Phe-L-Nva. All the conformations have been shown in the previous Fig. S3.
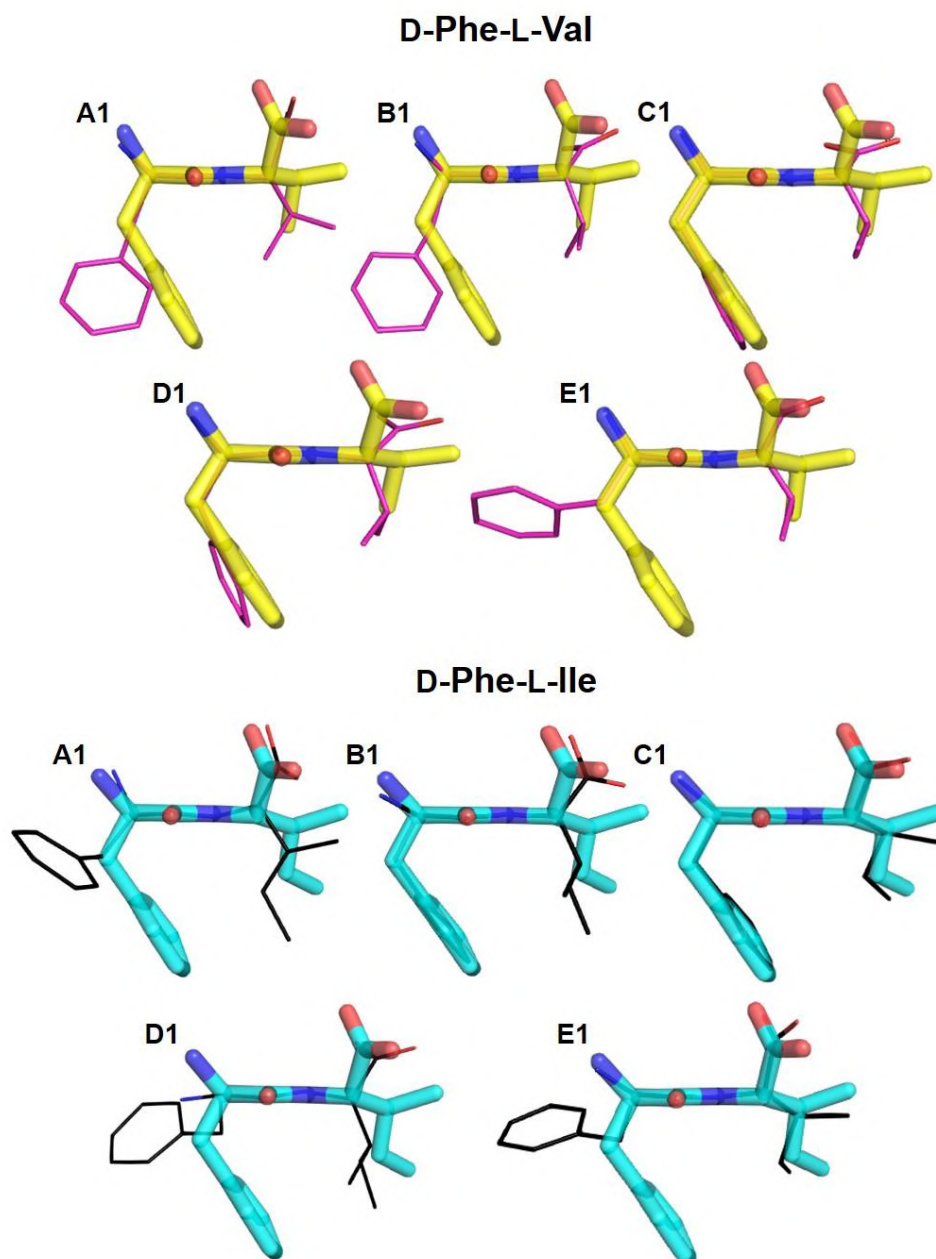
D-Phe-L-Val

D-Phe-L-Ile

Figure S5. Representative significant (ΔG ≤ 2.0 kJ/mol) conformers extracted from the corresponding 2D conformational landscape (see Fig. S2) of the D-Phe-L-Val (top), and D-Phe-L-Ile (bottom) dipeptide. All the conformations (shown as coloured wire lines) have been overlapped with the corresponding experimental crystal units (shown as coloured sticks).[1,3] The relative free-energy values associated to these conformations are reported in Table S2.

a) D-Phe-L-Val

Label	Relative Free Energy (kJ/mol)
A1	0.0
B1	1.18
C1	1.13
D1	0.074
E1	0.73

b) D-Phe-L-Ile

Label	Relative Free Energy (kJ/mol)
A1	0.0
B1	1.30
C1	0.21
D1	0.33
E1	1.70

Table S2. Relative Free Energy values of the significant conformers ($\Delta G \leq 2.0$ kJ/mol) extracted for: a) D-Phe-L-Val, and b) D-Phe-L-Ile. All the conformations have been shown in the previous Fig. S5.

D-Phe-L-Val
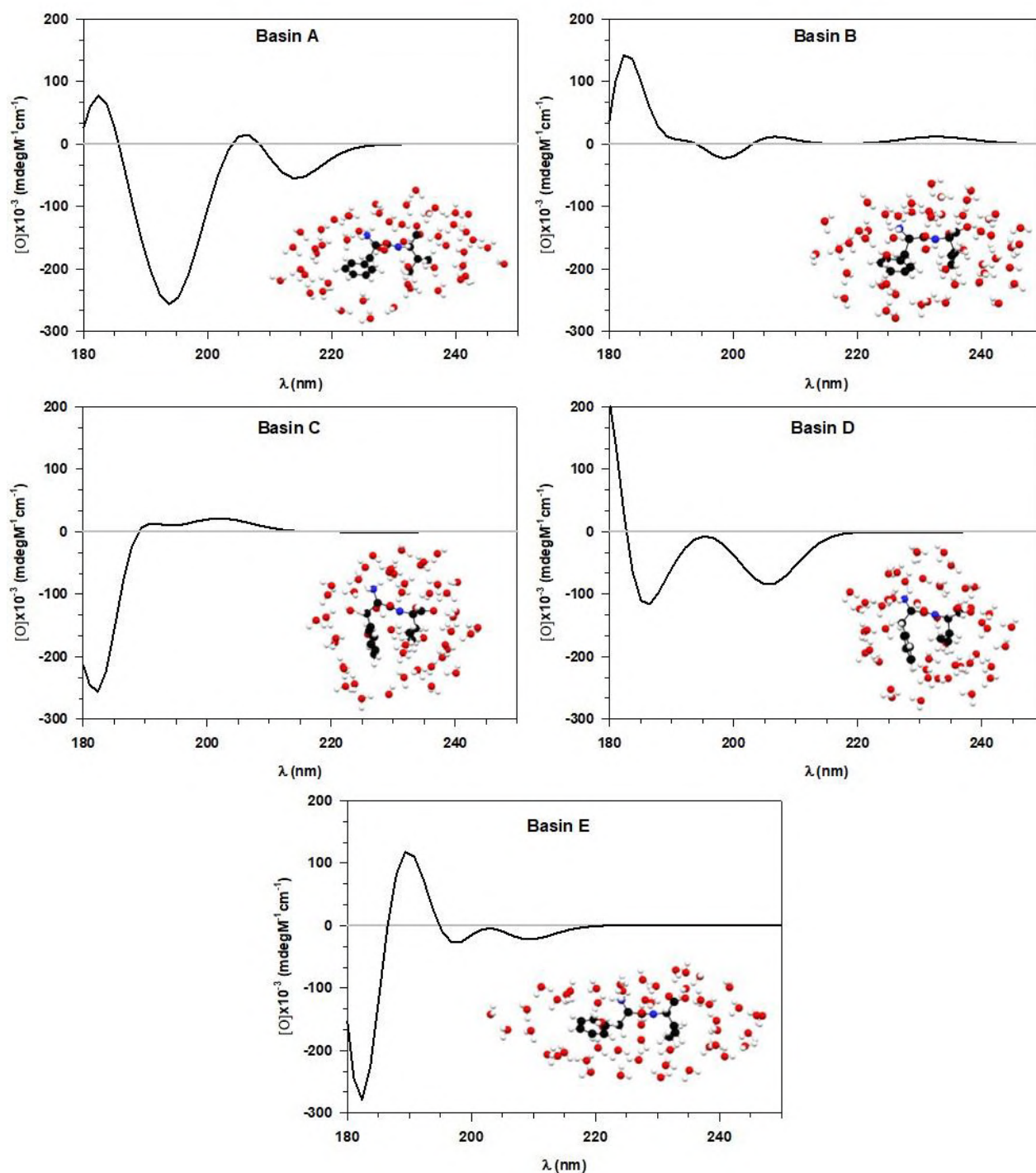
Figure S6. Individual calculated ECD spectra of the D-Phe-L-Val-(H₂O)₅₀ clusters. Here, we reported only the most probable structures obtained from the statistical analysis of both the dipeptide and the water conformational spaces. The labels of the basin refer to those defined for the conformational landscape of D-Phe-L-Val (see Fig. 1 in the main article). All the spectra are not weighted for their corresponding Boltzmann's factor. Technical details on the calculations have been reported above.
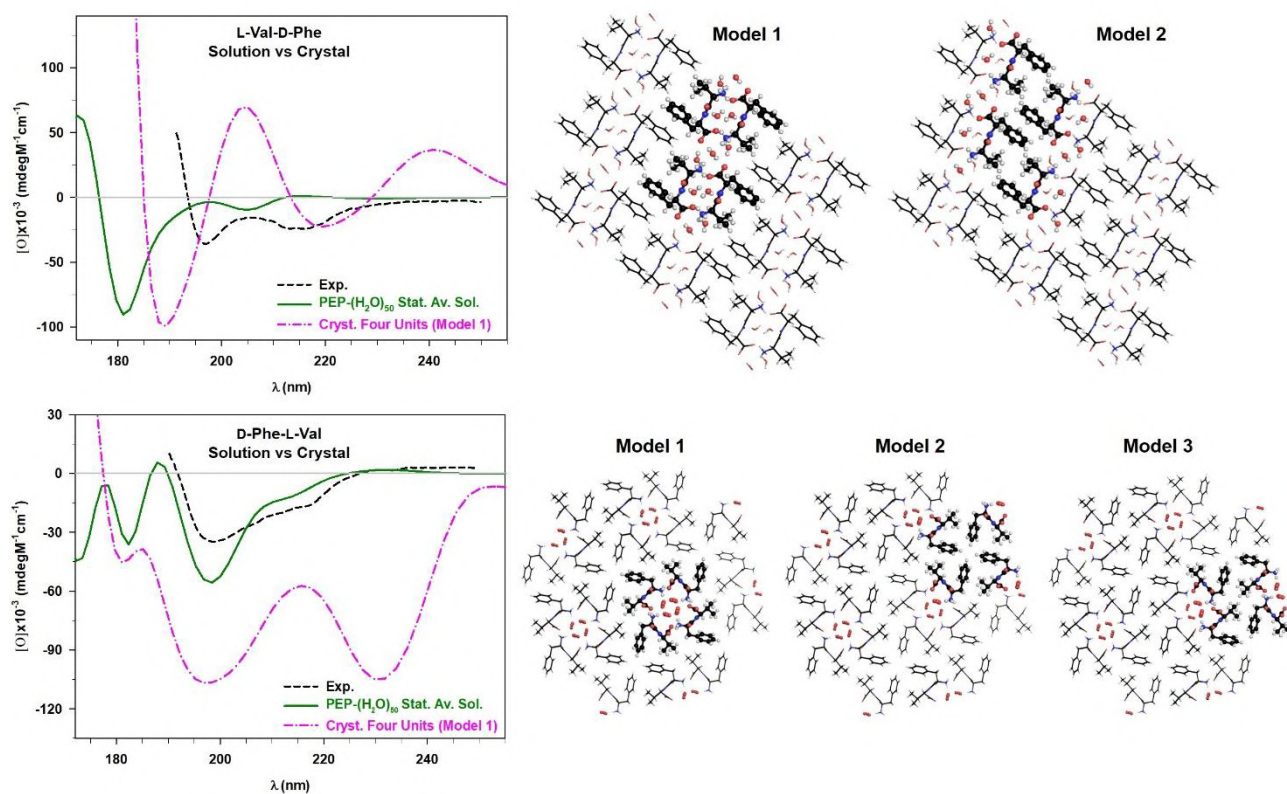
Figure S7. Comparison of the experimental (Exp.) ECD spectrum (dashed black line)[1] of L-Val-D-Phe, and D-Phe-L-Val, respectively, with those calculated: i) statistically averaging the ECD of the most probable conformations in solution (PEP + 50 H$_2$O mols., solid green line), ii) considering a four units crystal model (dash-dotted pink line). Different possible four-unit models are also shown on the right. The intensities of the crystal four units ECD have been divided by a factor 2. Technical details on the calculations have been reported above.
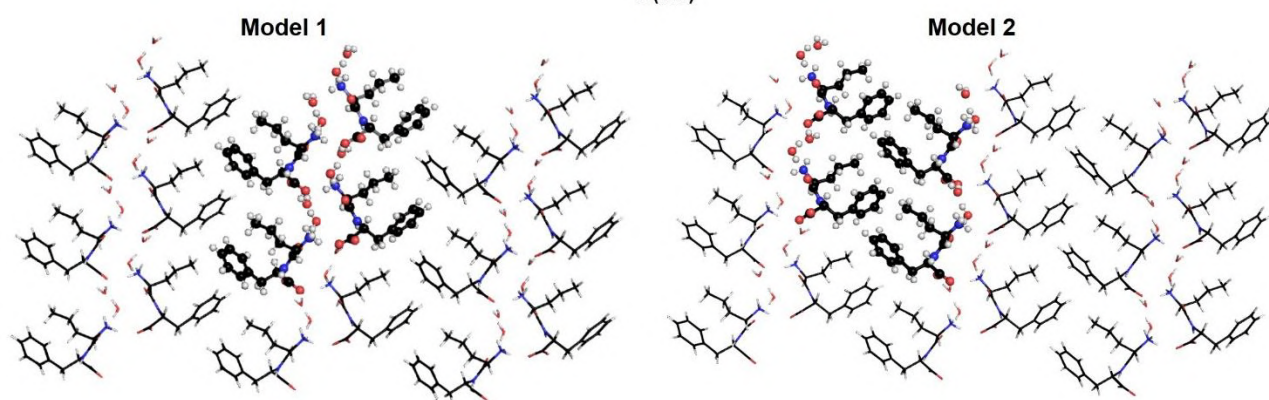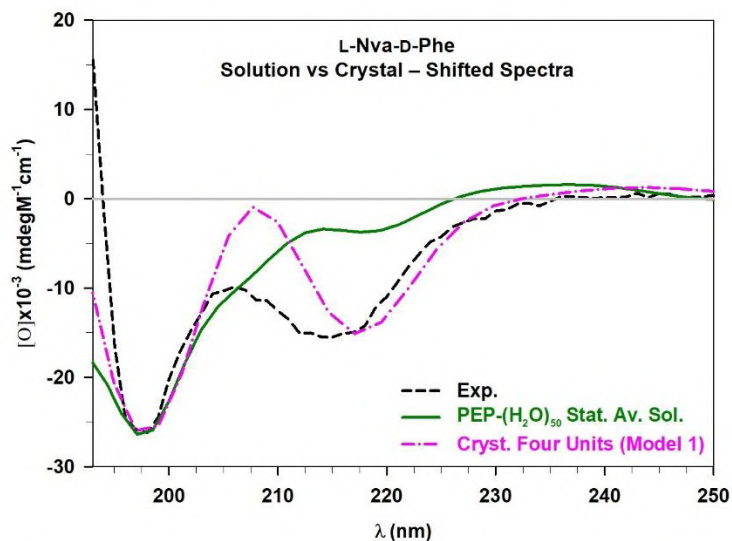
Figure S8. Comparison of the experimental (Exp.) ECD spectrum (dashed black line)[2] of L-Nva-D-Phe with those calculated: i) statistically averaging (stat. av.) the ECD of the most probable conformations in solution (PEP + 50 H_2O mols., solid green line), ii) considering a four units crystal model (dash-dotted pink line). Different possible four-unit models are also shown on the bottom. The stat. av. and the four units ECD have been shifted by +9.2 nm, and -28.5 nm, respectively, and their intensities have been normalised in order to match properly the experimental response. Technical details on the calculations have been reported above.
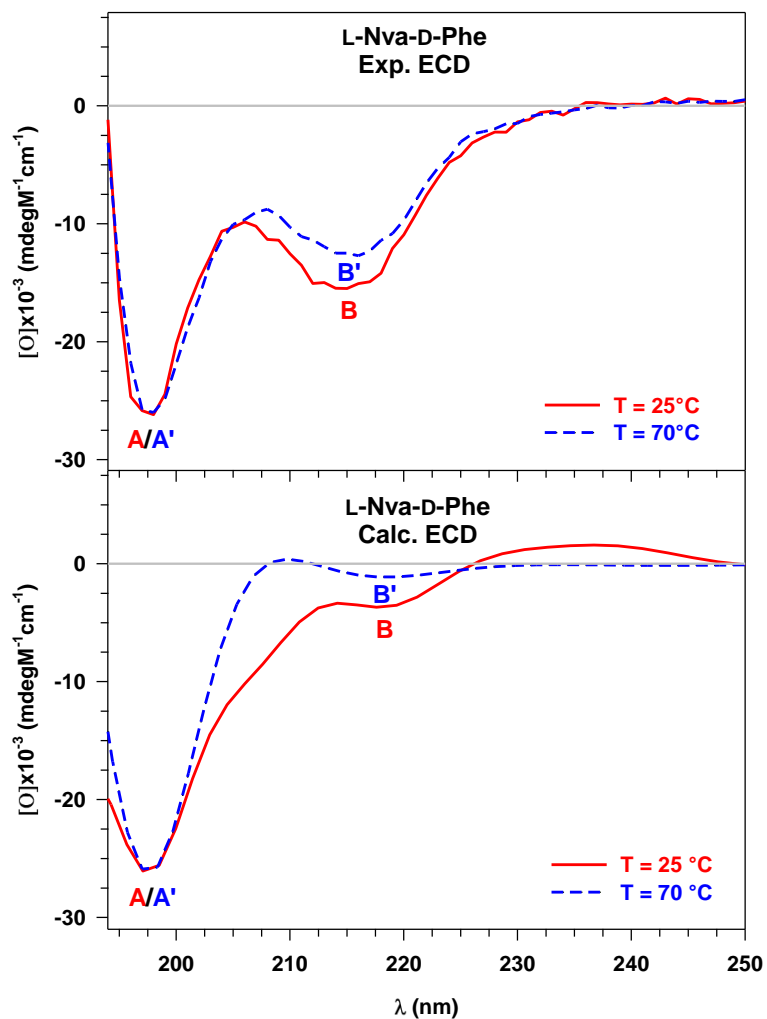
Figure S9. Experimental (Exp.) and calculated (Calc.) ECD spectra of L-Nva-D-Phe in aqueous solution at 25°C (solid red line) and 70°C (dashed blue line), respectively. The calculated spectra have been shifted by +9.2 nm (25°C) and +16 nm (70°C), and their intensities have been normalised in order to match properly the experimental response. Technical details on the calculations have been reported above.
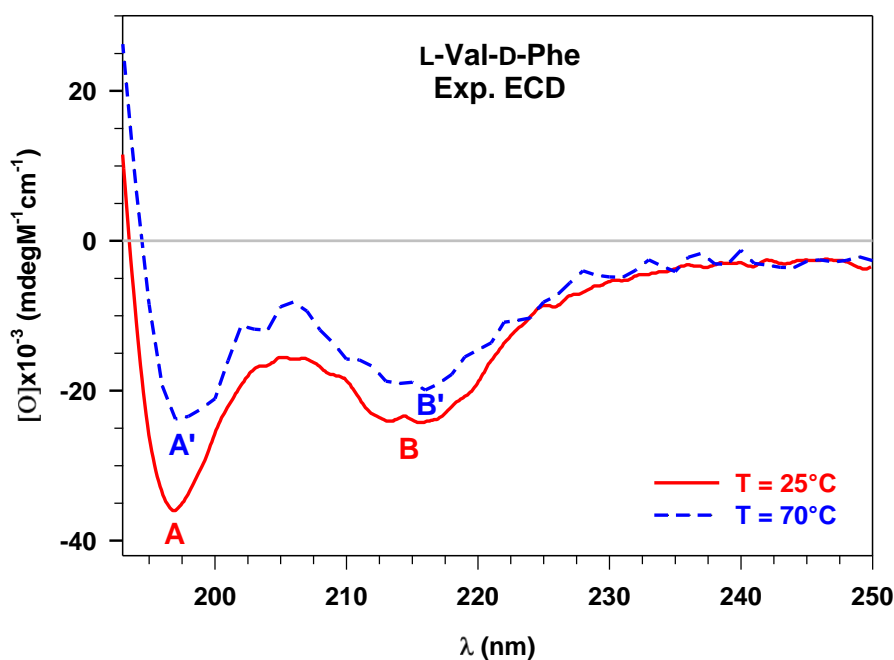
Figure S10. Experimental (Exp.) ECD spectra of L-Val-D-Phe in aqueous solution at 25°C (solid red line) and 70°C (dashed blue line), respectively.
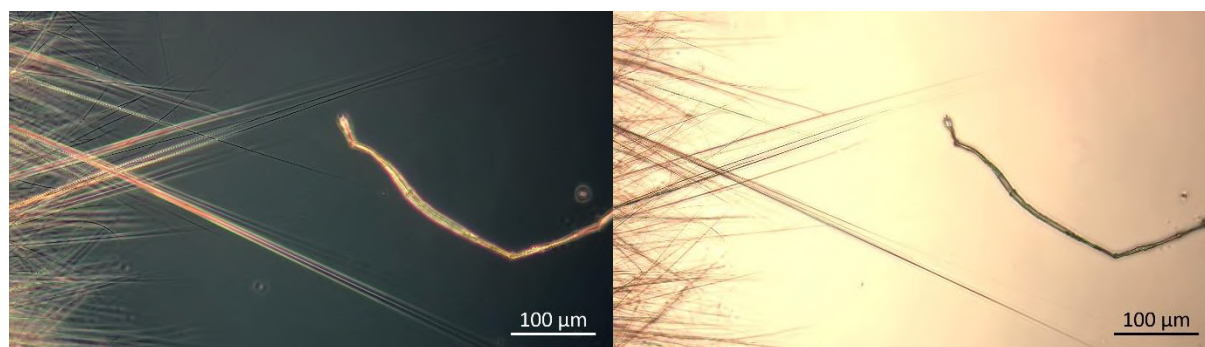
Figure S11: L-Val-D-Phe optical microscope images under polarised light (left) and normal light (right). During the collection of the fibre images after the high temperature experiment, the solution was left evaporate on the glass slide. Ordered crystalline needles appeared after several minutes on the left of the image. This highlights the different peptide behaviour at 25 °C.

References

[1] O. Bellotto, G. Pierri, P. Rozhin, M. Polentarutti, S. Kralj, P. D'Andrea, C. Tedesco, S. Marchesan, *Org. Biomol. Chem.,* 2022, **20**, 6211.

[2] E. Scarel, G. Pierri, P. Rozhin, S. Adorinni, M. Polentarutti, C. Tedesco, S. Marchesan, *Chemistry,* 2022, **4**, 1417.

[3] O. Bellotto, S. Kralj, M. Melchionna, P. Pengo, M. Kisovec, M. Podobnik, R. De Zorzi, S. Marchesan, *ChemBioChem.,* 2022, **23**, e202100518.

[4] M. Monti, M. Stener, M. Aschi, *J. Comput. Chem.,* 2022, **43**, 1997.

[5] M. Monti, G. Brancolini, E. Coccia, D. Toffoli, A. Fortunelli, S. Corni, M. Aschi, M. Stener, *J. Phys. Chem. Lett.,* 2023, **14**, 1941.

[6] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Comput. Phys. Commun.,* 1995, **91**, 43.

[7] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhot, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.,* 2005, **26**, 1701.

[8] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.,* 1988, **110**, 1657.

[9] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.

[10] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.,* 1983, **79**, 926.

[11] G. Bussi, M. Parrinello, *Comput. Phys. Commun.,* 2008, **179**, 26.

[12] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. Fraaije, *J. Comput. Chem.,* 1997, **18**, 1463.

[13] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.,* 1993, **98**, 10089.

[14] A. Amadei, B. M. Linssen, H. J. C. Berendsen, *Proteins,* 1993, **17**, 412.

[15] I. Daidone, A. Amadei, *WIREs Comput. Mol. Sci.,* 2012, **2**, 762.

[16] M. D'Alessando, A. Amadei, M. Stener, M. Aschi, *J. Comput. Chem.,* 2015, **36**, 399.

[17] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules,* Oxford University Press, New York, Oxford, 1989.

[18] G. Te Velde, M. Bickelhaupt, E. J. Baerendsen, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.,* 2001, **22**, 931.

[19] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.,* 2008, **10**, 6615.

[20] M. E. Casida, *Recent Advances in Density Functional Methods: (Part I) (World Scientifict),* 1995, 155.

[21] J. A. Schellman, *Chem. Rev.,* 1975, **75**, 323.

[22] O. Baseggio, M. De Vetta, G. Fronzoni, M. Stener, L. Sementa, A. Fortunelli, A. Calzolari, *J. Phys. Chem. C,* 2016, **120**, 12773.

[23] The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

4.1.5 Electronic circular dichroism from real-time propagation in state space

In collaboration with Prof. Emanuele Coccia, we proposed a new time-domain protocol for ECD calculations. This way, it is possible to go beyond the ground-state rotationally averaged ECD spectra that result from the adoption of frequency-domain methods in the linear response regime. In practice, we implemented an explicit interface between the AMS and WaveT codes to propagate the magnetic moment in the space of electronic states. Therefore, we defined a general approach to compute ECD spectra via the time-dependent Schrödinger equation, using the magnetic response to an electric perturbation. The procedure was tested on systems of different size and nature, such as *(R)/(S)*-methyloxirane, L-alanine, and the Λ-Co(acac)$_3$ (acac = acetylacetonato) complex. Frequency-domain spectra were calculated at the TDDFT level using Casida for reference. Moreover, the presented interface extracts from these TDDFT calculations both the excitation energies and transition dipole moments, then given in input for the real-time propagation in WaveT.

For all the test cases, an almost perfect overlap was obtained between the calculated frequency- and time-domain ECD spectra, thus confirming the correct simulation of the temporal evolution of the transition dipole moment. Furthermore, the comparison between the time-domain and experimental gas-phase ECD spectra showed the ability of our calculations to reproduce all the main chiroptical features. Some discrepancies were observed only for Λ-Co(acac)$_3$, but they can be related to the environment. Indeed, for the cobalt complex the experimental ECD is measured in ethanol, while our calculations are in gas-phase. Such hypothesis was corroborated by statically including the solvent in the wave function time evolution, providing a better description of the spectral features.

In addition to the validation of the protocol, excited spectra and the relation between the electric perturbation and the ECD response were investigated. For instance, we propagated the *(S)*-methyloxirane in its first excited-state but with its ground-state geometry (vertical excitation), thus resembling the conditions of a pump-probe experiment with a small delay between the two pulses. This investigation pointed out the significant optical changes that occur exciting the system. Furthermore, we decomposed the ECD of Λ-Co(acac)$_3$ in its components discussing the spectral response to different polarization of the external electric field.

Electronic circular dichroism from real-time propagation in state space

View Online   Export Citation   CrossMark

M. Monti, [iD] M. Stener, [iD] and E. Coccia[a) [iD]

AFFILIATIONS

Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Via L. Giorgieri 1, 34127 Trieste, Italy

[a)]Author to whom correspondence should be addressed: ecoccia@units.it

ABSTRACT

In this paper, we propose to compute the electronic circular dichroism (ECD) spectra of chiral molecules using a real-time propagation of the time-dependent Schrödinger equation (TDSE) in the space of electronic field-free eigenstates, by coupling TDSE with a given treatment of the electronic structure of the target. The time-dependent induced magnetic moment is used to compute the ECD spectrum from an explicit electric perturbation. The full matrix representing the transition magnetic moment in the space of electronic states is generated from that among pairs of molecular orbitals. In the present work, we show the ECD spectra of methyloxirane, of several conformers of L-alanine, and of the Λ-Co(acac)$_3$ complex, computed from a singly excited ansatz of time-dependent density functional theory eigenstates. The time-domain ECD spectra properly reproduce the frequency-domain ones obtained in the linear-response regime and quantitatively agree with the available experimental data. Moreover, the time-domain approach to ECD allows us to naturally go beyond the ground-state rotationally averaged ECD spectrum, which is the standard outcome of the linear-response theory, e.g., by computing the ECD spectra from electronic excited states.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0136392

I. INTRODUCTION

The electronic circular dichroism (ECD) is the absorption difference exhibited by enantiomers of chiral molecules when interacting with left- and right-circularly polarized light. The theoretical understanding of ECD has been widely reported in the literature over the years.[1–7] ECD is employed to assign the absolute configuration of chiral compounds.[8] One can compute the ECD spectrum by using the response of the time-dependent induced magnetic dipole to an electric-field perturbation or the response of the time-dependent electric dipole to a magnetic-field perturbation.

Plenty of theoretical developments to compute the ECD spectra of chiral molecules have been published over the years: wavefunction methods,[9–13] such as the time-dependent Hartree–Fock[10,14] and coupled cluster,[10,15–18] and parameterized approaches, such as the semiempirical method[19] and tight-binding based on Tamm–Dancoff approximation.[20] Among the various electronic-structure approaches, the time-dependent density functional theory (TDDFT) has been extensively employed for ECD applications and method development.[6,10,21–38]

Most of these methods rely on frequency domain, while less effort has been devoted to time-domain methods.[21,23–25] Here, we propose an approach based on the propagation of the time-dependent Schrödinger equation (TDSE) in the presence of an explicit electromagnetic field, in length gauge.[39–46] The time-dependent wave function is expanded into the set of eigenstates of the field-free Hamiltonian, making the approach general and applicable, in principle, to any level of theory for the electronic structure.[47–61] The time evolution of the magnetic dipole moment is also computed in the basis of field-free eigenstates. Propagating TDSE in the state space allows us to decouple the electronic-structure treatment from how TDSE is propagated in time, at variance with the methods based on the dynamics in the space of molecular orbitals:[62–64] first, the field-free problem is solved using the chosen methodology, providing ingredients for TDSE such as eigenenergies and transition dipole moments; second, TDSE is propagated in the presence of an explicit pulse to simulate the desired electron dynamics. Details are given in Sec. II.

In the present work, we computed the ECD spectrum of (R)- and (S)-methyloxirane, of various conformers of L-alanine, and of the Λ-Co(acac)$_3$ complex (Fig. 1), using TDDFT eigenstates, formulated within a singly excited ansatz, for defining the propagation space. We first compared the real-time ECD spectra with those obtained by a standard frequency-domain linear response and then with the experimental data, when available. Moreover, we studied the linear absorption and ECD of excited methyloxirane, assum-
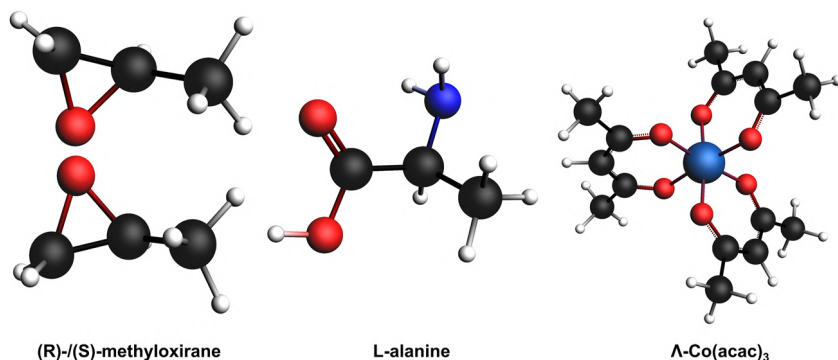
FIG. 1. Stick-and-ball representation of (R)-/(S)-methyloxirane, L-alanine, and Λ-Co(acac)$_3$. C, O, N, H, and Co atoms are reported in black, red, blue, white, and cobalt-blue, respectively.

(R)-/(S)-methyloxirane   **L-alanine**   **Λ-Co(acac)$_3$**

ing as the initial condition of the propagation the molecule being in its first bright electronic excited state,[64] and analyzed the contribution to the ECD spectrum of Λ-Co(acac)$_3$ due to the different linear polarizations of the incoming pulse.

This article is organized as follows: in Sec. II, we present our method based on TDSE propagation; the computational details are given in Sec. III; the results are shown and discussed in Sec. IV; and then, the conclusions and perspectives are given in Sec. V.

II. THEORY

A. Time propagation

Here, we focus on the electric perturbation of the magnetic response of the system. TDSE in length gauge is given by (atomic units are used in this work)

$$i\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle, \qquad (1)$$

where $|\psi(t)\rangle$ is the time-dependent wave function and $\hat{H}(t)$ is the time-dependent Hamiltonian, which is composed of the field-free Hamiltonian $\hat{H}_0$ and the coupling between the molecular electric dipole operator $\vec{\hat{\mu}}$ and the external field $\vec{F}(t)$,

$$\hat{H}(t) = \hat{H}_0 - \vec{\hat{\mu}} \cdot \vec{F}(t). \qquad (2)$$

$|\Psi(t)\rangle$ is practically expressed as a linear combination of the N_states field-free eigenstates (in this work, the DFT ground state and the $N_\text{states}-1$ TDDFT eigenstates) as

$$|\psi(t)\rangle = \sum_{M=0}^{N_\text{states}-1} C_M(t)|M\rangle. \qquad (3)$$

Here, $C_M(t)$ are time-dependent coefficients and $|M\rangle$ is the Mth eigenstate of the system, with the eigenvalue E_M. In the space of such eigenstates, TDSE in Eq. (1) becomes

$$i\frac{d\mathbf{C}(t)}{dt} = \mathbf{H}(t)\mathbf{C}(t), \qquad (4)$$

with $\mathbf{C}(t)$ being the vector of the expansion coefficients and $\mathbf{H}(t)$ being the matrix representation at time t of $\hat{H}(t)$, i.e., $(\mathbf{H}(t))_{LM} = \langle L|\hat{H}(t)|M\rangle$. The time-dependent Hamiltonian $\hat{H}(t)$ is diagonal for the field-free part of the eigenenergies E_M and characterized

by the transition (electric) dipole moments $\langle L|\hat{\mu}|M\rangle$ between states,[44]

$$\langle L|\hat{H}(t)|M\rangle = E_M \delta_{LM} - \sum_\gamma F_\gamma(t)\langle L|\hat{\mu}_\gamma|M\rangle, \qquad (5)$$

where $\gamma = x, y$, or z corresponds to the Cartesian component of the dipole and of the field. The numerical gauge-invariance for the ECD spectra has been previously tested and verified.[22] We use the TDDFT eigenvectors within a configuration-interaction singles ansatz[44,65] for the excited states,

$$|M\rangle = \sum_i^{occ} \sum_a^{vir} d_{i,M}^a |\Phi_i^a\rangle, \qquad (6)$$

where $|\Phi_i^a\rangle$ is the singly excited Slater determinant, with an electron promoted from the occupied orbital i to the virtual one a, while $d_{i,M}^a$ are the amplitudes of the expansion for the state $|M\rangle$. Expansion coefficients, $d_{i,M}^a$, and the excited-state eigenenergies are computed by the Amsterdam Modeling Suite (AMS),[66] as reported in Sec. III.

B. Computing ECD spectrum

In order to compute the ECD spectrum in a large frequency range, we apply a kick pulse, modeled by a narrow Gaussian

$$\vec{F}(t) = \vec{F}_\text{max} \exp\left(-\frac{(t-t_0)^2}{2\sigma^2}\right), \qquad (7)$$

where $\vec{F}_\text{max}$ is the field amplitude (the intensity I is equal to $\frac{1}{2}|\vec{F}_\text{max}|^2$) and t_0 and σ are the center and the amplitude of the Gaussian, respectively.

The magnetic dipole moment operator $\vec{m}$ is defined, neglecting the spin contribution, as[2]

$$\vec{m} = -\frac{1}{2c}\vec{r} \times \vec{p}, \qquad (8)$$

with $\vec{r}$ and $\vec{p}$ being the position and momentum operators, respectively, and c being the light speed.

At each time t of the simulation, the induced magnetic dipole $\Delta\vec{m}(t)$ is defined as the difference between the time-dependent

magnetic moment $\vec{m}(t)$ at time t and that at initial time ($t = 0$), $\vec{m}(0)$,

$$\Delta \vec{m}(t) = \vec{m}(t) - \vec{m}(0). \tag{9}$$

Explicitly, $\vec{m}(t)$ and $\vec{m}(0)$ are computed from the TDSE propagation in state space,

$$\vec{m}(t) = \sum_{LM} C_L^*(t) C_M(t) \langle L|\vec{m}|M\rangle, \tag{10}$$

$$\vec{m}(0) = \sum_{LM} C_L^*(0) C_M(0) \langle L|\vec{m}|M\rangle. \tag{11}$$

The transition magnetic moment $\langle L|\vec{m}|M\rangle$ can be expressed on the basis of Slater determinants. When one of the electronic states is the ground state $|0\rangle = |\Phi_0\rangle$ (with $|\Phi_0\rangle$ the DFT Slater determinant, in our specific case), one finds, using Eq. (6),

$$\langle L|\vec{m}|0\rangle = \sum_i^{occ} \sum_a^{vir} d_{i,L}^a \langle \Phi_i^a|\vec{m}|\Phi_0\rangle, \tag{12}$$

while the transition magnetic moment between two excited states is

$$\langle L|\vec{m}|M\rangle = \sum_i^{occ} \sum_a^{vir} \sum_j^{occ} \sum_b^{vir} d_{i,L}^a d_{j,M}^b \langle \Phi_i^a|\vec{m}|\Phi_j^b\rangle. \tag{13}$$

The generic transition magnetic moment $\langle \Phi_i^a|\vec{m}|\Phi_j^b\rangle$ is, in turn, obtained by applying the Slater–Condon rules,[44]

$$\langle \Phi_i^a|\vec{m}|\Phi_j^b\rangle = \begin{cases} \vec{M}_0 - \vec{m}_{ii} + \vec{m}_{aa}, & i = j, a = b, \\ \vec{m}_{ab}, & i = j, a \neq b, \\ -\vec{m}_{ji}, & i \neq j, a = b, \\ 0, & i \neq j, a \neq b, \end{cases} \tag{14}$$

where $\vec{M}_0 = \langle \Phi_0|\vec{m}|\Phi_0\rangle = \sum_i \vec{m}_{ii}$ and the generic $\vec{m}_{st}$ is the magnetic moment matrix element in the molecular-orbital (MO) representation. In practice, AMS explicitly computes the transition matrix elements in the MO space, which are then transformed into $\langle \Phi_i^a|\vec{m}|\Phi_j^b\rangle$ and into $\langle L|\vec{m}|M\rangle$ by means of an interface between AMS and WaveT, which accounts for the antisymmetric nature of the magnetic matrix. The ECD spectrum is finally computed as the imaginary part of the following:

$$P_{nl}^{ECD}(\omega) = -\frac{i}{2\pi\omega F_n^0(\omega)} \int_0^{+\infty} - \Delta m_l(t) e^{i(\omega+i\Gamma)t} dt. \tag{15}$$

In Eq. (15), $F_n^0(\omega)$ is the Fourier transform of the component n of the external field, $\Delta m_l(t)$ is the component l of the induced magnetic dipole, and Γ is a damping parameter reproducing an exponential decay of the excited-state population.[22] The comparison with the experimental spectra is performed using a rotational average, i.e.,

$$\bar{P}^{ECD}(\omega) = \frac{1}{3}\left(P_{xx}^{ECD}(\omega) + P_{yy}^{ECD}(\omega) + P_{zz}^{ECD}(\omega)\right). \tag{16}$$

III. COMPUTATIONAL DETAILS

The ground-state geometries of the (S)-methyloxirane and Λ-Co(acac)$_3$ system have been optimized at the DFT level[67] by the Amsterdam Density Functional (ADF) engine[68] of the AMS code.[66] Both the optimizations were realized using the hybrid B3LYP[69] exchange-correlation (xc) functional and a triple-zeta plus polarization TZP basis set of Slater-type orbitals. The geometry of the (R)-methyloxirane has been obtained by inverting the x coordinates of the optimized S enantiomer. For the L-alanine, we have considered seven low-energy conformers, taking both the relative energies and the geometry structures, optimized at the df-CCSD(T)-F level with the aug-cc-pVDZ-F12 basis set, from a recent study.[70] For all the systems, we have then calculated the ECD spectra in the frequency domain at the time-dependent DFT (TDDFT) level by using the Casida approach.[71] The eigenvalue problem is solved by means of the Davidson algorithm. We have extracted the lowest 20, 40, and 100 excited-states for the methyloxirane, L-alanine, and Λ-Co(acac)$_3$, respectively. These choices have been made in order to cover the experimental frequency range. The ECD calculations have been performed using the B3LYP and the long-range corrected CAM-B3LYP[72] functionals. Moreover, for the L-alanine, we have tested the range-separated wB97X-D functional, which includes the dispersion and non-local exchange corrections, allowing a proper treatment of the amino acid intra-molecular H bonds.[32,73] All the ECD spectra have been calculated in gas phase. For the Λ-Co(acac)$_3$ complex, we have additionally performed the ECD calculation, simulating the solvent (ethanol) effect with the conductor like screening model (COSMO) of solvation[74] available in AMS. Finally, all the ECD spectra have been convoluted by using Gaussian functions with a half width at half maximum (HWHM) equal to 0.15 eV.

The time-domain ECD calculations have been performed with the in-house WaveT package,[43] whose interface with AMS has been recently proposed for the electric transition dipole moment[44] and here extended for the magnetic transition dipole moment, as explained in Sec. II. In practice, excitation energies and transition dipole moments (both electric and magnetic, after the transformation performed by the interface) are provided by the TDDFT calculations and then employed as input parameters for the real-time propagation. For all the systems in Fig. 1, 250-fs dynamics have been simulated with a time step δt of 0.12×10^{-2} fs. Different values of time length (100, 150, 250, and 300 fs) and time step (0.12×10^{-2}, 0.24×10^{-2}, and 0.24×10^{-3} fs) have been tested to ensure the convergence of the dynamics. The results of these convergence tests are shown in Figs. S1 and S2 of the supplementary material. The time-dependent external field [see Eq. (7)] has been reproduced with a kick pulse of 10^5 W/cm^2 intensity and a full width at half maximum of 0.094 fs for the Gaussian envelope. The value of the damping parameter Γ in Eq. (15) is 5 fs for all the cases. The ECD spectrum in Eq. (15) has been computed with a post-processing tool implemented in WaveT, using the TDSE dynamics as input. The final spectra have been obtained averaging the results, according to Eq. (16).

For all the cases, the energy range of both frequency- and time-domain ECD calculations has been adapted to the experimental one (i.e., eV, nm, and cm^{-1}) for a suitable comparison of the results.

IV. RESULTS AND DISCUSSION

This section is organized as follows: for each molecule in Fig. 1, we first compare the time-domain ECD spectra [i.e., obtained by Eqs. (15) and (16)] with the standard frequency-domain ones and

then with the experimental data available in the literature,[70,75,76] using the units of the original works. Additional analyses on the excited-state absorption, the role of the solvent in ECD, and the interplay between the $\vec{F}$ polarization and magnetic transition dipole moment are reported for S-methyloxirane and Λ-Co(acac)$_3$ in the corresponding subsections.
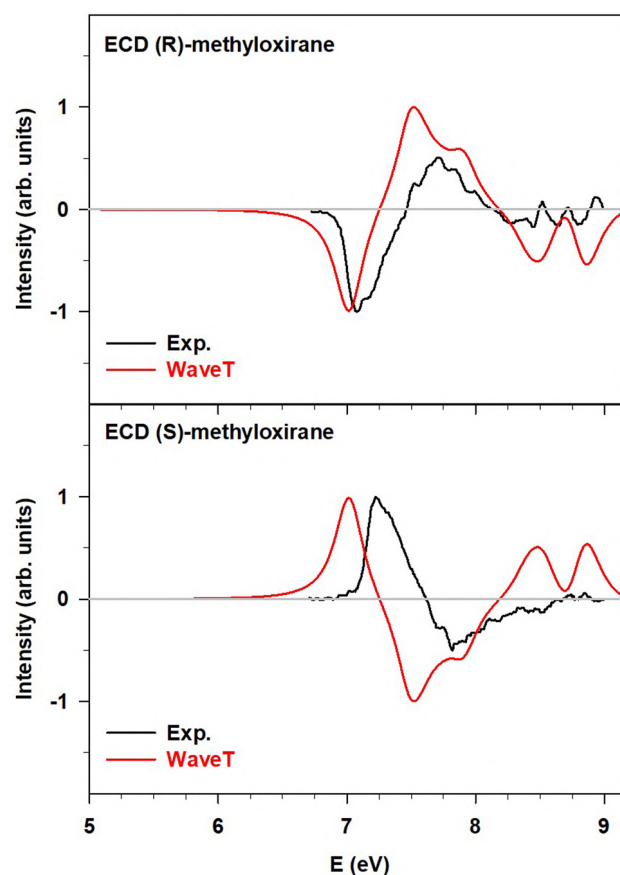
A. (R)- and (S)-methyloxirane

The ECD spectra obtained with AMS (frequency domain) and WaveT (time domain), using the B3LYP functional, reveal an almost perfect overlap for both the enantiomers of the methyloxirane, as shown in Fig. 2: the shape and relative intensities of the peaks are conserved moving from a frequency to a real-time calculation. The small discrepancies could be due to the different smoothing techniques applied. These results clearly show that the interface between the two codes produces the right transition magnetic moments in the state space of Eqs. (12) and (13) and, in turn, that WaveT computes the right temporal evolution of the magnetic transition dipole moment. Considering the CAM-B3LYP functional

(see Fig. S3 of the supplementary material), we still find a good agreement between the frequency-domain and time-domain calculations, with only the small discrepancies that can be associated again with the smoothing strategy. We can also notice that the overall ECD pattern calculated with CAM-B3LYP resembles that obtained with the B3LYP functional, although blueshifted by around 0.6 eV and with some differences in terms of the intensity ratio of the peaks. As a matter of fact, such discrepancies make the agreement between the CAM-B3LYP calculations and the experimental spectra less accurate (Fig. S4 of the supplementary material). Therefore, from here on, we will only consider the real-time calculations carried out with B3LYP for the comparison with the experimental ECD.

We can notice in Fig. 3 an overall good agreement between the experimental ECD spectra measured in the gas-phase[75] and our time-domain ECD propagated in vacuum, with the discrepancies being more marked for the (S)-methyloxirane. Indeed, the experimental positive feature located at 7.20 eV is slightly redshifted in the calculation and the same goes for the following peak, whose experimental negative maximum is found at around 7.80 eV. In addition to that, the calculated peaks from 8.5 eV onward show
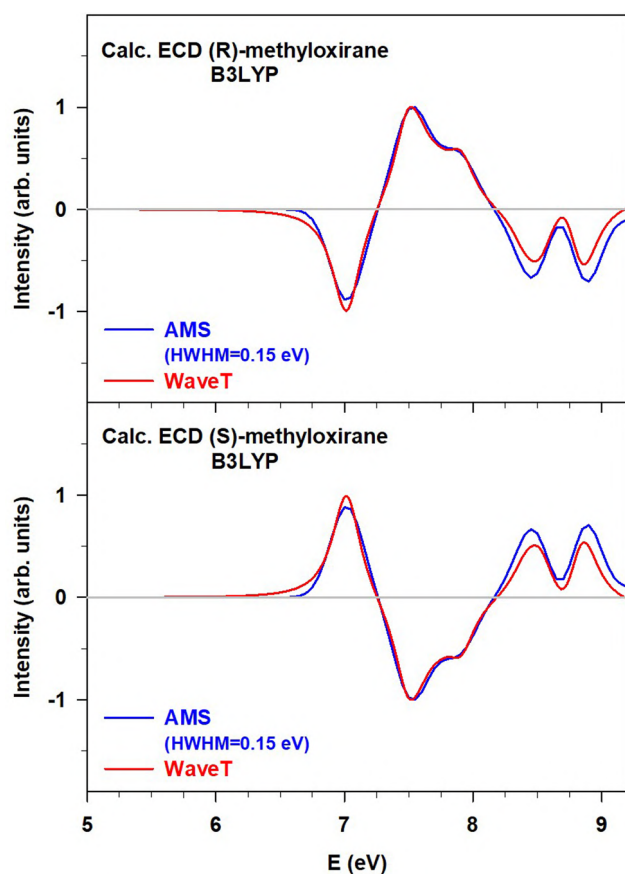
FIG. 2. Comparison of the ECD spectra calculated for (R)-methyloxirane (top panel) and (S)-methyloxirane (bottom panel) using the B3LYP functional and the AMS (frequency-domain, solid blue line) and the WaveT (time-domain, solid red line) code. All the intensities have been normalized and reported in arbitrary units (arb. units).


FIG. 3. Comparison between the experimental (solid black line)[75] and real-time calculated (solid red line) (B3LYP functional) ECD spectra of (R)-methyloxirane (top panel) and (S)-methyloxirane (bottom panel). All the intensities have been normalized and reported in arbitrary units (arb. units).

a higher intensity with respect to the corresponding experimental ones. Nevertheless, most of the discrepancies are reconciled, noticing that while the calculated features of the two enantiomers have equal and opposite signs, the experimental patterns slightly deviate from being mirror images. Therefore, considering this experimental limitation, both the calculations properly reproduce all the main features of the system.

Besides checking the implementation, we also calculated the absorption and ECD spectra of (S)-methyloxirane from the first (bright) excited state. For computing the optical properties from ground state, the initial condition in the electron dynamics is given by $|C_0(t=0)|^2 = 1$; see Eq. (3). For a generic $|M\rangle$ state, with $M > 0$, running an excited-state dynamics corresponds to $|C_M(t=0)|^2 = 1$. In our case, $M = 1$. The simulation was run on the ground state geometry, thus simulating a vertical excitation. This scheme resembles a pump–probe experiment, where the delay time between the two pulses is small enough to measure the observable of interest without changing the geometry of the system. Indeed, time-resolved circular dichroism spectroscopy has recently become a powerful experimental tool to study the dynamics in chiral molecular systems.[77–82] The influence of the electronic state on the optical

response is well highlighted in Fig. 4, where significant changes have been found populating the first excited state for both the absorption and ECD spectra. Indeed, looking at the top panel of Fig. 4, we observe that the ground state absorption peaks, located in the vacuum-UV region, are drastically shifted toward the near-IR/visible energy range when the system is excited. Furthermore, the inset in the top panel of Fig. 4 reveals the presence of the absorption and emission peaks involving the $|0\rangle$ and the $|1\rangle$ states, both located at 7 eV. A significant change in the intensity scale is also noticed, with a strong $|0\rangle \rightarrow |3\rangle$ transition of absorption (i.e., peak A′ in Fig. 4, and Table I). Additional details on the principal transitions of both the absorption spectra are reported in Table I. The chiral features of the (S)-methyloxirane are also markedly dependent on the electronic state, as shown in the bottom panels of Fig. 4. In this case, the optical response in the vacuum-UV range is still observable for the first excited state, even though the main feature lies between (0–2) eV, but the pattern is diametrically opposite with respect to the ground-state one. Indeed, focusing on the peak centered at ~7 eV, where the $|1\rangle \rightarrow |0\rangle$ emission peak is observed (see above), it can be suggested that an ECD inversion occurs when the system is excited. A similar spectrum for the first-excited state has been recently reported by
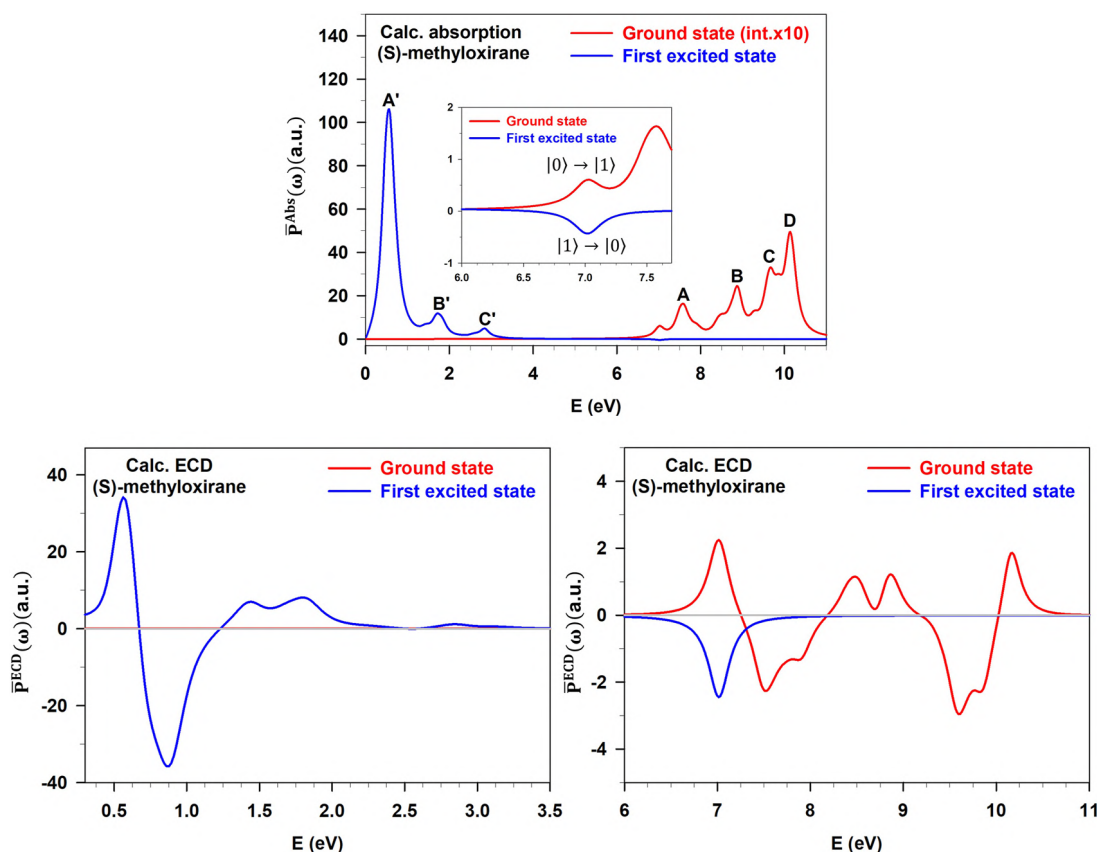
FIG. 4. Calculated (Calc.) absorption (top panel) and ECD (bottom panels) spectra of (S)-methyloxirane considering the ground state (solid red lines) and the first excited state (solid blue lines). In the inset of the top panel, the region between 6.0 and 7.7 eV has been highlighted to show the absorption and emission peaks involving the $|0\rangle$ and $|1\rangle$ states. Details of the significant absorption transitions are reported in Table I. Lower panels represent the same ECD spectrum up to 3.5 eV (left) and between 6 and 11 eV (right), with different vertical scales. No ECD signal is present between 3.5 and 6 eV. Intensities are in atomic units.

TABLE I. Transition and excitation energy for each significant peak of absorption of (S)-methyloxirane considering both the ground state and the first excited state.

Label	Transition	Excitation energy (eV)
A	$\lvert 0\rangle \rightarrow \lvert 3\rangle$	7.58
B	$\lvert 0\rangle \rightarrow \lvert 11\rangle$	8.87
C	$\lvert 0\rangle \rightarrow \lvert 15\rangle$	9.68
D	$\lvert 0\rangle \rightarrow \lvert 17\rangle$	10.11
A$'$	$\lvert 1\rangle \rightarrow \lvert 3\rangle$	0.56
B$'$	$\lvert 1\rangle \rightarrow \lvert 9\rangle$	1.72
C$'$	$\lvert 1\rangle \rightarrow \lvert 16\rangle$	2.83

Andersen and co-workers,[18] even though only the 0–2 eV range is shown in that work.

B. L-alanine

For L-alanine, we compared the frequency-domain and time-domain ECD spectra considering the B3LYP, CAM-B3LYP, and wB97X-D xc functionals, as explained in detail in Sec. III and as shown in Fig. 5. We have used the most stable conformer of the amino acid, as reported in Ref. 70. All the results support the consistency between the Casida and real-time calculations. Differences arise in the ECD spectrum according to the chosen xc functional. For instance, after redshifting the frequency-domain and time-domain wB97X-D spectra by 10 nm, we can notice two distinguishable positive peaks in the 140–160 nm energy range, while only a single positive feature is centered at around 145/150 nm for the B3LYP/CAM-B3LYP ECD spectra, respectively. From 170 nm onward, some similarities for the features calculated with the range-separated functionals (CAM-B3LYP and wB97X-D) are found, although they differ in terms of intensity values, while the response is less defined with B3LYP. Since the calculation performed with the wB97X-D shows closer resemblances with the experimental spectrum (as it will be shown below), consistently with the data available in the literature for amino acids and peptides,[32,73] it will be the only one considered from here on for L-alanine.
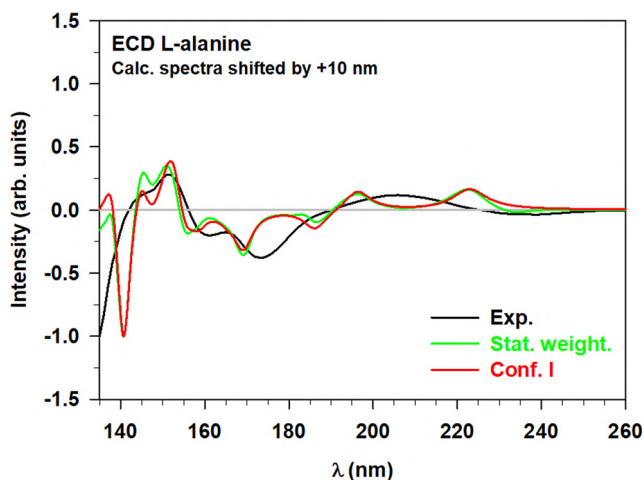
Before comparing the experimental and calculated spectra, it is worth emphasizing the high sensitivity of the ECD toward conformational changes. This aspect becomes particularly relevant when a flexible system, e.g., an amino acid, is treated, since a significant number of conformations will contribute to the overall optical response. Therefore, a balance must be found between the inclusion of, at least, the low-energy conformations and the computational effort, which increases proportionally to the number of structures considered. The conformational investigation of the gas-phase alanine has been carried out in several studies,[83,84] including the recent work of Meinert et al.,[70] where seven stable conformers have been extracted and used in the ECD calculations. Starting from these available data, we have calculated a statistically weighted ECD at 300 K and then compared with the experimental response and the calculated spectrum for the most probable conformation (as already shown in Fig. 5). All the results are collected in Fig. 6, while details on the conformers are reported in Table S1 of the supplementary material. We can notice that the experimental features are well reproduced by both the redshifted computed spectra, with only some
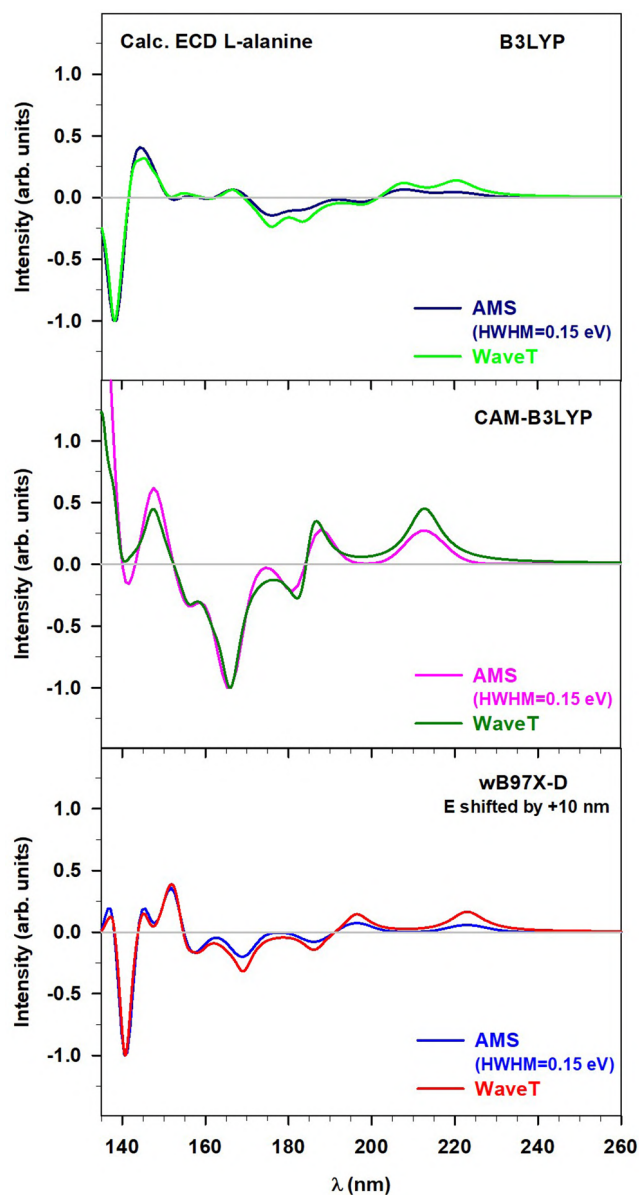
FIG. 5. Comparison of the ECD spectra calculated with AMS (frequency-domain) and WaveT (time-domain) for L-alanine. The B3LYP (top panel), CAM-B3LYP (middle panel), and wB97X-D (bottom panel) xc functionals have been used. The ECD spectra calculated with the wB97X-D functional have been redshifted by +10 nm. All the intensities have been normalized and reported in arbitrary units (arb. units).

differences in the intensity around 180 nm. Figure 6 also reveals strong similarities between the statistical approach and the inclusion of only the most probable conformer, but they are justified considering the energy and probability values reported in Table S1 of the supplementary material. Indeed, since the lowest energy conformer is much more stable with respect to the other structures, its optical response becomes predominant in the averaged ECD spectrum. However, the individual time-domain ECD spectra reported in

FIG. 6. Comparison between the experimental (solid black line)[70] and real-time calculated (wB97X-D functional) ECD spectra of L-alanine. The spectrum calculated statistically weighting the seven lowest energy conformations (solid green line) is reported together with that calculated considering only the most probable conformation (solid red line). All the intensities have been normalized and reported in arbitrary units (arb. units).

Fig. S5 of the supplementary material confirm the significant interplay between conformations and chirality.

C. Λ-Co(acac)₃

Also for the Λ-Co(acac)$_3$ complex, we started from the comparison between the frequency-domain and time-domain spectra, which confirms the good quality agreement between the AMS and WaveT results obtained employing the B3LYP functional, as shown
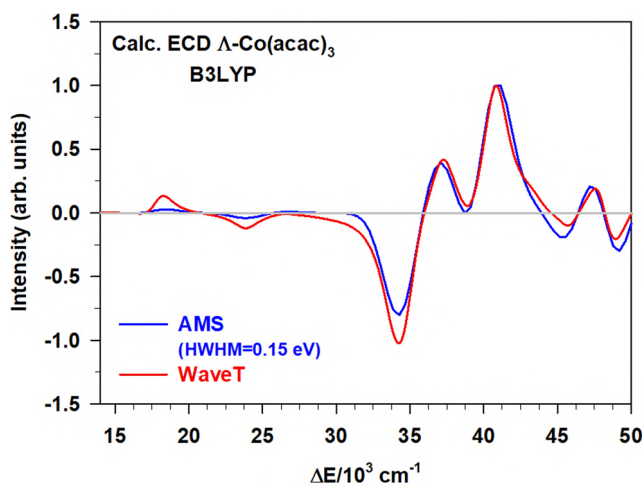
in Fig. 7. Indeed, despite the small differences in intensity, particularly focused around 17×10^3 and 35×10^3 cm^{-1}, the Casida and real-time calculations match properly. The same comment can be applied to the CAM-B3LYP results (see the top panel of Fig. S6 of the supplementary material), with only small discrepancies in terms of intensity, particularly in the high energy range, between the frequency-domain and time-domain data. The CAM-B3LYP spectrum is blueshifted with respect to the B3LYP one. In addition to that, the intensity ratio of the two positive features centered at 37×10^3 and 41×10^3 cm^{-1}, respectively, in the B3LYP spectra is inverted in the CAM-B3LYP ones. Such differences lead to a worse agreement with the experimental ECD for the real-time calculation with CAM-B3LYP (Fig. S6 of the supplementary material, bottom panel); thus, only the B3LYP results will be discussed from here on.

The overall quality agreement is also observable comparing the experimental spectrum[76] with the real-time propagation (Fig. 8), which, hence, reproduces the main chiroptical properties of the Λ-Co(acac)$_3$, despite some energy and intensity discrepancies. Indeed, the experimental negative peaks centered at 31×10^3 and 45×10^3 cm^{-1}, respectively, find the theoretical correspondence in the peaks located at around 34×10^3 and 46×10^3 cm^{-1}, with the first calculated one exceeding in intensity. Instead, the positive features at ~17×10^3 and 39×10^3 cm^{-1} are properly reproduced by the small and double peak at around 19×10^3 and 40×10^3 cm^{-1}, respectively. It is worth noting that some limitations arise from the difference in the environment, since the experimental ECD has been measured in ethanol, while our dynamics is realized in vacuum (Fig. 8).

Therefore, we partially investigated the solvent role propagating the time-dependent wave function expanded in the TDDFT eigenstates equilibrated with an implicit description of ethanol at the AMS level and then propagated in vacuum. The corresponding result calculations are reported together with the experimental findings and the gas-phase ECD in Fig. 9. We can notice that even the partial treatment of the solvent provides a better description,
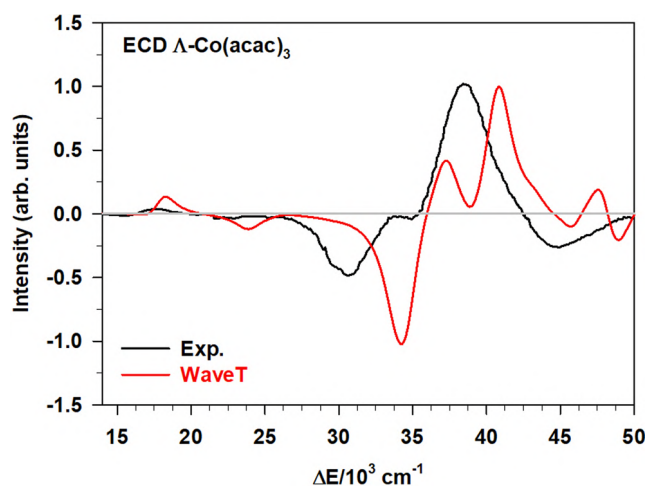

FIG. 7. Comparison of the ECD spectra calculated with AMS (frequency-domain, solid blue line) and WaveT (time-domain, solid red line) for the Λ-Co(acac)$_3$ complex. All the intensities have been normalized and reported in arbitrary units (arb. units).
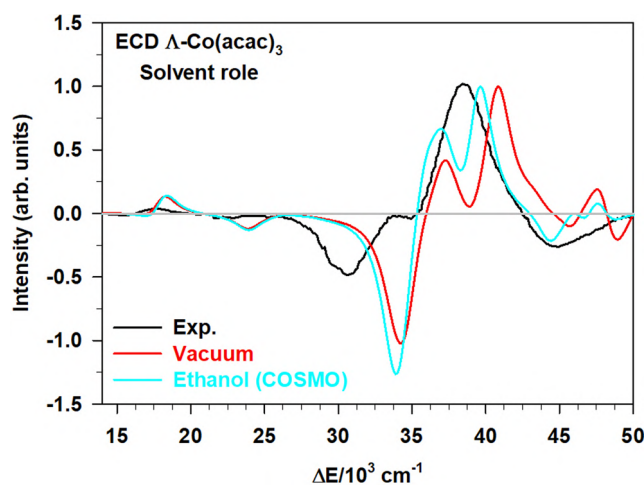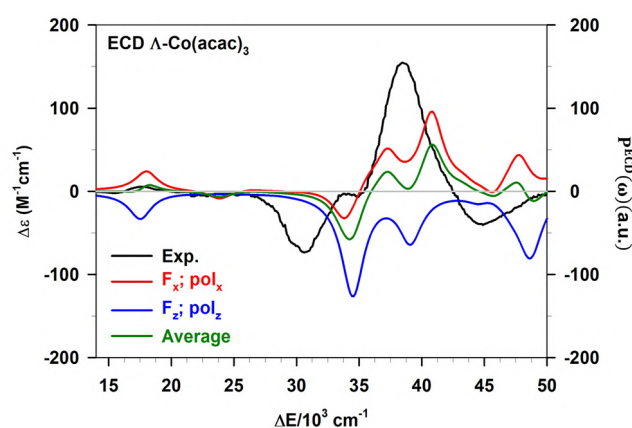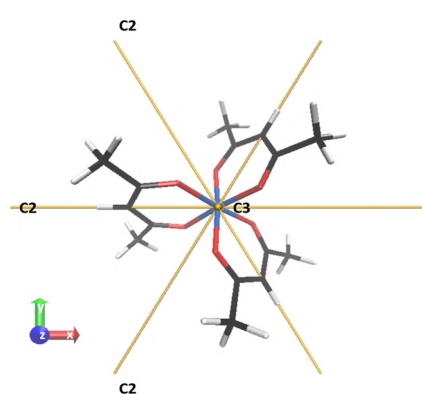

FIG. 8. Comparison between the experimental (solid black line)[76] and real-time calculated (solid red line) ECD spectra of the Λ-Co(acac)$_3$ complex. All the intensities have been normalized and reported in arbitrary units (arb. units).

FIG. 9. The experimental ECD (solid black line)[76] spectrum of the Λ-Co(acac)$_3$ complex is reported together with the real-time ECD calculated in the gas phase (solid red line) and ethanol (solid cyan line), respectively. The solvent has been implicitly treated with the COSMO model available in AMS. All the intensities have been normalized and reported in arbitrary units (arb. units).

particularly in the 35–45 $\times 10^3$ cm^{-1} energy range where the double peak is well aligned with the corresponding experimental feature.

It is worth noting that a previous theoretical study on the same Co complex[85] furnished an ECD spectrum in satisfactory agreement with respect to the experiment,[76] despite an energy shift of 4×10^3 cm^{-1}. Considering that our calculations have not been shifted, the agreement of the present results is slightly better than the previous ones, with especially the features at 17×10^3 and 39×10^3 cm^{-1} improved by theory. Such an improvement should be ascribed to the more accurate hybrid exchange-correlation functional, which

performs better than the GGA-BP86 functional used by Fan and Ziegler.[85]

We concluded our investigation on the Λ-Co(acac)$_3$ complex analyzing how each polarization direction of the external field $\vec{F}$ influences the temporal evolution of the corresponding component of the magnetic transition dipole moment and, thus, the ECD. Indeed, while the spectrum is usually shown, or directly obtained, as averaged on the three components, the contributions can be decomposed in the WaveT package since they are calculated independently. Considering, in particular, the Λ-Co(acac)$_3$ with its D$_3$ symmetry (see the left panel of Fig. 10), it is interesting to perform a directional analysis. The ECD spectra reported in Fig. 10 (right panel) confirm the ability of our calculations to discriminate among the components, with the magnetic response to F_z along z different from those to F_x and F_y along x and y, respectively.

V. CONCLUSIONS

We have implemented an explicit interface between the AMS and WaveT codes in order to propagate the magnetic moment in the space of the electronic states. In this way, we have defined a general approach to compute the ECD spectra via TDSE, using the magnetic response to an electric perturbation.

We have applied this scheme to the (R)-/(S)-methyloxirane, L-alanine, and Λ-Co(acac)$_3$ chiral systems to validate its correct functioning. The frequency-domain (AMS) and time-domain (WaveT) ECD calculations are in excellent agreement, hence confirming the correct simulation of the temporal evolution of the magnetic transition dipole moment. Furthermore, for each system, the comparison between the real-time ECD spectra and the experimental data shows the good overall quality of our calculations, which all reproduce the main spectral features. Instead, for the cobalt complex, the differences between the experimental ECD, measured in an ethanol solution, and our propagation in vacuum can be related to the environmental conditions. This hypothesis has been corroborated with the static inclusion of the solvent in the wave function

FIG. 10. Left panel: schematic representation of the D$_3$ symmetry of the Λ-Co(acac)$_3$ complex with the C$_3$ axis coincident with the z axis. Right panel: the experimental ECD (solid black line) spectrum of the Λ-Co(acac)$_3$ complex is reported together with those simulated considering the x (solid red line) and z (solid blue line) directions of the external field $\vec{F}$, respectively, and the real-time ECD averaged on the three components (solid dark green line). The experimental and calculated spectra are reported in M^{-1} cm^{-1} and atomic units (a.u.), respectively, without any normalization.

time evolution, which provides a better agreement between the experimental and time-domain ECD. In addition to this successful validation, we have investigated the excited state spectra and the relation between the $\vec{F}$ polarization and the ECD response. Propagating the wave packet of (S)-methyloxirane in its first excited state but with the ground state geometry, we have pointed out the significant changes that occur in the optical response following the system excitation. Indeed, while the absorption spectrum is entirely shifted toward the near IR/vis range, the ECD one reveals a sign inversion. For the symmetric Λ-Co(acac)$_3$, instead, we have decomposed the rotationally averaged ECD in its components, highlighting the optical changes in response to a different polarization of the external field.

SUPPLEMENTARY MATERIAL

See the supplementary material for the convergence tests of real-time propagations, the CAM-B3LYP ECD spectra of (R)- and (S)-methyloxirane, the statistical weights of the L-alanine conformers at 300 K, the ECD spectra of the L-alanine higher-energy conformers, and the CAM-B3LYP ECD spectra of Λ-Co(acac)$_3$.

ACKNOWLEDGMENTS

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

M. Monti: Formal analysis (equal); Investigation (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **M. Stener**: Writing – review & editing (equal). **E. Coccia**: Conceptualization (lead); Investigation (lead); Methodology (lead); Supervision (lead); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] P. Lazzeretti, Adv. Chem. Phys. **75**, 507 (1989).

[2] P. Lazzeretti, *Handbook of Molecular Physics and Quantum Chemistry* (Wiley, 2003), Vol. 3, p. 53.

[3] C. Diedrich and S. Grimme, J. Phys. Chem. A **107**, 2524 (2003).

[4] T. D. Crawford, Theor. Chem. Acc. **115**, 227 (2006).

[5] T. D. Crawford, M. C. Tam, and M. L. Abrams, J. Phys. Chem. A **111**, 12057 (2007).

[6] I. Warnke and F. Furche, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 150 (2012).

[7] S. S. Andrews and J. Tretton, J. Chem. Educ. **97**, 4370 (2020).

[8] N. Berova, L. D. Bari, and G. Pescitelli, Chem. Soc. Rev. **36**, 914 (2007).

[9] M. Scott, D. R. Rehn, P. Norman, and A. Dreuw, J. Phys. Chem. Lett. **12**, 5132 (2021).

[10] N. Niemeyer, M. Caricato, and J. Neugebauer, J. Chem. Phys. **156**, 154114 (2022).

[11] D. M. Rogers, H. Do, and J. D. Hirst, Mol. Phys. **120**, e2133748 (2022).

[12] M. Schreiber, R. Vahrenhorst, V. Buss, and M. P. Fülscher, Chirality **13**, 571 (2001).

[13] F. Gendron, B. Moore II, O. Cador, F. Pointillart, J. Autschbach, and B. Le Guennic, J. Chem. Theory Comput. **15**, 4140 (2019).

[14] A. E. Hansen and T. D. Bouman, J. Am. Chem. Soc. **107**, 4828 (1985).

[15] T. B. Pedersen, H. Koch, and K. Ruud, J. Chem. Phys. **110**, 2883 (1999).

[16] J. Kongsted, T. B. Pedersen, A. Osted, A. E. Hansen, K. V. Mikkelsen, and O. Christiansen, J. Phys. Chem. A **108**, 3632 (2004).

[17] D. A. Fedotov, S. Coriani, and C. Hättig, J. Chem. Phys. **154**, 124110 (2021).

[18] J. H. Andersen, K. D. Nanda, A. I. Krylov, and S. Coriani, J. Chem. Theory Comput. **18**, 1748 (2022).

[19] G. Bringmann, K.-P. Gulden, B. Holger, J. Fleischhauer, B. Kramer, and E. Zobel, Tetrahedron **49**, 3305 (1993).

[20] S. Grimme and C. Bannwarth, J. Chem. Phys. **145**, 054103 (2016).

[21] D. Varsano, L. A. Espinosa-Leal, X. Andrade, M. A. L. Marques, R. di Felice, and A. Rubio, Phys. Chem. Chem. Phys. **11**, 4481 (2009).

[22] S. Pipolo, S. Corni, and R. Cammi, Comput. Theor. Chem. **1040-1041**, 112 (2014).

[23] J. J. Goings and X. Li, J. Chem. Phys. **144**, 234102 (2016).

[24] J. Mattiat and S. Luber, Chem. Phys. **527**, 110464 (2019).

[25] J. Mattiat and S. Luber, J. Chem. Theory Comput. **18**, 5513 (2022).

[26] L. Konecny, M. Kadek, S. Komorovsky, K. Ruud, and M. Repisky, J. Chem. Phys. **149**, 204104 (2018).

[27] E. Makkonen, T. P. Rossi, A. H. Larsen, O. Lopez-Acevedo, P. Rinke, M. Kuisma, and X. Chen, J. Chem. Phys. **154**, 114102 (2021).

[28] W. D. Liu, J. Q. Wang, S. F. Yuan, X. Chen, and Q. M. Wang, Angew. Chem., Int. Ed. Engl. **60**, 11430 (2021).

[29] O. Baseggio, D. Toffoli, G. Fronzoni, M. Stener, L. Sementa, and A. Fortunelli, J. Phys. Chem. C **120**, 24335 (2016).

[30] L. Chang, O. Baseggio, L. Sementa, D. Cheng, G. Fronzoni, D. Toffoli, E. Aprà, M. Stener, and A. Fortunelli, J. Chem. Theory Comput. **14**, 3703 (2018).

[31] D. Toffoli, A. Russi, G. Fronzoni, E. Coccia, M. Stener, L. Sementa, and A. Fortunelli, J. Phys. Chem. Lett. **12**, 5829 (2021).

[32] M. Monti, M. Stener, and M. Aschi, J. Comput. Chem. **43**, 2023 (2022).

[33] D. Toffoli, M. Medves, G. Fronzoni, E. Coccia, M. Stener, L. Sementa, and A. Fortunelli, Molecules **27**, 93 (2022).

[34] H. Jang, N. J. Kim, and J. Heo, Comput. Theor. Chem. **1125**, 63 (2018).

[35] E. Molteni, G. Onida, and G. Tiana, J. Phys. Chem. B **119**, 4803 (2015).

[36] J. Autschbach, T. Ziegler, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **116**, 6930 (2002).

[37] J. Autschbach and T. Ziegler, J. Chem. Phys. **116**, 891 (2002).

[38] J. Autschbach, S. Patchkovskii, T. Ziegler, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **117**, 581 (2002).

[39] S. Pipolo and S. Corni, J. Phys. Chem. C **120**, 28774 (2016).

[40] E. Coccia, J. Fregoni, C. A. Guido, M. Marsili, S. Pipolo, and S. Corni, J. Chem. Phys. **153**, 200901 (2020).

[41] E. Coccia, F. Troiani, and S. Corni, J. Chem. Phys. **148**, 204112 (2018).

[42] E. Coccia and S. Corni, J. Chem. Phys. **151**, 044703 (2019).

[43] E. Coccia, Mol. Phys. **118**, e1769871 (2020).

[44] P. Grobas Illobre, M. Marsili, S. Corni, M. Stener, D. Toffoli, and E. Coccia, J. Chem. Theory Comput. **17**, 6314 (2021).

[45] G. Dall'Osto, E. Coccia, C. A. Guido, and S. Corni, Phys. Chem. Chem. Phys. **22**, 16734 (2020).

[46]G. Dall'Osto and S. Corni, J. Phys. Chem. A **126**, 8088 (2022).

[47]J. C. Tremblay, T. Klamroth, and P. Saalfrank, J. Chem. Phys. **129**, 084302 (2008).

[48]J. C. Tremblay, S. Klinkusch, T. Klamroth, and P. Saalfrank, J. Chem. Phys. **134**, 044311 (2011).

[49]P. Saalfrank, F. Bedurke, C. Heide, T. Klamroth, S. Klinkusch, P. Krause, M. Nest, and J. C. Tremblay, Adv. Quantum Chem. **81**, 15 (2020).

[50]F. Bedurke, T. Klamroth, and P. Saalfrank, Phys. Chem. Chem. Phys. **23**, 13544 (2021).

[51]E. Coccia, B. Mussard, M. Labeye, J. Caillat, R. Taïeb, J. Toulouse, and E. Luppi, Int. J. Quantum Chem. **116**, 1120 (2016).

[52]E. Coccia and E. Luppi, Theor. Chem. Acc. **135**, 43 (2016).

[53]E. Luppi and M. Head-Gordon, Mol. Phys. **110**, 909 (2012).

[54]E. Luppi and M. Head-Gordon, J. Chem. Phys. **139**, 164121 (2013).

[55]M. Labeye, F. Zapata, E. Coccia, V. Véniard, J. Toulouse, J. Caillat, R. Taïeb, and E. Luppi, J. Chem. Theory Comput. **14**, 5846 (2018).

[56]E. Coccia and E. Luppi, Theor. Chem. Acc. **138**, 96 (2019).

[57]A. F. White, C. J. Heide, P. Saalfrank, M. Head-Gordon, and E. Luppi, Mol. Phys. **114**, 947 (2016).

[58]E. Coccia, R. Assaraf, E. Luppi, and J. Toulouse, J. Chem. Phys. **147**, 014106 (2017).

[59]J. A. Sonk, M. Caricato, and H. B. Schlegel, J. Phys. Chem. A **115**, 4678 (2011).

[60]J. A. Sonk and H. B. Schlegel, J. Phys. Chem. A **115**, 11832 (2011).

[61]H. B. Schlegel, S. M. Smith, and X. Li, J. Chem. Phys. **126**, 244110 (2007).

[62]N. Tancogne-Dejean, M. J. T. Oliveira, X. Andrade, H. Appel, C. H. Borca, G. Le Breton, F. Buchholz, A. Castro, S. Corni, A. A. Correa *et al.*, J. Chem. Phys. **152**, 124119 (2020).

[63]T. P. Rossi, M. Kuisma, M. J. Puska, R. M. Nieminen, and P. Erhart, J. Chem. Theory Comput. **13**, 4779 (2017).

[64]S. A. Fischer, C. J. Cramer, and N. Govind, J. Chem. Theory Comput. **11**, 4294 (2015).

[65]P. Hoerner, M. K. Lee, and H. B. Schlegel, J. Chem. Phys. **151**, 054102 (2019).

[66]R. Rüger, M. Franchini, T. Trnka, A. Yakovlev, E. van Lenthe, P. Philipsen, T. van Vuren, B. Klumpers, and T. Soini, AMS 2022.1, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, 2022, http://www.scm.com.

[67]R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1989), Vol. 1, p. 1989.

[68]G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).

[69]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[70]C. Meinert, A. D. Garcia, J. Topin, N. C. Jones, M. Diekmann, R. Berger, L. Nahon, S. V. Hoffmann, and U. J. Meierhenrich, Nat. Commun. **13**, 502 (2022).

[71]M. E. Casida, *Recent Advances in Density Functional Methods: (Part I)* (World Scientific, 1995), pp. 155–192.

[72]T. Yanai, D. P. Tew, and N. C. Handy, Chem. Phys. Lett. **393**, 51 (2004).

[73]A. Kumar, S. E. Toal, D. DiGuiseppi, R. Schweitzer-Stenner, and B. M. Wong, J. Phys. Chem. B **124**, 2579 (2020).

[74]A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2 **1993**, 799–805.

[75]M. Carnell, S. D. Peyerimhoff, A. Breest, K. H. Gödderz, P. Ochmann, and J. Hormes, Chem. Phys. Lett. **180**, 477 (1991).

[76]R. C. Fay and R. B. Von Dreele, J. Am. Chem. Soc. **93**, 4936 (1971).

[77]K. Hiramatsu and T. Nagata, J. Chem. Phys. **143**, 121102 (2015).

[78]V. Stadnytskyi, G. S. Orf, R. E. Blankenship, and S. Savikhin, Rev. Sci. Instrum. **89**, 033104 (2018).

[79]M. Oppermann, B. Bauer, T. Rossi, F. Zinna, J. Helbing, J. Lacour, and M. Chergui, Optica **6**, 56 (2019).

[80]M. Kuronuma, T. Sato, Y. Araki, T. Mori, S. Sakamoto, Y. Inoue, O. Ito, and T. Wada, Chem. Lett. **48**, 357 (2019).

[81]M. Schmid, L. Martinez-Fernandez, D. Markovitsi, F. Santoro, F. Hache, R. Improta, and P. Changenet, J. Phys. Chem. Lett. **10**, 4089 (2019).

[82]A. Sharma, S. Athanasopoulos, E. Kumarasamy, C. Phansa, A. Asadpoordarvish, R. P. Sabatini, R. Pandya, K. R. Parenti, S. N. Sanders, D. R. McCamey *et al.*, J. Phys. Chem. A **125**, 7226 (2021).

[83]M. Cao, S. Q. Newton, J. Pranata, and L. Schäfer, J. Mol. Struct.: THEOCHEM **332**, 251 (1995).

[84]S. Blanco, A. Lesarri, J. C. López, and J. L. Alonso, J. Am. Chem. Soc. **126**, 11675 (2004).

[85]J. Fan and T. Ziegler, Inorg. Chem. **47**, 4762 (2008).

Supplementary Material: Electronic circular dichroism from real-time propagation in state space

M. Monti, M. Stener, and E. Coccia[*]

Dipartimento di Scienze Chimiche e Farmaceutiche,

Università di Trieste, via L. Giorgieri 1, 34127, Trieste, Italy

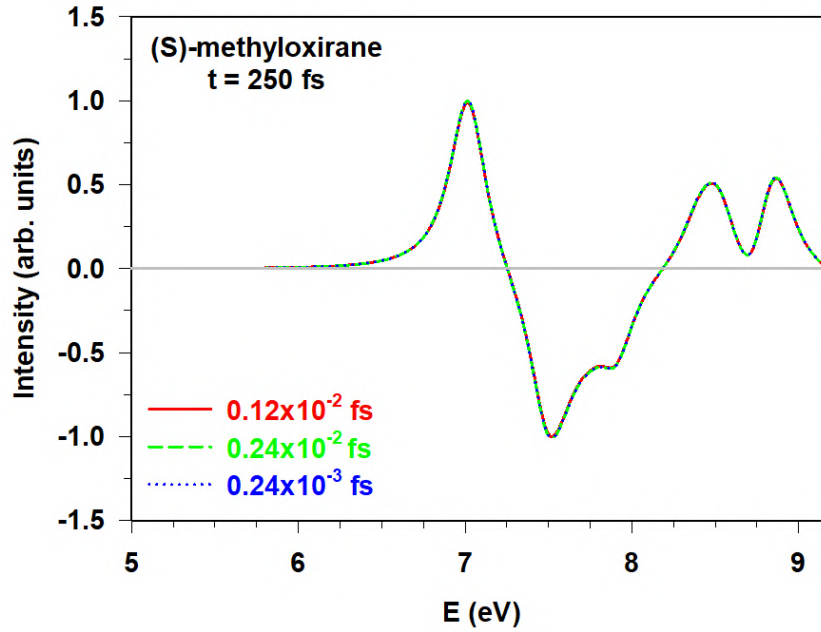[*]Electronic address: ecoccia@units.it

FIG. S1: Time-domain ECD spectra of (S)-methyloxirane calculated for 250 fs with a time step δt of 0.12×10^{-2} (solid red line), 0.24×10^{-2} (dashed green line), and 0.24×10^{-3} (dotted blue line) fs.

Conformation	ΔE (kJ/mol)	$P(i)_{norm}$ (T=300 K)
1	0.0	0.60
2	4.7	0.092
3	4.9	0.083
4	5.1	0.077
5	5.5	0.067
6	5.8	0.058
7	8.3	0.022

TABLE S1: Conformation, relative energy and normalized probability values ($\sum_{i=1}^{7} P(i)_{norm} = 1$) for the 7 stable structures investigated in the work by Meinert *et al.* [1].
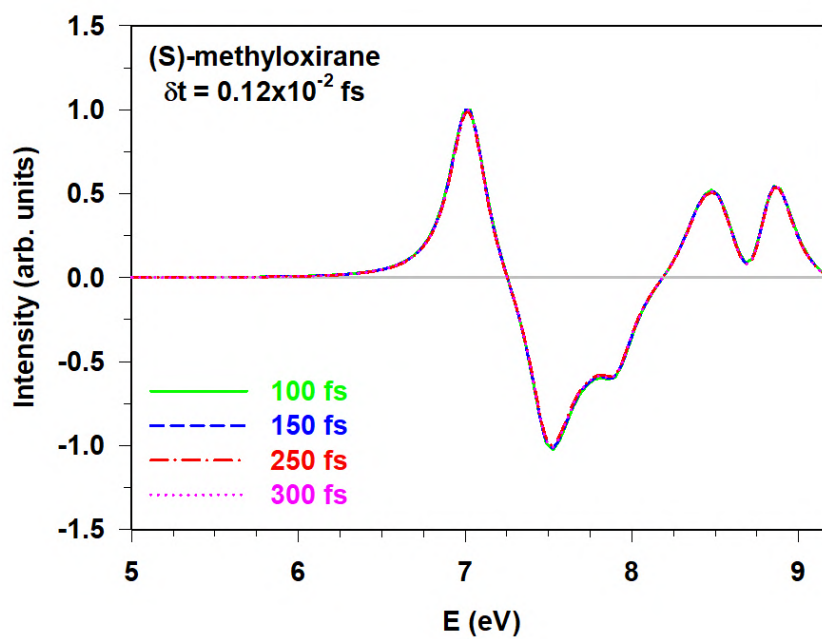
FIG. S2: Time-domain ECD spectra of (S)-methyloxirane calculated for 100 (solid green line), 150 (dashed blue line), 250 fs (dashed and dotted red line), and 300 (dotted pink line) fs with a time step δt of 0.12×10^{-2} fs.
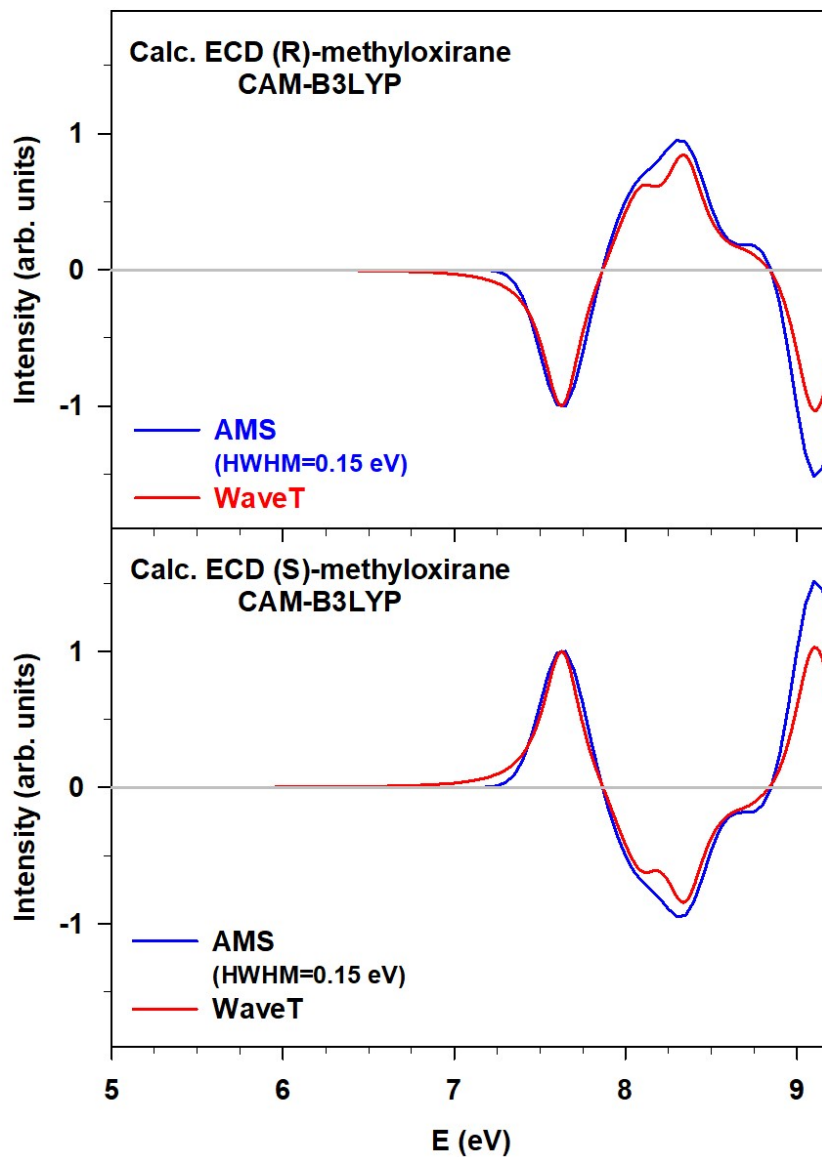
FIG. S3: Comparison of the ECD spectra calculated for the (R)-, (top panel) and (S)- (bottom panel) methyloxirane using the CAM-B3LYP functional and the AMS (frequency-domain, solid blue line) and WaveT (time-domain, solid red line) code. All the intensities have been normalized and reported in arbitrary units (arb. units).

FIG. S4: Comparison between the experimental (solid black line) and calculated with WaveT (CAM-B3LYP) (solid red line) ECD spectra of the (R)-, (top panel) and (S)- (bottom panel) methyloxirane. All the intensities have been normalized and reported in arbitrary units (arb. units).
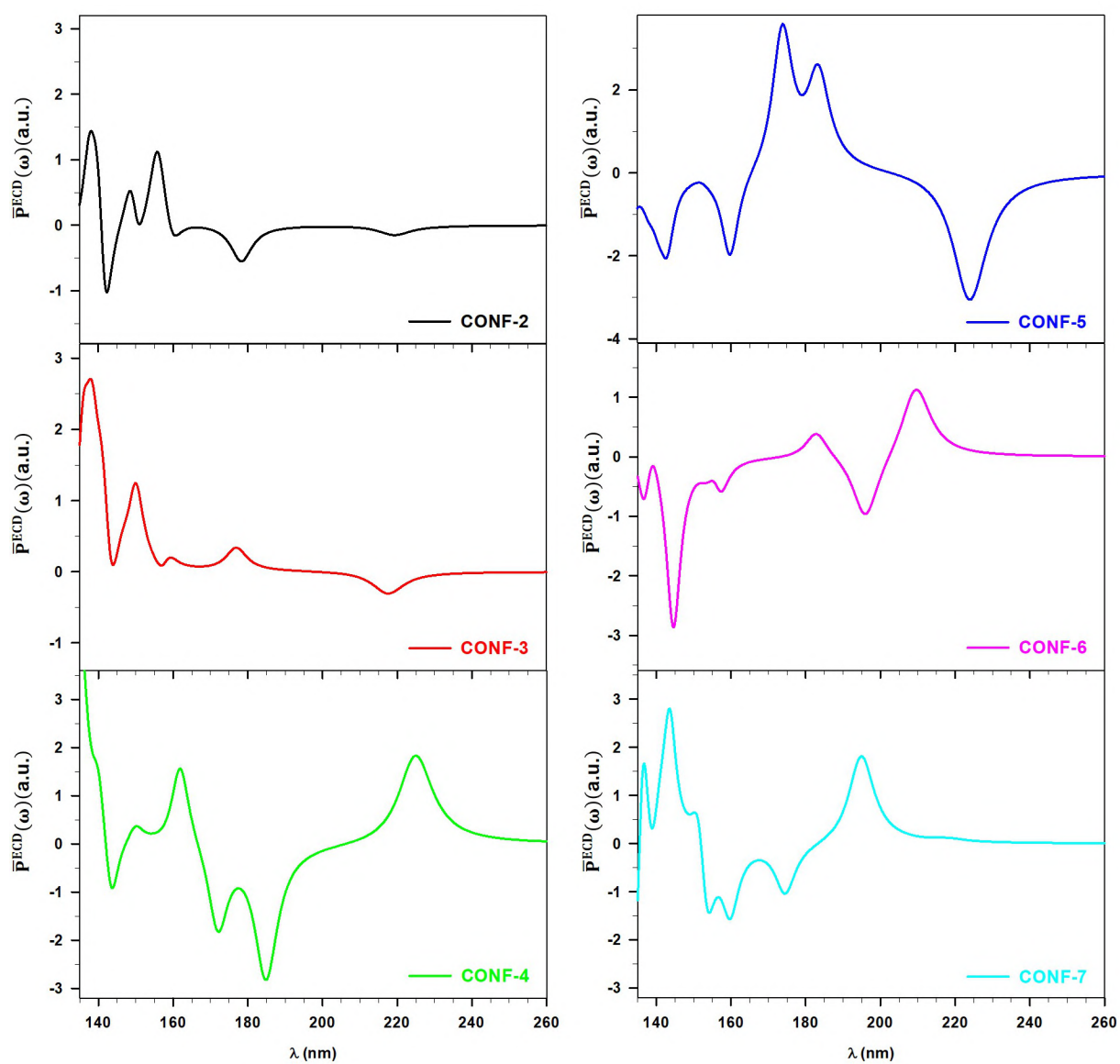
FIG. S5: Individual time-domain ECD spectra for the L-alanine higher-energy conformers (from 2 to 7). All the intensities have been reported in atomic units (a.u.)

6
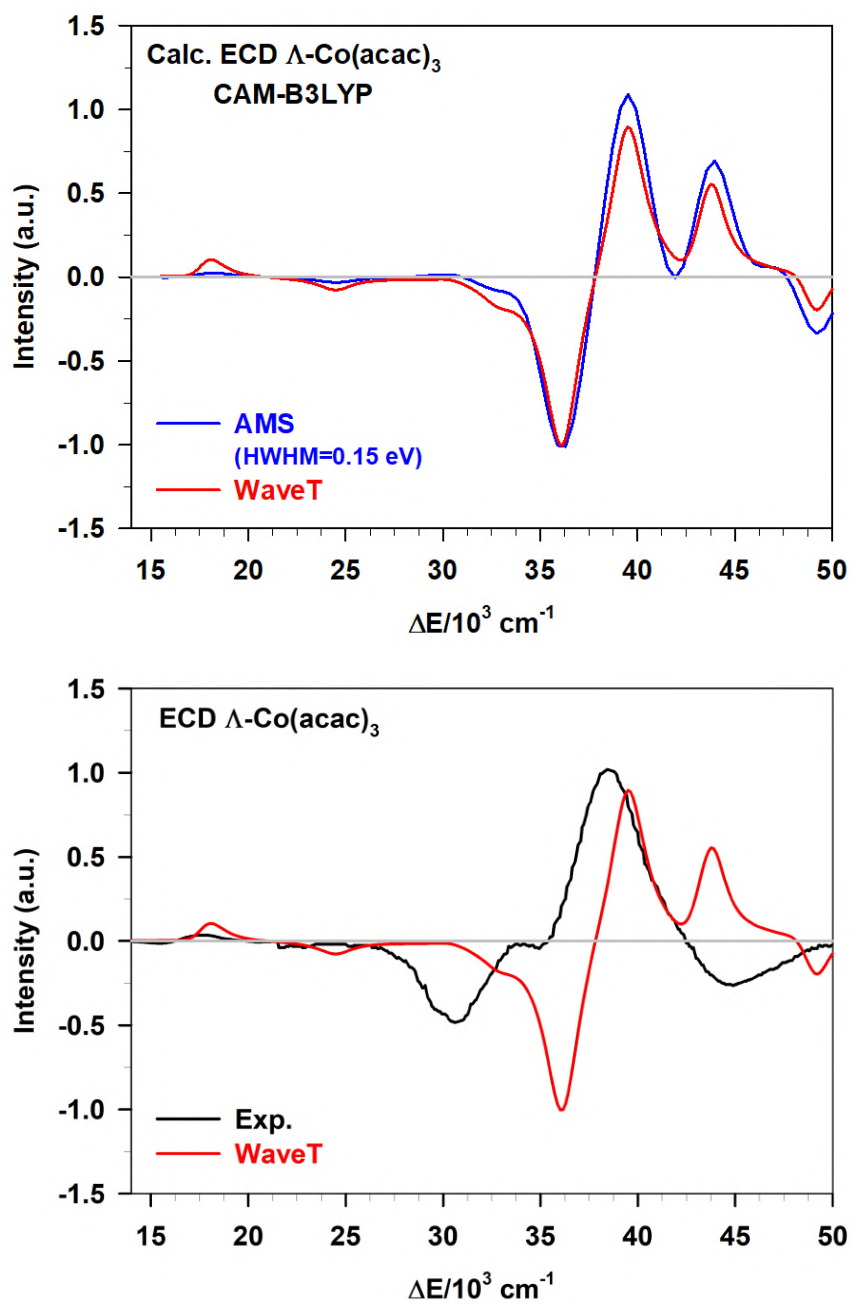
FIG. S6: Top panel: Comparison of the ECD spectra calculated for the Λ-Co(acac)$_3$ using the CAM-B3LYP functional and the AMS (frequency-dom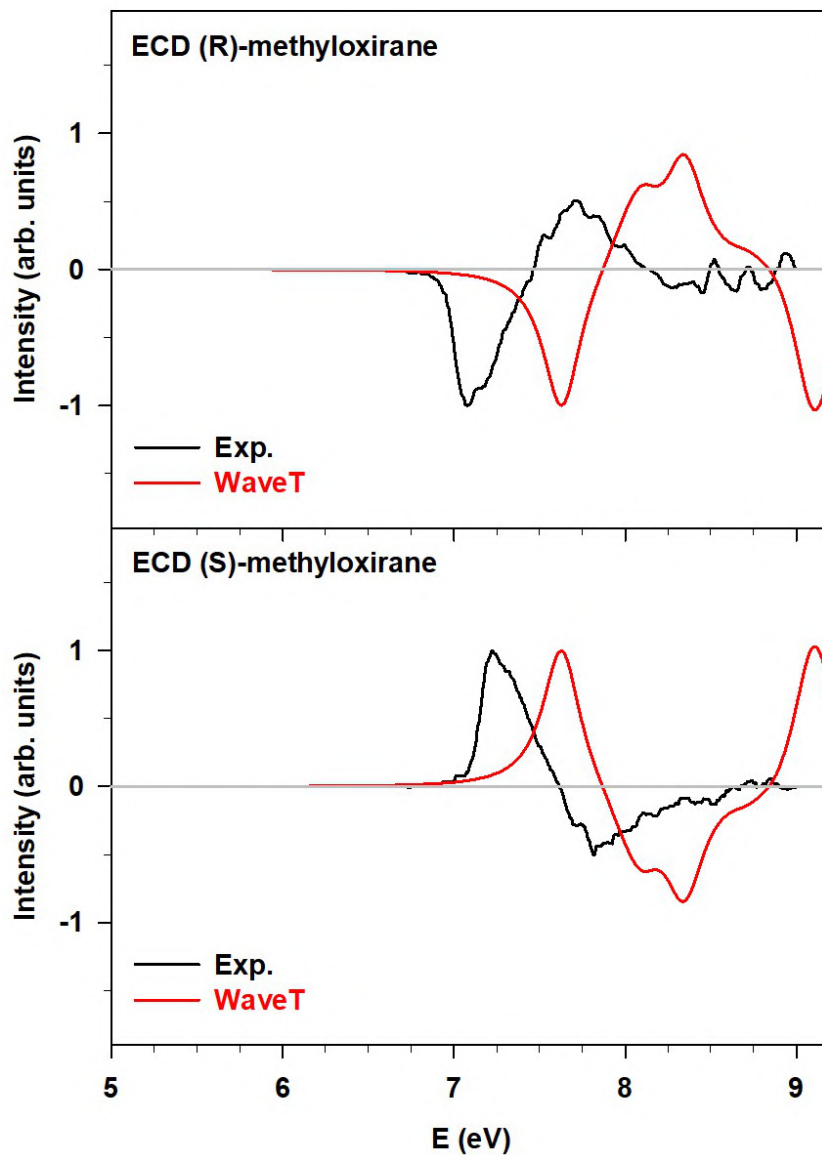ain, solid blue line) or WaveT (time-domain, solid red line) code. Bottom panel: Comparison between the experimental (solid black line) and calculated with WaveT (CAM-B3LYP) (solid red line) ECD spectra of Λ-Co(acac)$_3$. All the intensities have been normalized and reported in arbitrary units (arb. units).

[1] C. Meinert, A. D. Garcia, J. Topin, N. C. Jones, M. Diekmann, R. Berger, L. Nahon, S. V. Hoffman, and U. J. Meierhenrich, Nat. Commun. **13**, 502 (2022).

4.2 Additional Publications

Along with the main activity, it is worth mentioning some additional projects that were carried out over these three years. All the studies are still focused on optical properties, spanning from valence-shell spectroscopies, as photoabsorption, to core-shell ones, as X-ray Photoelectron (XPS) and Near Edge X-ray Absorption Fine Structure (NEXAFS) spectroscopy. The resulting publications are reported below.

4.2.1 Theoretical Investigation of Photoinduced Processes in Subnanometer Oxide-Supported Metal Catalysts

The project goal was to investigate the optical absorption and photo-decay process in two subnanometer metal complexes supported on MgO(100), namely, $Ag_3(HCO_3)(C_2H_4)_2(O)$ and $Ag_3(CO_2F)(C_2H_4)_2(O)$. These aggregates were selected as derivatives of $Ag_3(CO_3)(C_2H_4)_2(O)$, which has revealed to play a key role in the partial oxidation of ethylene to ethylene epoxide catalyzed by $Ag_3/MgO(100)$.

In order to calculate both the photoabsorption spectra of the aggregates and the evolution of their excited states, we first built suitable cluster models taking into account the effect of the charge-separated metal oxide support. Practically, the MgO(100) cluster is divided into three regions as follows: (i) ions that are directly involved into the chemical process, or are close enough to the adsorbed metal complex, are treated explicitly, (ii) Mg^{2+} cations at the interface are treated in a hybrid way, and (iii) both Mg^{2+} and O^{2-} far enough from the reacting regions are treated as pure charges. Such models were built as a compromise between accuracy and computational cost, compatibly with the AMS software which does not work with PBC. Afterwards, the photoabsorption spectra were calculated by means of TDDFT in its standard Casida's formulation. From the TDDFT spectrum, a specific excited state was then selected for a given geometry optimization with the Broyden-Fletcher-Goldfarb-Shanno algorithm. My contribution to this project was focused on implementing the code that allows us to build the suitable MgO(100) clusters and the calculation of TDDFT spectra.

Thanks to this protocol, designed in collaboration with Dr. Alessandro Fortunelli and collaborators, it was shown that complexes containing bicarbonate and fluorocarbonate groups lead to cluster disaggregation and/or ligand detachment. Thus, an optical frailty of these subnanometer systems is revealed. Additionally, the nature of the given excitations was correlated to the corresponding photo-induced reaction products by analyzing Overlap Population Density of States (OP-DOS) plots, as well as geometrical features of the MOs involved in the excitations. This way, all the methodological tools needed to investigate this novel field are provided.

Article

Theoretical Investigation of Photoinduced Processes in Subnanometer Oxide-Supported Metal Catalysts

Published as part of The Journal of Physical Chemistry virtual special issue "Metal Clusters, Nanoparticles, and the Physical Chemistry of Catalysis".

Luca Sementa, Marta Monti, Daniele Toffoli, Alvaro Posada-Amarillas, Mauro Stener,* and Alessandro Fortunelli*

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report a computational study and analysis of the optical absorption and photodecay processes in two subnanometer metal complexes deposited on an oxide support, the regular MgO(100) surface: (i) $Ag_3(HCO_3)(C_2H_4)_2(O)$ and (ii) $Ag_3(CO_2F)(C_2H_4)_2(O)$. These aggregates are chosen as derivatives of a $Ag_3(CO_3)(C_2H_4)_2(O)$ ligand/metal-cluster/support complex, previously singled out as a key intermediate in the path of ethylene partial oxidation to ethylene epoxide catalyzed by $Ag_3/$ MgO(100), and serve as model systems to investigate photochemical phenomena in ligand/metal-cluster/support complexes by subnanometer metal catalysts, an appealing field for future research. After generating optimized initial configurations and building cluster models that take properly into account the effect of the charge-separated oxide support, we use time-dependent density-functional theory (TDDFT) to determine first the photoabsorption spectra of the two aggregates and then to follow the evolution of their excited states in the optical region. We show that complexes containing such bicarbonate and fluorocarbonate groups are sensitive to optical adsorption, often leading to ligand detachment and/or cluster disaggregation, thus pointing to an "optical frailty" of these subnanometer cluster species, possibly rationalizing previous experimental observations. Additionally, we correlate the nature of the given excitations and of the corresponding photoinduced reaction products via an analysis of overlap population-density of states (OP-DOS), geometric parameters, and spatial distribution of the molecular orbitals involved in the excitation, thus providing the set of methodological tools needed to explore this novel field.
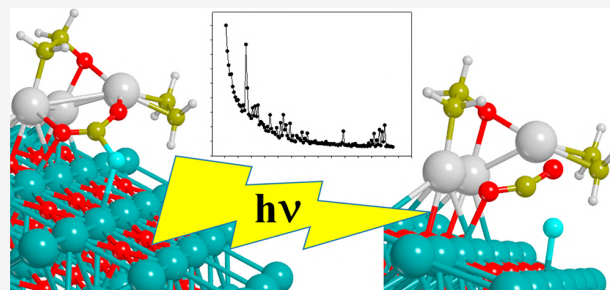
1. INTRODUCTION

The field of subnanometer metal clusters (also called ultrananoclusters, as they will be referred to hereafter), i.e., very small (1−20 atoms) metal clusters either free or deposited on a support, presents attractive features in several fields including catalysis, from both a scientific and a technological perspective.[1−4] In the last 2 decades dedicated experiments have shown that supported ultrananoclusters (in particular, oxide-supported) can be the true actors of catalysis with different and superior catalytic behavior with respect to larger nanoparticles,[1−4] therefore being of interest in all those cases in which an impasse is faced in the use of traditional systems and new catalytic species are required and actively sought for. In addition to challenging but exciting scientific problems,[5] appealing possibilities in terms of practical applications are offered by the combination of novel reaction mechanisms,[4] precise control of the catalyst's size and composition,[3] and an atom-economic use of metals (often precious or endangered elements),[6] together with the fact that these systems can be

thoroughly characterized and investigated by advanced computational methods,[7,8] thus achieving a high degree of basic knowledge and understanding. Among the conclusions reached, of particular relevance is the idea that the catalytically active species in heterogeneous ultrananocatalysis are not the bare metal clusters but "ligand/metal-cluster/support catalytic complexes", i.e., complex aggregates generated *in situ* under reaction conditions by the metal cluster, the support, and the ligands (reactants, intermediates, products), typically in the high coverage regime but still able to absorb and chemically transform incoming species.[7] In addition to chemical ones, the optical properties of some ultrananoclusters have also been
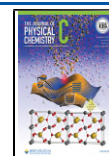
studied and predicted also via computational methods, extensively in the case of free clusters[9−12] and in the more recently investigated field of ligand-protected atomically precise clusters,[13−16] including dynamic processes.[17−22] Much less is known instead on the optical properties of ultrananoclusters deposited on a support, although few studies have appeared in the literature.[23−25]

In parallel in recent years, photochemistry has developed at an equally impressive pace,[26] so that light-induced processes, including chemical reactions using sensitizers[27] and photo-catalysts,[28] are nowadays extensively utilized even at the industrial level.[29−36] Unfortunately, although detailed mechanistic studies exist for some of the simpler reactions,[37] many photochemical and photocatalytic reactions are still often employed in a heuristic way without mastering mechanistic steps. A prevailing opinion in the field holds that an improved atomic-level understanding of light-induced photochemical and photocatalytic processes is needed to extend the scope and reliability of these methods beyond current limits and that a fundamental knowledge of light absorption and evolution processes should lay the ground to the rational design of photoreactions.[38]

In this situation, it appears as an intriguing possibility to combine the potentialities of supported ultrananocatalysts and photochemistry, therefore achieving systems with both novel and potentially well understood photochemical mechanisms, which could have a variety of practical applications as sensitizers, photocatalysts, etc. To make this lap, a number of steps are needed at both the experimental and theoretical levels. In particular, at the theory level methods should be developed and tested to predict the optics and photochemistry of supported ultrananoclusters and their catalytically active ligand/cluster/support complexes, which is at present a largely unexplored field.

In the present work, we address this problem and make a step in this direction to lay the ground for future progress. We take an example of supported ultrananocatalyst and its corresponding ligand/cluster/support catalytic complex from previous studies on the partial oxidation of ethylene to ethylene epoxide over a $Ag_3(CO_3)$ cluster catalyst deposited on the MgO(100) surface (which ensures additional stability to the catalyst),[39,40] and we combine it with the recent possibility of accurately predicting absorption spectra of subnanometer clusters deposited on an oxide substrate[24] to explore the absorption properties and the fate of the absorbed photons on model but basically realistic ultrananocatalyst systems. Our goal in this exploratory study is to start drawing correlations between the initial excited state and the resulting photoinduced mechanism, a key step in view of understanding and controlling photochemical processes in this field.

The article is organized as follows. Section 2 describes the choice of investigated systems and the computational approach. Section 3 implements the approach, focusing attention on two distinct cases and on how the photochemical steps can be rationalized in terms of initial excitations. Section 4 summarizes our conclusions.

2. CHOICE OF INVESTIGATED SYSTEMS AND COMPUTATIONAL APPROACH

In previous studies,[7,39−41] we investigated oxidation processes over a Ag_3 cluster catalyst supported on the MgO(100) surface, including the oxidation of carbon monoxide to carbon dioxide ($CO + 1/2 O_2 \rightarrow CO_2$, or COox reaction)[41] and the partial oxidation of ethylene to ethylene epoxide ($C_2H_4 + 1/2 O_2 \rightarrow C_2H_4O$, or EtOx reaction).[7,39,40] We showed that under anhydrous (no moisture) COox conditions, the Ag_3 supported cluster transforms into a $Ag_3(CO_3)$ or silver-trimer/carbonate catalytic complex, which then acts as the COox catalytically active species.[41] We then considered using this $Ag_3(CO_3)$-(C_2H_4) complex as a catalyst in the EtOx reaction, still under anhydrous conditions. This is an example of and corresponds to the idea of using a ligand/cluster/support catalytic complex from one reaction as the catalyst in a different but akin reaction.[39] It should be stressed that $Ag_3(CO_3)$ does not need CO and O_2 to form, but it will be in any case generated directly *in situ* under EtOx conditions via reaction of gas-phase CO_2 (an unavoidable side-product in the oxidation of organic species) with $Ag_3(O)$ intermediates.[41] Moreover, the presence of the carbonate group in $Ag_3(CO_3)/MgO(100)$ is justified by the fact that bare $Ag_3/MgO(100)$ clusters are not stable against disaggregation or diffusion/sintering, as shown in refs 7 and 39−41 so that $Ag_3/MgO(100)$ cannot work as a catalyst if bare but needs stabilizing ligands. We showed that under EtOx conditions $Ag_3(CO_3)/MgO(100)$ first transforms into a $Ag_3(CO_3)(C_2H_4)$ complex. We then assessed the EtOx reaction cycle for this catalyst, involving the adsorption of an oxygen molecule (O_2) and an ethylene molecule (C_2H_4), the reaction between C_2H_4 and O_2 to evolve one ethylene epoxide (C_2H_4O) molecule, leaving behind a $Ag_3(CO_3)(C_2H_4)(O)$ aggregate, the adsorption of a third ethylene molecule next to the O adatom, and the transfer of the O adatom from one C_2H_4 molecule to the other with the final evolution of a second ethylene epoxide (C_2H_4O) molecule to regenerate the catalytic complex, see Figure 1e,f in ref 40 for a pictorial illustration of the catalytic cycle. For convenience of the reader we summarize and illustrate this previous work in the Supporting Information, Figures S1 and S2. The main point to be recalled here from these previous studies is that the rate-determining mechanistic step for this system is the last one here described, i.e., it involves as a starting species a $Ag_3(CO_3)(C_2H_4)_2(O)/MgO(100)$ aggregate, in which $Ag_3/MgO(100)$ is bound to a carbonate group (CO_3), two ethylene molecules ($C_2H_4)_2$, and an oxygen adatom (O) and corresponds to the transfer of the oxygen adatom to one of the ethylene molecules, with the final evolution of the C_2H_4O epoxide, as illustrated in Figure S2 of the Supporting Information.

We note that the $Ag_3(CO_3)(C_2H_4)_2(O)/MgO(100)$ aggregate is a spin-unpaired compound (spin doublet), with one unpaired electron mostly residing on the CO_3 residue.[40] Also due to technical reasons, in this work we limit ourselves to spin-restricted simulations. We therefore transformed the $Ag_3(CO_3)(C_2H_4)_2(O)/MgO(100)$ aggregate into two distinct spin-restricted analogues: (i) $Ag_3(CO_3H)(C_2H_4)_2(O)/MgO(100)$, where the carbonate group is replaced by a bicarbonate (Figure 1a), and (ii) $Ag_3(CO_2F)(C_2H_4)_2(O)/MgO(100)$, in which a fluorine atom (F) replaces one of the oxygen atoms of the carbonate (Figure 1b). This choice of investigated systems, in particular of the bicarbonate complex (Figure 1a,c), is connected with experimental evidence. In previous experiments, in fact, it has been demonstrated not only that carbonate species often play a crucial role in oxidation reactions of organic species on metal catalysts (Ag, Au, Pd, Pt), including the ethylene partial oxidation catalytic reaction (EtOx) over Ag systems,[42] which is not surprising considering that carbonate groups in general mediate silver-particles/organic-environment interfaces[43] but also that bicarbonates are
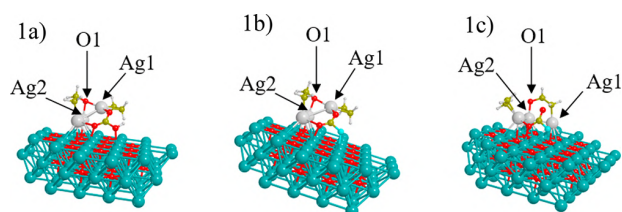
Figure 1. Pictorial illustrations of (a) the initial Ag$_3$(HCO$_3$)-(C$_2$H$_4$)$_2$(O) structure, (b) the initial Ag$_3$(CO$_2$F)(C$_2$H$_4$)$_2$(O) structure, and (c) the intermediate Ag$_3$(CO$_2$F)(C$_2$H$_4$)$_2$(O) structure (exhibiting an oxametallocycle adduct).

invariably associated with carbonate species under hydrous conditions and play an equally important role in the catalytic cycle. For example, for Au catalysts deposited on a wide set of both reducible and nonreducible oxides in the COox reaction, a commonly accepted mechanism grounded on kinetic and spectroscopic data foresees that stable carbonate species can form and poison the catalysts.[44,45] However, in the presence of water (hydrous conditions), either as a medium or in catalytic amounts, carbonate groups convert to some degree into bicarbonate groups, and this is decisive for the catalytic activity, because bicarbonates are assumed to decompose easily, thus hindering the carbonate poisoning effect and leading to recovery of catalytic activity under hydrous conditions.[46–51] Analogous observations on the importance of bicarbonate species have been reported for Pd[52–54] catalysts, while a poisoning effect of carbonates has been proposed also for subnanometer Pt catalysts.[55] Finally, it is important to underline that bicarbonate species have been claimed to assist or play an important role also in the reverse process to COox, the reduction reaction of carbon dioxide (CO2RR), especially to formate on Au, Pd, and Sn catalysts[56–60] (we will come back to this point below).

Considering then the novelty of these species, we first made sure that the Ag$_3$(CO$_3$H)(C$_2$H$_4$)$_2$(O)/MgO(100) and Ag$_3$(CO$_2$F)(C$_2$H$_4$)$_2$(O)/MgO(100) clusters were stable aggregates, as they indeed turned out to be (see subsection 2.1), and then we proceeded with the investigation of their photochemistry (see subsection 2.3). What we aim at investigating in the present work is the adsorption spectrum of such Ag$_3$(HCO$_3$)(C$_2$H$_4$)$_2$(O)/MgO(100) and Ag$_3$(CO$_2$F)-(C$_2$H$_4$)$_2$(O)/MgO(100) complex aggregate species as initial structures of the catalytic rate-determining step (so that they should be the most abundant surface species) and the fate of their optical excited states. As we will see in section 3, the complexes exhibiting bicarbonate and fluorocarbonate groups have somewhat modified chemical features and reactive possibilities with respect to their carbonate analogues. The interest of the present investigation is therefore both as proof-of-principle and methodological study and also in view of rationalizing and interpreting experimental evidence on bicarbonate species, as discussed in section 3.

2.1. Structure Generation of Periodic Models. We prepared two spin-restricted systems by modifying the carbonate group that anchors the Ag$_3$ nanocluster of the Ag$_3$CO$_3$(Et)$_2$O aggregate to the oxide support. In one case, we bonded one H atom to the oxygen atom farthest from the Ag$_3$ moiety; in another case, we substituted the same oxygen with a fluorine atom. The structures thus obtained were relaxed keeping the MgO support frozen, and the structural stability of the obtained local minima was tested by performing NVT

molecular dynamics (MD) lasting 2.5 ps at 400 K, without finding significant changes in the geometry, which confirm the stability of the chosen systems. Structure relaxations and MD were performed using periodic models and the CP2K software,[61–63] which allows one to predict the DFT-electronic structure of periodic systems thanks to the use of plane wave auxiliary basis sets. The code implements the Gaussian and plane wave approach (GPW),[63,64] which uses Gaussian localized basis sets to expand the electronic orbitals and an auxiliary set of plane waves to efficiently evaluate, exploiting periodic boundary conditions and the Hartree and exchange-correlation contributions to the DFT energy. Calculations were performed by employing the PBE-functional,[65] DZVP basis sets,[66] and a Fourier basis including plane waves whose kinetic energy was less than 300 Ry. Since we employed a large square 4 × 4 supercell of MgO-(100) with a lattice parameter of about 1.5 nm to reduce the interaction between replicated images of the clusters supported on the MgO surface, we could limit the sampling of the Brillouin zone[67] to the gamma point only without losing accuracy. We further reduced the computational burden by modeling the core electrons of each atom with norm-conserving GTH-pseudopotentials optimized for the PBE functional.[68] As for the MD setup, the equations of motion were integrated by using time a step of 1 fs, and the average temperature was controlled by employing a Noose–Hover thermostat.[69]

2.2. From Periodic to Cluster Models. While we used periodic models in the original studies[39,40] and in the structure generation step in section 2.1, and while these models are convenient to speed up the structural search, most of the tools we have available to describe the optical response are implemented within the ADF software,[70] which does not treat periodic boundary conditions. We therefore convert the periodic-model configurations obtained from the QE simulations into cluster-model configurations appropriate for ADF simulations in an accurate but also computationally efficient way. To this end, we use a protocol tested and validated in previous work,[24,71] by which only part of the system is described explicitly at the Quantum-Mechanical (QM) level. The protocol is based on the following steps. First, we build a large cluster representing the MgO(100) surface, extending by 35.768 × 35.768 Å in the directions parallel to the surface and 14.728 Å in the direction perpendicular to the surface so as that the final model is made of 2592 atoms. The atoms in this large cluster are then distinguished into three groups:

(i) a first group contains atoms sufficiently far from the adsorbed species that their contribution can be considered as purely electrostatic;

(ii) a second group of Mg^{2+} cations, interfacing the region in which the chemistry occurs, that are treated in a hybrid way;

(iii) atoms that take part directly in the chemical process or which are anyway close to the reacting species, which are explicitly included in the QM region.

The atoms in the first group are described as point charges, with charge $q = +2$ e for the Mg^{2+} cations and $q = -2$ e for the O^{2-} anions. The Mg^{2+} cations in the second group must be dealt with appropriately to ensure that they do not perturb the neighboring QM-described O^{2-} anions, and specifically one has to ensure that the tails of the wave function of the QM-described O^{2-} anions do not interact with the $q = +2$ charges of the Mg^{2+} cations generating spurious energy terms. This can

be achieved by describing these Mg^{2+} cations with centers in which a repulsive pseudopotential is positioned in addition to a $q = +2$ charge (see ref 71 and references therein) or by using a frozen-core or completely free Mg^{2+} wave function in addition to a $q = +2$ charge: the latter possibility is technically easier in the ADF code and is employed in the present work. Finally, the atoms in group (iii) are treated explicitly at the QM level. This last group of atoms must naturally include the metal cluster and the ligands but also some neighboring atoms of the support whose QM contribution must be explicitly taken into account. In detail, for an accurate reproduction of the energetics, it was found necessary[71] that all the atoms in the support that are at a bonding distance from the metal cluster and/or the ligands and additionally all the atoms in the support that are first-neighbors of the atoms of the support directly interacting with the reaction moiety be included into the QM simulation. This core region of explicit QM-level atoms also defines the size of the overall cluster model. Our protocol foresees in fact that, for a precise reproduction of the QE energetics, all the atoms explicitly described by a QM wave function (i.e., the adsorbed reacting species and their neighbors in the MgO support) must be at least 3 MgO units far from the boundary of the finite cluster model. Finally, we took special care at the level of both point-charge and explicit-atoms models to avoid *polar borders* or any uncompensated electrostatic contribution which may give rise to electrostatic divergences.[72] Once the model has been built, both the ground state and the excited states (see the next subsection) were optimized keeping the position of all the MgO ions frozen, while the atoms of the cluster+ligand moiety were left free to relax.

For convenience of the reader, report that the number of atoms in each group are 2491, 51, and 50 for the first, second and third group, respectively. Also, the full support including point charges is pictorially illustrated in Figure S3 of the Supporting Information. As previously discussed,[71] this large ensemble of point charges allows us to accurately represent the Madelung potential, still representing a minor part of the computational effort, since point charges only enter one-electron integrals.
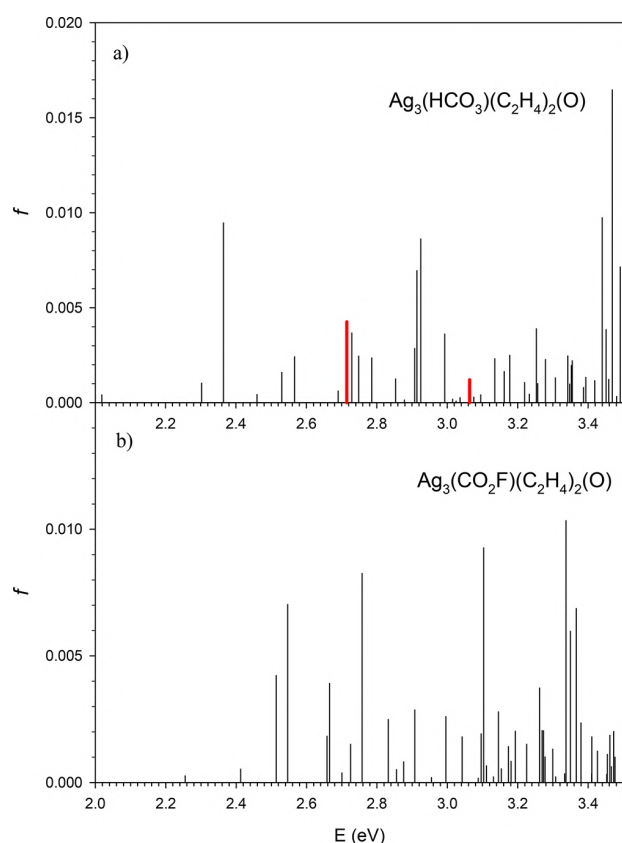
2.3. Excited State Simulations. Excitation spectra were calculated at the time-dependent DFT (TDDFT) level, employing the Casida formalism[73] as implemented in the ADF program.[70] An all-electron basis set consisting of Slater type orbitals (STO) of TZP quality was employed. Relativistic effects were treated within the zero order regular approximation (ZORA) formalism at the scalar relativistic (SR) level.[74] The PBE exchange-correlation functional[65] was employed both in the DFT Kohn−Sham equations as well as in the TDDFT kernel, which was approximated at the ALDA level.[75] Since we are interested in photochemical processes, we need to optimize the geometry of a series of excited states. The TDDFT spectrum was calculated and a specific excited state was selected for a given geometry optimization. We then employed the TDDFT formalism to calculate the energy and energy gradients and minimized the TDDFT energy to obtain the local energy minimum for each selected excited state, following the procedure developed by Ziegler as implemented in the ADF code.[76] For such geometry optimizations, the ground state equilibrium configuration was taken as the initial structure. Since during the geometry optimization process the excited states can change their relative energy ordering, we have followed each excited state during geometry optimization,

comparing the excited state nature for each geometry update and taking in the next step the excited state which was most similar to the previous one (and therefore recursively to the initial excited state). In other words, once the geometry is updated, a new TDDFT spectrum is calculated, and a comparison is done in terms of overlap between the DFT/KS wave functions of the new excited states with the DFT/KS wave function of the excited state of the previous step, and the excited state of the updated geometry is selected which has the maximum overlap with the original excited state. The geometry optimization method corresponds to a Broyden−Fletcher−Goldfarb−Shanno (BFGS) algorithm.

3. RESULTS AND DISCUSSION

The catalytic conversion of ethylene to ethylene oxide (ethylene partial oxidation or ethylene epoxidation, EtOx)[77] is an important reaction in the petrochemical industry because of the numerous uses of ethylene epoxide as an intermediate chemical. Alumina-supported silver nanoparticles of a few tens nanometer in size are typically used in the industrial context,[78,79] but the reaction can be conducted efficiently also over subnanometer cluster catalysts, as in our previous studies on Ag_3 cluster catalysts supported on the MgO(100) regular surface.[39,40,80] As discussed in detail in section 2, here we focus exclusively on the rate-determining step of the EtOx mechanism on $Ag_3(CO_3)(C_2H_4)$, i.e., the step: "$Ag_3(CO_3)$-$(C_2H_4)_2(O) \rightarrow Ag_3(CO_3)(C_2H_4) + C_2H_4O$" (the barriers for the other steps being sufficiently low that the population of the corresponding intermediates can be considered as negligible) and we modify the catalyst of our original study[39] in two different ways, studying (i) $Ag_3(HCO_3)(C_2H_4)_2(O)$ and (ii) $Ag_3(CO_2F)(C_2H_4)_2(O)$ aggregates supported on the regular MgO(100) surface, and possibly how they might realize the conclusive catalytic step and evolve the C_2H_4O molecule as a photoassisted process. Moreover, since this "$Ag_3(XCO_x)$-$(C_2H_4)_2(O) \rightarrow Ag_3(XCO_x)(C_2H_4) + C_2H_4O$" step can actually be realized either as a single-step process or as a two-step process in which (1) the aggregate first rearranges to an oxametallocycle species to allow for oxygen-adatom migration from one C_2H_4 to the other, followed by (2) the detachment of the C_2H_4O epoxide molecule and its evolution into the gas phase, in the case of $Ag_3(CO_2F)(C_2H_4)_2(O)$ we calculated both the initial configuration and its isomeric oxametallocycle variant with the migrated O-adatom. In Figure 1 we then depict the initial $Ag_3(HCO_3)(C_2H_4)_2(O)$ isomer, Figure 1a, the initial $Ag_3(CO_2F)(C_2H_4)_2(O)$ isomer, Figure 1b, and the $Ag_3(CO_2F)(C_2H_4)_2(O)$ isomer intermediate between the initial configuration and the epoxide evolution, Figure 1c. The reaction energies for the $Ag_3(CO_2F)$-$(C_2H_4)_2(O)$ case are similar to the original $Ag_3(CO_3)$-$(C_2H_4)_2(O)$ case,[39,40] with the oxygen-migrating isomerization and the epoxide evolution costing 0.09 and 0.69 eV, respectively, using the CP2K code, whereas in the $Ag_3(HCO_3)(C_2H_4)_2(O)$ case the protonation of the carbonate increases the energy of the oxygen-migrating step to 0.16 eV: this is why we considered the oxygen-migrating isomer only for the more favorable fluorinated complex. Despite these limitations, the systems here investigated can be considered good tests of an overall strategy to the photochemistry of heterogeneous ultrananocatalysts.

The optical absorption spectra of the $Ag_3(HCO_3)$-$(C_2H_4)_2(O)$, Figure 1a, and $Ag_3(CO_2F)(C_2H_4)_2(O)$, Figure 1b, initial isomers are reported in Figure 2a,b, respectively.

Figure 2. Photoabsorption spectrum of (a) $Ag_3(HCO_3)(C_2H_4)_2(O)$ (b) $Ag_3(CO_2F)(C_2H_4)_2(O)$ simulated at the TDDFT level at the geometry of the respective electronic ground states. In part a, excitation nos. 8 and 21 are highlighted with red vertical bars and discussed in detail in the text.

In this work, we are interested to identify links between the characteristics of a given excited state and the photochemical reaction path its excitation gives rise to. Knowledge of such links would allow one to promote selectively a specific reaction by changing the energy of the light employed in the photoexcitation. We now therefore focus on the analysis tools available to identify these links.

Since ethylene oxidation is the reaction of interest, we started with the analysis of the C−O fragment, and in particular of its σ^*-antibonding molecular orbital, to verify if there were excitations populating such antibonding orbitals and thus promoting an undesired cleavage of the C−O bond. From an analysis of the DFT ground state orbitals, we found that such σ^* antibonding orbitals lie very high in energy (even beyond the ionization limit) and are therefore not accessible with UV−vis photoexcitation, so that we can safely ignore them.

An interesting and potentially reactive possibility instead corresponds to promote electrons into virtual orbitals with Ag−O antibonding character: in this way one could expect the structural relaxation following the electron excitation to go in the direction of a weaker interaction between the metal cluster and the oxygen adatom, therefore making the interaction of the latter with the ethylene molecules stronger and hence favoring the desired C_2H_4O-evolution reaction mechanism. With this idea in mind, as a first step, we identified the molecular orbitals (MO) with antibonding character between the oxygen atom

involved in the ethylene oxidation (labeled as O1 in Figure 1) and the two Ag atoms directly bonded to it, labeled Ag1 and Ag2 in Figure 1. This analysis can be easily performed using the overlap population analysis of the density of states (OP-DOS)[81] as reported in Figures S4 and S5 for the structures containing HCO_3 and CO_2F moieties, respectively. In the OP-DOS analysis plots, bonding and antibonding orbitals are characterized by positive and negative signs, respectively. For the HCO_3^- system, we find Ag1−O1 and Ag2−O1 antibonding interactions corresponding to the LUMO+2 and LUMO+3 orbitals, respectively. The model with the CO_2F fragment instead exhibits Ag1−O1 and Ag2−O1 antibonding interactions corresponding to the LUMO orbital for both Ag1−O1 and Ag2−O1. Schematic contour plots of the corresponding virtual MOs are reported in Figure 3, and as
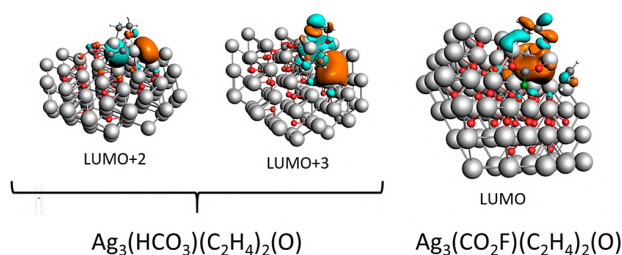
Figure 3. Molecular orbitals of the clusters $Ag_3(HCO_3)(C_2H_4)_2(O)$ and $Ag_3(CO_2F)(C_2H_4)_2(O)$ involved in the photoexcitation excitations involving Ag1−O1 and Ag2−O1 bonds.

expected from the OP-DOS analysis, these MOs are all localized mainly on the Ag_3 cluster fragment or very close to it and also display several nodal surfaces confirming their antibonding character. The following step consists in analyzing the photoabsorption spectra at the ground state geometry, reported in Figure 2 for both HCO_3 and CO_2F models, to find excitations corresponding to populating the Ag−O antibonding orbitals thus singled out. For HCO_3, we could identify two excitations (excited state nos. 8 and 21) which correspond to populating the LUMO+2 and LUMO+3 orbitals as suggested by the OP-DOS analysis. These excitations have been highlighted with red vertical bars in Figure 2. Interestingly, the situation is completely different for the CO_2F model: for this system, all the excitation peaks in the optical region below 3 eV have a strong component onto the LUMO, with the only exceptions of the excited state no. 6 at 2.72 eV and the excited state no. 9 at 2.83 eV. So while for HCO_3 only excitation nos. 8 and 21 seem promising, for CO_2F there are many excitations with a potential photochemical effect. For this reason, also considering that this field is still largely unexplored so that we do not have previous experience to guide simplifications, we found it interesting to do a systematic study of the structural relaxation of all excited states of the $Ag_3(CO_2F)(C_2H_4)_2(O)$ aggregate, and also for the $Ag_3(HCO_3)(C_2H_4)_2(O)$ aggregate we did not focus exclusively on the most promising excitations. In practice, starting from the optimized geometry of the ground state, we conducted a geometry optimization for each excited state in the optical spectrum of $Ag_3(CO_2F)$-$(C_2H_4)_2(O)$, while for $Ag_3(HCO_3)(C_2H_4)_2(O)$ we conducted a more limited search.

Figure 4 shows typical photoproducts coming out of our geometry relaxations. Starting the description of the results with the HCO_3 model, we found that the structural relaxation
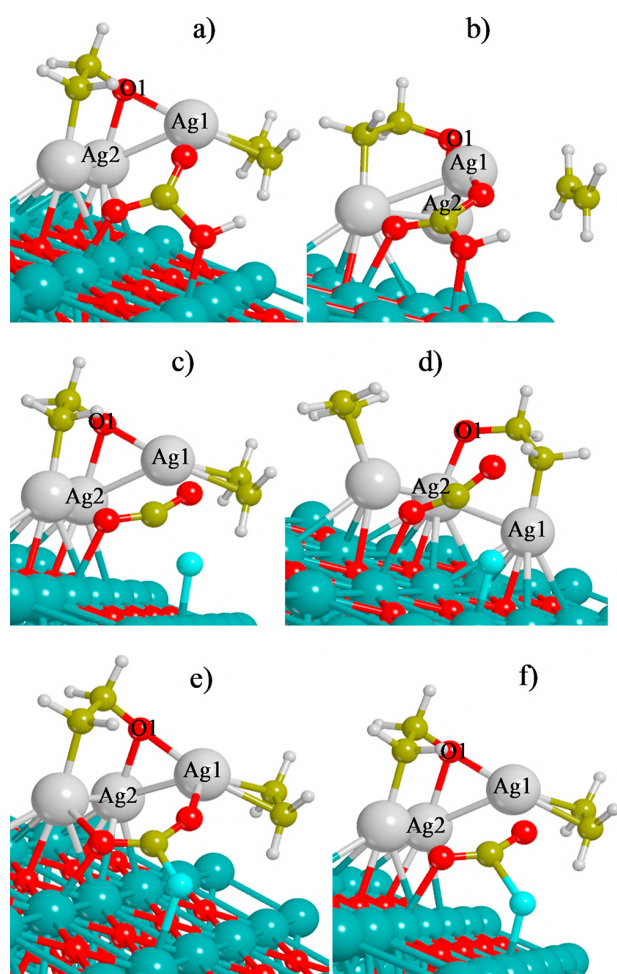
Figure 4. Schematic depiction of photoproducts obtained from the geometry optimizations of the excited states. $Ag_3HCO_3(Et)_2O//$ $MgO(100)$: (a) excited state no. 8 and (b) excited state no. 21. $Ag_3CO_2F(Et)_2O//MgO(100)$: (c) excited state no. 1 and (d) excited state no. 1 of the intermediate structure (Figure 1c). (e) Excited state no. 3 and (f) excited state no. 1 starting from geometry optimized for state no. 3. The initial structures (reactant) have been considered, except for the cases in Figure 4d,f in which intermediate species have been considered.

following the excitation no. 8 brings to the detachment of the HCO_3^- fragment from the Ag_3 cluster (see Figure 4a). Although this product is not interesting from a catalytic point of view, it demonstrates that the computational protocol is able to follow a reaction path after photoabsorption. Going to the higher excited state no. 21, we find that relaxation instead promotes the detachment of an ethylene molecule (Figure 4b). Also in this case the photoreaction product is not the desired one; however, again a chemical reaction is observed as a consequence of photoinduced relaxation effects. To explore in a more systematic way the geometry optimization of the excited states, we have investigated all the excited states up to the no. 12. In Table S1 we report, for each optimization, the optimized Ag1−O1 and Ag2−O1 distances as well as the final energy difference (relaxation + excitation) with respect to the optimized ground state, which corresponds to the energy fed into the system, i.e., the part of the photon energy which is stored as chemical energy and not dissipated in phononic

terms. Interestingly, we find that, while most excited-state relaxations fall back onto a geometry similar to the ground state, there is another excited state (the no. 6) which lead to HCO_3 decomposition. Also, it can be noted that for excited state nos. 6 and 21. we find an elongation of Ag1−O1 and Ag2−O1 distances, in agreement with the antibonding character of the populated orbitals. This confirms that a correlation can be found between the character of the excited state and the ensuing relaxation path. Naturally, the "specificity" of such a correlation will depend on the chemical system under investigation, and it will be important to search for systems that exhibit enhanced correlations and photo-catalytic yields. A more detailed structural description of the excitation process for $Ag_3(HCO_3)(C_2H_4)_2(O)$ is reported in Figure 5 for the optimized excited state nos. 8 and 21 (the MgO support is not shown for the sake of clarity). Their energies are reported in Table S3. Interestingly, the structural relaxation involves Ag−Ag distances as well as the Ag−$O(C_2H_4)$ distance, while Ag−C, C−C, and C−O distances show very small differences. In Figure 5, we also report the optimization path, i.e., the energy as a function of the optimization cycle. Occasional jumps are evident: this happens since the optimization procedure must follow the state maximizing the overlap between two successive optimization geometry. When the order of the state changes during the optimization, the algorithm switches to the new order following the requested state maximizing the overlap between states of two successive optimization geometry, and this reflects in the "jumps" observed occasionally during the optimization. Toward the end of the optimization some "ripples" are evident, due to the fact that the HCO_3^- fragment is basically detached from the silver cluster and is relaxing with very shallow minima over the MgO surface.

In passing, we recall from a previous analysis[24] that in view of controlling and exploiting the interaction between the excitations of the aggregate and those of the support, of possible importance in photocatalysis,[29−36] it is interesting to catalogue the excited states of the aggregate in terms of their transition dipole moment, distinguishing those with a moment parallel and perpendicular to the oxide surface. In Tables S3 and S4 and Figure S6, we therefore report for reference the parallel and perpendicular components of the excitation dipole moments. It can be noted that the perpendicular component is not dominant in this case, being always smaller or at least not greater than the parallel one, due to the elongated arrangement of these ligand/cluster complexes. However, significant differences are present: for example, the ratio between perpendicular to parallel component differs by a factor of 2 between excited state nos. 8 and 21, with the latter exhibiting a more pronounced perpendicular character.

Considering now the compound containing the CO_2F fragment, here, in contrast with the HCO_3 case, it is difficult to preselect a set of promising excitations, since the same virtual orbital (LUMO) is the leading component in most of the excitations in the optical region, as discussed above. Therefore, we have conducted a complete series of optimization of excited states in the optical region (and we report in Table S2 the corresponding distances). Moreover, here we also tried two further variants, as discussed below.

First, we note that five excited states among the first 12 lowest-energy ones (excited state nos. 1, 2, 4, 6, 9) lead to dissociation or decomposition of the CO_2F fragment. As a typical example, the result of the relaxation of the first excited
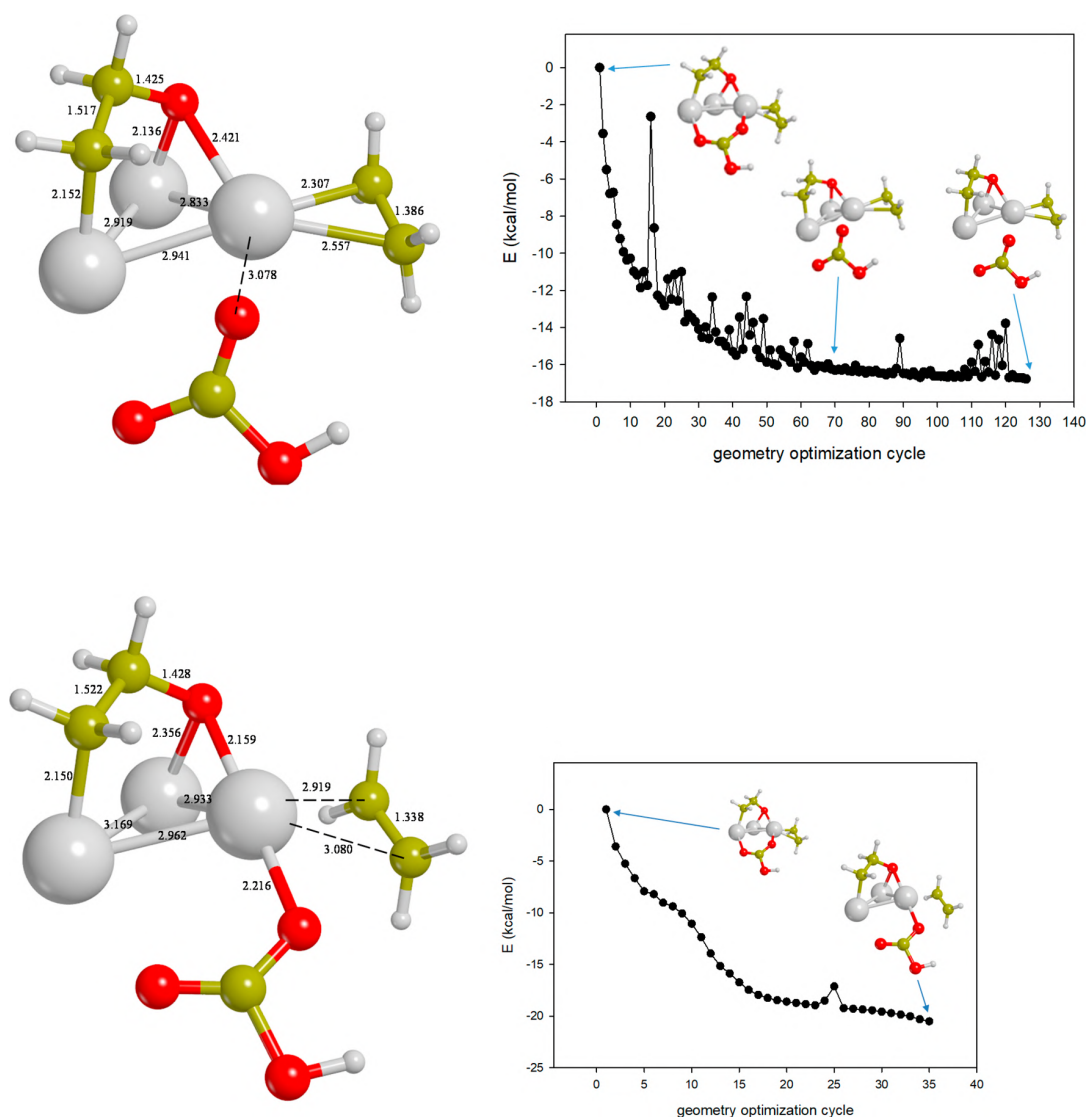
Figure 5. Schematic depiction of photoproducts obtained from the geometry optimizations of excited states of $Ag_3HCO_3(Et)_2O//MgO(100)$: excited state no. 8 (upper panels) and excited state no. 21 (lower panel). Structural details are shown on the left, and the optimization path is shown on the right.

state (no. 1) is reported in Figure 4c, where the decomposition of the CO_2F fragment into a CO_2 fragment (an eventually evolving CO_2 molecule) and a F^- anion supported onto MgO(100) can be clearly appreciated.

A more detailed structural description of the excitation process for $Ag_3(CO_2F)(C_2H_4)_2(O)$ is reported in Figure S7 for the optimized excited state nos. 1 and 2 (the MgO support is again not shown for the sake of clarity). Their energies are reported in Table S4. For these clusters, the structural differences are smaller than in the previous case with the HCO_3 group. The optimization paths are more regular toward the end of the optimization, but at the beginning the second state displays big jumps. For $Ag_3(CO_2F)(C_2H_4)_2(O)$, we also tried two further possibilities.

First, we optimized the lowest excited state of the intermediate structure instead of the lowest excited state of the initial structure of $Ag_3(CO_2F)(C_2H_4)_2(O)$, i.e., the isomer shown in Figure 1c. As shown in Figure 4d, also in this case a

CO_2F decomposition is observed as an outcome of this relaxation rather than epoxide evolution.

Then we considered a further possibility. The pictorial illustration in Figure 4e corresponds to the optimization of excited state no. 3, and in this case photoexcitation does not bring any significant reaction, only a slight structural rearrangement, with a shortening of the Ag1−O1 bond and a lengthening of the Ag2−O1 one. Nonetheless, we thought this could be a good case in which to test what can happen when the final geometry coming out from an excited state-relaxation is long-lived enough to sustain a successive photoexcitation. Note that to this purpose, we have chosen the final geometry of excited state no. 3 since it preserves the integrity of the CO_2F fragment, while relaxation of both excited state nos. 1 and 2 lead to CO_2F detachment or decomposition. Starting from the relaxed geometry of excited state no. 3 (Figure 4e), we have performed a TDDFT simulation of the optical spectrum at this geometry and then have optimized the lowest-energy excited state no. 1 of this

aggregate, thus producing a final geometry as illustrated in Figure 4f. As can be seen from Figure 4f and also from Table S2 (see row SS3_SS1 to be compared with row SS1 in the same table), the final result of this optimization is similar (in terms of both distances as well as energies) to the simple relaxation of the first excited state, so the net result of this test is that in this case the excitation of the lowest-energy state at the relaxed geometry of a higher excited-state (no. 3) "brings the system back" and is essentially the same as exciting directly the lowest-energy excitation from the ground-state geometry.

To conclude this discussion, we trace back the present findings to experimental observations. As discussed in section 2, carbonate and bicarbonate functional groups have been shown to play an important role in general in the catalytic oxidation of organic species over Ag, Au, Pd, and Pt catalysts, supporting our choice of investigated systems.

On Au, Pd, and Pt systems as COox catalysts, carbonates have been shown to have a poisoning effect, rather than being reactive intermediates.[44−55] However, according to our previous theoretical modeling, this does not occur on $Ag_3(CO_3)/MgO(100)$ complexes, which are catalytically active with reasonably low energy barriers in COox and also (after addition of an ethylene molecule) in EtOx. In this connection, we note that, with respect to the literature systems, our investigated systems are made of subnanometer (different size) and silver (different catalyst) clusters,[7,39−42] and we can also recall as a crucial observation that, whereas the catalytic activity of Pt_{10}/a-Al_2O_3 clusters in COox is sensitive to the generation of carbonate ligands,[55] similarly sized $Ag_9Pt_{2−3}/a$-Al_2O_3 clusters do not suffer from this phenomenon,[82] suggesting that Ag systems, in particular subnanometer ones, may be less prone to carbonate poisoning.

Under hydrous COox conditions, carbonate invariably converts to some degree into bicarbonates, which are assumed to be labile and decompose, thus hindering the carbonate poisoning effect. In contrast with this picture, in the present work we found that bicarbonates per se can actually be stable, and their decomposition may not be so easy, unless promoted by optical excitation. The discrepancy between previous work and the present study can be therefore reconciled noting that the literature experimental studies were all conducted under normal light irradiation: on the basis of the present findings, we expect that bicarbonates form from carbonates under hydrous conditions and are stable per se, but their decomposition can be caused by optical absorption and it could therefore be avoided working under dark conditions. In this connection, it can also be recalled as an important observation that bicarbonates have been proposed to play an important role also in the reverse process to COox, the reduction reaction of carbon dioxide (CO2RR), as discussed in section 2.[56−60] This role of bicarbonate intermediates is a strong point in favor of their possessing some stability: if they were at very high energy, their presence would be too labile to play any catalytically relevant role.

In conclusion, the present work points to the possible "optical frailty" of complexes involving bicarbonate species. Optical degradation has been recently demonstrated for other ultrananomaterials,[83] and an analogy with the present work might be conducted. This "optical frailty" might also be a weak feature of subnanometer catalytic complexes and appears as an important topic for future research.

4. CONCLUSIONS

In the present work, we have developed and tested an approach for investigating the photochemistry of subnanometer (or ultranano) oxide-supported metal clusters. The proposed protocol consists in designing a realistic model of the catalytic system, comprising the metal cluster and a shell of adsorbates (either reactants or intermediates or products) and in which the oxide support is described by a small number of explicit, QM-treated ions augmented by a large array of point charges to model the Madelung potential of the environment. While the support is kept frozen, all the remaining atoms are free to relax. Photochemistry is then studied by TDDFT relaxation of each relevant excited state. The excited states most promising for photochemistry are selected by first inspecting overlap population of density of states (OP-DOS) plots[81] to single out which molecular orbitals (MOs) have a proper bonding or antibonding character which can be expected to produce the desired photochemical process. The orbitals identified via OP-DOS are then searched as components of the excited states as derived from TDDFT simulations at the ground-state equilibrium geometry. Finally, the excited states with the proper MOs as main components are subjected to full geometry relaxation. The computational effort entailed by the present approach is affordable according to current standards, with a typical excited-state optimization requiring about 36 h of elapsed time using 12 cores of a ProLiant DL580 Gen10 HP server equipped with Intel Xeon Gold 6140 CPU @ 2.30 GHz. The present protocol is thus able to treat accurately realistic systems with a feasible computational effort.

In the present study, we focused on two examples which, based on previous work,[39,40] correspond to plausible intermediates derived from initial structures of the rate-determining step in the ethylene partial oxidation by Ag_3 species supported on the MgO(100) surface: the $Ag_3(HCO_3)$-$(C_2H_4)_2(O)$ and $Ag_3(CO_2F)(C_2H_4)_2(O)$ aggregates and are intended as realistic tests of an overall strategy to the photochemistry of heterogeneous ultrananocatalysts. The OP-DOS between the oxygen atom which participates in the oxidation reaction and the Ag atoms were selected as descriptors to single out which MOs and hence which excited states are more promising in terms of possible photoinduced reactions. The results demonstrated the capability of the present approach to simulate reactive paths following photoexcitation, thus providing the tools for a line of research potentially of great interest, although in the present case we found that such paths mainly resulted in the detachment of HCO_3^-, CO_2F^-, or ethylene groups or the decomposition of the CO_2F^- ligand into an adsorbed F^- anion and a CO_2 molecule, pointing to a possible "optical frailty" of ultrananocatalysts upon photoexcitation. Nevertheless, a correlation was found between reaction paths and the nature of the excited states, thus suggesting that these correlations exist and can be equally or even more pronounced for other systems, so as to be useful in the design of active photocatalytic systems.

In terms of perspectives, further applications to systems with "real" photocatalytic activity will be needed to tune the present protocol. In particular, we have preliminarily tested with success the application of the method to open-shell systems employing an unrestricted spin-polarized scheme, an essential step to use the method in many catalytically relevant cases. Also, the use of analysis tools more general than OP-DOS such

as individual component map of oscillatory strength (ICM-OS) plots[84] and perhaps the extension to chiral systems using individual component map of oscillatory strength (ICM-RS) plots[85] could also be useful in realistic cases. Naturally, the accuracy of the approach is a potential issue and, to increase it, the use of hybrid DFT exchange-correlation functionals should be considered. This however seems feasible especially if helped by implementing simplified but computationally efficient schemes.[86] A final more delicate issue corresponds to a well-known still open problem in photochemistry, i.e., the fact that the procedure to follow a given excited state may fail in the presence of a very congested potential energy surface, for example, in the presence of conical intersections or avoided crossings.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.0c08815.

> Formation of the ligand/cluster/support catalytic complex; depiction of the point charge and explicit atom models; overlap population of density of states (OP-DOS) plots and tables with parallel and perpendicular components of the excitation dipole moments of the various excited states; geometry optimizations of selected excited states of $Ag_3(CO_2F)(C_2H_4)_2(O)$; tables with selected Ag−O distances in the ground-state configurations and in the configurations resulting from excited-state structural optimization; and Cartesian coordinated of the excited states here considered (PDF)

AUTHOR INFORMATION

Corresponding Authors

Mauro Stener − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste I-34127, Italy;* ⓞ orcid.org/0000-0003-3700-7903; Email: stener@units.it

Alessandro Fortunelli − *CNR-ICCOM & IPCF, Consiglio Nazionale delle Ricerche, ThC2-Lab and Molecular Modelling Team, 56124 Pisa, Italy;* ⓞ orcid.org/0000-0001-5337-4450; Email: alessandro.fortunelli@cnr.it

Authors

Luca Sementa − *CNR-ICCOM & IPCF, Consiglio Nazionale delle Ricerche, ThC2-Lab and Molecular Modelling Team, 56124 Pisa, Italy*

Marta Monti − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste I-34127, Italy*

Daniele Toffoli − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste I-34127, Italy;* ⓞ orcid.org/0000-0002-8225-6119

Alvaro Posada-Amarillas − *Departamento de Investigación en Física, Universidad de Sonora, 83000 Hermosillo, Sonora, Mexico;* ⓞ orcid.org/0000-0001-7533-4820

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcc.0c08815

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Heiz, U.; Sanchez, A.; Abbet, S.; Schneider, W.-D. CO Oxidation on Monodispersed Pt Clusters. *J. Am. Chem. Soc.* **1999**, *121*, 3214−3217.

(2) Herzing, A. A.; Kiely, C. J.; Carley, A. F.; Landon, P.; Hutchings, G. J. Identification of Active Gold Nanoclusters on Iron Oxide Supports for CO Oxidation. *Science* **2008**, *321*, 1331−1335.

(3) Tyo, E. C.; Vajda, S. Catalysis by clusters with precise numbers of atoms. *Nat. Nanotechnol.* **2015**, *10*, 577−588.

(4) Luo, Z.; Castleman, A. W.; Khanna, S. N. Reactivity of Metal Clusters. *Chem. Rev.* **2016**, *116*, 14456−14492.

(5) Thomas, J. M. Heterogeneous catalysis: Enigmas, illusions, challenges, realities, and emergent strategies of design. *J. Chem. Phys.* **2008**, *128*, 182502.

(6) Bihouix, P.; de Guillebon, B. *Quel Futur pour les Métaux?* EDP Sciences: Paris, France, 2010.

(7) Negreiros, F. R.; Aprà, E.; Barcaro, G.; Sementa, L.; Vajda, S.; Fortunelli, A. A first-principles theoretical approach to heterogeneous nanocatalysis. *Nanoscale* **2012**, *4*, 1208−1219.

(8) Zhang, Z.; Zandkarimi, B.; Alexandrova, A. N. Ensembles of Metastable States Govern Heterogeneous Catalysis on Dynamic Interfaces. *Acc. Chem. Res.* **2020**, *53*, 447−458.

(9) Bonacic-Koutecky, V.; Veyret, V.; Mitric, R. Ab initio study of the absorption spectra of Ag-n (n = 5−8) clusters. *J. Chem. Phys.* **2001**, *115*, 10450−10460.

(10) Harb, M.; Rabilloud, F.; Simon, D.; Rydlo, A.; Lecoultre, S.; Conus, F.; Rodrigues, V.; Felix, C. Optical absorption of small silver clusters: Ag(n), (n = 4−22). *J. Chem. Phys.* **2008**, *129*, 194108.

(11) Barcaro, G.; Sementa, L.; Fortunelli, A.; Stener, M. Optical properties of nanoalloys. *Phys. Chem. Chem. Phys.* **2011**, *17*, 27952−27967.

(12) Fortunelli, A.; Stener, M. Optical Properties of Metal Nanoclusters—Theory. In *Encyclopedia of Interfacial Chemistry: Surface Science and Electrochemistry*; Elsevier, 2018, DOI: 10.1016/B978-0-12-409547-2.12960-9.

(13) Dass, A. Faradaurate Nanomolecules: A Superstable Plasmonic 76.3 kDa Cluster. *J. Am. Chem. Soc.* **2011**, *133*, 19259−19261.

(14) Negishi, Y.; Nakazaki, T.; Malola, S.; Takano, S.; Niihori, Y.; Kurashige, W.; Yamazoe, S.; Tsukuda, T.; Häkkinen, H. A Critical Size for Emergence of Nonbulk Electronic and Geometric Structures in Dodecanethiolate-Protected Au Clusters. *J. Am. Chem. Soc.* **2015**, *137*, 1206−1212.

(15) Jin, R. C.; Zeng, C. J.; Zhou, M.; Chen, Y. X. Atomically Precise Colloidal Metal Nanoclusters and Nanoparticles: Fundamentals and Opportunities. *Chem. Rev.* **2016**, *116*, 10346−10413.

(16) Dass, A.; Sakthivel, N. A.; Jupally, V. R.; Kumara, C.; Rambukwella, M. Plasmonic Nanomolecules: Electrochemical Resolution of 22 Electronic States in $Au_{329}(SR)_{84}$. *ACS Energy Lett.* **2020**, *5*, 207−214.

(17) Link, S.; El-Sayed, M. A. Optical Properties and Ultrafast Dynamics of Metallic Nanocrystals. *Annu. Rev. Phys. Chem.* **2003**, *54*, 331−366.

(18) Hartland, G. V. Optical Studies of Dynamics in Noble Metal Nanostructures. *Chem. Rev.* **2011**, *111*, 3858−3887.

(19) Yau, S. H.; Varnavski, O.; Goodson, T. An Ultrafast Look at Au Nanoclusters. *Acc. Chem. Res.* **2013**, *46*, 1506−1516.

(20) Jin, R. C. Atomically precise metal nanoclusters: stable sizes and optical properties. *Nanoscale* **2015**, *7*, 1549−1565.

(21) Shabaninezhad, M.; Abuhagr, A.; Sakthivel, N. A.; Kumara, C.; Dass, A.; Kwak, K.; Pyo, K.; Lee, D.; Ramakrishna, G. Ultrafast

Electron Dynamics in Thiolate-Protected Plasmonic Gold Clusters: Size and Ligand Effect. *J. Phys. Chem. C* **2019**, *123*, 13344−13353.

(22) Weinstein, J. A. The future iron age. *Nat. Chem.* **2020**, *12*, 789−790.

(23) Lünskens, T.; Heister, P.; Thämer, M.; Walenta, C. A.; Kartouzian, A.; Heiz, U. Plasmons in supported size-selected silver nanoclusters. *Phys. Chem. Chem. Phys.* **2015**, *17*, 17541−17544.

(24) Luque-Ceballos, J. C.; Sementa, L.; Aprà, E.; Fortunelli, A.; Posada-Amarillas, A. TDDFT Study of the Optical Spectra of Free and Supported Binary Coinage Metal Hexamers: Effect of Doping and Support. *J. Phys. Chem. C* **2018**, *122*, 23143−23152.

(25) Lopez-Caballero, P.; Ramallo-Lopez, J. M.; Giovanetti, L. J.; Buceta, D.; Miret-Artes, S.; Lopez-Quintela, M. A.; Requejo, F. G.; de Lara-Castells, M. P. Exploring the properties of Ag$_5$−TiO$_2$ interfaces: stable surface polaron formation, UV-Vis optical response, and CO$_2$ photoactivation. *J. Mater. Chem. A* **2020**, *8*, 6842−6853.

(26) Balzani, V.; Ceroni, P.; Juris, A. *Photochemistry and Photophysics: Concepts, Research, Applications*; Wiley-VCH: Weinheim, Germany, 2014.

(27) Gratzel, M. Dye-sensitized solar cells. *J. Photochem. Photobiol., C* **2003**, *4*, 145−153.

(28) Linic, S.; Christopher, P.; Ingram, D. B. Plasmonic-metal nanostructures for efficient conversion of solar to chemical energy. *Nat. Mater.* **2011**, *10*, 911−921.

(29) Diebold, U. The surface science of titanium dioxide. *Surf. Sci. Rep.* **2003**, *48*, 53−229.

(30) Watanabe, K.; Menzel, D.; Nilius, N.; Freund, H.-J. Photochemistry on Metal Nanoparticles. *Chem. Rev.* **2006**, *106*, 4301−4320.

(31) Shimura, K.; Yoshida, H. Heterogeneous Photocatalytic Hydrogen Production from Water and Biomass Derivatives. *Energy Environ. Sci.* **2011**, *4*, 2467−2481.

(32) Murdoch, M.; Waterhouse, G. I. N.; Nadeem, M. A.; Metson, J. B.; Keane, M. A.; Howe, R. F.; Llorca, J.; Idriss, H. The Effect of Gold Loading and Particle Size on Photocatalytic Hydrogen Production from Ethanol over Au/TiO2 Nanoparticles. *Nat. Chem.* **2011**, *3*, 489−492.

(33) Montini, T.; Gombac, V.; Sordelli, L.; Delgado, J. J.; Chen, X. W.; Adami, G.; Fornasiero, P. Nanostructured Cu/TiO2 Photocatalysts for H2 Production from Ethanol and Glycerol Aqueous Solutions. *ChemCatChem* **2011**, *3*, 574−577.

(34) Letfullin, R. R.; George, T. F. Plasmonic nanomaterials for nanomedicine. In *Springer Handbook: Nanomaterials*; Vajtai, R., Ed.; Springer: Berlin, Germany, 2013.

(35) Su, R.; Tiruvalam, R.; Logsdail, A. J.; He, Q.; Downing, C. A.; Jensen, M. T.; Dimitratos, N.; Kesavan, L.; Wells, P. P.; Bechstein, R.; et al. Designer Titania-Supported Au-Pd Nanoparticles for Efficient Photocatalytic Hydrogen Production. *ACS Nano* **2014**, *8*, 3490−3497.

(36) Low, J.; Yu, J.; Ho, W. Graphene-Based Photocatalysts for CO2 Reduction to Solar Fuel. *J. Phys. Chem. Lett.* **2015**, *6*, 4244−4251.

(37) Hoffmann, N. *Chem. Rev.* **2008**, *108*, 1052−1103.

(38) Masson, G.; König, B. Chemical Photocatalysis − Do It Right! *Eur. J. Org. Chem.* **2020**, *2020*, 1191−1192.

(39) Sementa, L.; Barcaro, G.; Negreiros, F. R.; Fortunelli, A. Ligand/cluster/support catalytic complexes in heterogeneous ultra-nanocatalysis: NO oxidation on Ag3/MgO(100). *Phys. Chem. Chem. Phys.* **2014**, *16*, 26570−26577.

(40) Sementa, L.; Barcaro, G.; Fortunelli, A. Analogy between homogeneous and heterogeneous catalysis by subnanometer metal clusters: Ethylene oxidation on Ag trimers supported on MgO(100). *Inorg. Chim. Acta* **2015**, *431*, 150−155.

(41) Negreiros, F. R.; Sementa, L.; Barcaro, G.; Vajda, S.; Aprà, E.; Fortunelli, A. CO Oxidation by Subnanometer AgxAu3−x Supported Clusters via Density Functional Theory Simulations. *ACS Catal.* **2012**, *2*, 1860−1864.

(42) Bulushev, D. A.; Khasin, A. V. Effect of carbon dioxide on the catalytic oxidation of ethylene over silver. CO2 adsorption. *React. Kinet. Catal. Lett.* **1991**, *44*, 421.

(43) Sambalova, O.; Thorwarth, K.; Heeb, N. V.; Bleiner, D.; Zhang, Y.; Borgschulte, A.; Kroll, A. Carboxylate Functional Groups Mediate Interaction with Silver Nanoparticles in Biofilm Matrix. *ACS Omega* **2018**, *3*, 724−733.

(44) Schryer, D. R.; Upchurch, B. T.; Sidney, B. D.; Brown, K. G.; Hoflund, G. B.; Herz, R. K. A proposed mechanism for Pt/SnOx-catalyzed CO oxidation. *J. Catal.* **1991**, *130*, 314−317.

(45) Hoflund, G. B.; Gardner, S. D.; Schryer, D. R.; Upchurch, B. T.; Kielin, E. J. Effect of CO2 on the performance of Au/MnOx AND Pt/SnOx low-temperature CO oxidation catalysts. *Langmuir* **1995**, *11*, 3431.

(46) Kung, H. H.; Kung, M. C.; Costello, C. K. Supported Au catalysts for low temperature CO oxidation. *J. Catal.* **2003**, *216*, 425−432.

(47) Date, M.; Okumura, M.; Tsubota, S.; Haruta, M. Vital Role of Moisture in the Catalytic Activity of Supported Gold Nanoparticles. *Angew. Chem., Int. Ed.* **2004**, *43*, 2129−2132.

(48) Schumacher, B.; Denkwitz, Y.; Plzak, V.; Kinne, M.; Behm, R. J. Kinetics, Mechanism, and the Influence of H2 on the Co Oxidation Reaction on a Au/Tio2 Catalyst. *J. Catal.* **2004**, *224*, 449−462.

(49) Azar, M.; Caps, V.; Morfin, F.; Rousset, J.; Piednoir, A.; Bertolini, J.; Piccolo, L. Insights into Activation, Deactivation and Hydrogen-Induced Promotion of a Au/TiO2 Reference Catalyst in CO Oxidation. *J. Catal.* **2006**, *239*, 307−312.

(50) Ojeda, M.; Zhan, B. Z.; Iglesia, E. Mechanistic Interpretation of CO Oxidation Turnover Rates on Supported Au Clusters. *J. Catal.* **2012**, *285*, 92−102.

(51) Chen, S.; Luo, L.; Jiang, Z.; Huang, W. Size-dependent reaction pathways of low-temperature CO oxidation on Au/CeO2 catalysts. *ACS Catal.* **2015**, *5*, 1653−1662.

(52) Föttinger, K.; Schlögl, R.; Rupprechter, G. The Mechanism of Carbonate Formation on Pd-Al2O3 Catalysts. *Chem. Commun.* **2008**, *3*, 320−322.

(53) Parker, S. F. The role of hydroxyl groups in low temperature carbon monoxide oxidation. *Chem. Commun.* **2011**, *47*, 1988−1990.

(54) Oh, S. H.; Hoflund, G. B. Low-Temperature Catalytic Carbon Monoxide Oxidation over Hydrous and Anhydrous Palladium Oxide Powders. *J. Catal.* **2007**, *245*, 35−44.

(55) Yin, C.; Negreiros, F. R.; Barcaro, G.; Beniya, A.; Sementa, L.; Tyo, E. C.; Bartling, S.; Meiwes-Broer, K.-H.; Seifert, S.; Hirata, H.; et al. Alumina-supported sub-nanometer Pt10 clusters: amorphization and role of the support material in a highly active CO oxidation catalyst. *J. Mater. Chem. A* **2017**, *5*, 4923−4931.

(56) Sreekanth, N.; Phani, K. L. Selective Reduction of CO2 to Formate through Bicarbonate Reduction on Metal Electrodes: New Insights Gained from SG/TC Mode of SECM. *Chem. Commun.* **2014**, *50*, 11143−1114.

(57) Chen, Y.; Li, C. W.; Kanan, M. W. Aqueous CO2 reduction at very low overpotential on oxide-derived Au nanoparticles. *J. Am. Chem. Soc.* **2012**, *134*, 19969−19972.

(58) Baruch, M. F.; Pander, J. E.; White, J. L.; Bocarsly, A. B. Mechanistic Insights into the Reduction of CO2 on Tin Electrodes using in Situ ATR-IR Spectroscopy. *ACS Catal.* **2015**, *5*, 3148−3156.

(59) Bai, X.; Chen, W.; Zhao, C.; Li, S.; Song, Y.; Ge, R.; Wei, W.; Sun, Y. Exclusive Formation of Formic Acid from CO$_2$ Electro-reduction by a Tunable Pd-Sn Alloy. *Angew. Chem., Int. Ed.* **2017**, *56*, 12219−12223.

(60) Alfath, M.; Lee, C. W. Recent Advances in the Catalyst Design and Mass Transport Control for the Electrochemical Reduction of Carbon Dioxide to Formate. *Catalysts* **2020**, *10*, 859.

(61) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15−25.

(62) Kuhne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schutt, O.; Schiffmann, F.; et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.

(63) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.

(64) Lippert, G.; Hutter, J.; Parrinello, M. A hybrid Gaussian and plane wave density functional scheme. *Mol. Phys.* **1997**, *92*, 477−487.

(65) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(66) VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127*, 114105.

(67) Kratzer, P.; Neugebauer, J. The Basics of Electronic Structure Theory for Periodic Systems. *Front. Chem.* **2019**, *7*, 106.

(68) Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 1703−1710.

(69) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé−Hoover Chains: The Canonical Ensemble Via Continuous Dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.

(70) Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem. Phys.* **1973**, *2*, 41−51. Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor. Chem. Acc.* **1998**, *99*, 391−403.

(71) Barcaro, G.; Apra, E.; Fortunelli, A. Structure of Ag Clusters Grown on Fs-Defect Sites of an MgO(100) Surface. *Chem. - Eur. J.* **2007**, *13*, 6408−6418.

(72) Saunders, V. R.; Freyria-Fava, C.; Dovesi, R.; Salasco, L.; Roetti, C. On the electrostatic potential in crystalline systems where the charge density is expanded in Gaussian functions. *Mol. Phys.* **1992**, *77*, 629−665.

(73) Casida, M. E. In *Recent Advances in Density-Functional Methods*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; p 155.

(74) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two - Component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597−4610.

(75) Gross, E. K. U.; Kohn, W. Time-Dependent Density-Functional Theory. *Adv. Quantum Chem.* **1990**, *21*, 255−291.

(76) Seth, M.; Mazur, G.; Ziegler, T. Time-dependent density functional theory gradients in the Amsterdam density functional package: geometry optimizations of spin-flip excitations. *Theor. Chem. Acc.* **2011**, *129*, 331−342.

(77) Weissermel, K.; Arpe, H. *Industrial Organic Chemistry*; VCH: New York, 1993.

(78) Hassani, S. S.; Ghasemi, M. R.; Rashidzadeh, M.; Sobat, Z. Nano-sized silver crystals and their dispersion over α-alumina for ethylene epoxidation. *Cryst. Res. Technol.* **2009**, *44*, 948−952.

(79) Kim, Y.-C.; Park, N.-C.; Shin, J.-S.; Lee, S. R.; Lee, Y. J.; Moon, D. J. Partial oxidation of ethylene to ethylene oxide over nanosized Ag/α-Al2O3 catalysts. *Catal. Today* **2003**, *87*, 153−162.

(80) Lei, Y.; et al. Increased Silver Activity for Direct Propylene Epoxidation via Subnanometer Size Effects. *Science* **2010**, *328*, 224−228.

(81) Mulliken, R. S. Electronic Population Analysis on LCAO−MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(82) Negreiros, F. R.; Halder, A.; Yin, C.; Singh, A.; Barcaro, G.; Sementa, L.; Tyo, E. C.; Pellin, M. J.; Bartling, S.; Meiwes-Broer, K.-H.; et al. Bimetallic Ag-Pt Subnanometer Supported Clusters as Highly Efficient and Robust Oxidation Catalysts. *Angew. Chem., Int. Ed.* **2018**, *57*, 1209−1213.

(83) Kotsakidis, J. C.; Zhang, Q.; Vazquez de Parga, A. L.; Currie, M.; Helmerson, K.; Gaskill, D. K.; Fuhrer, M. S. Oxidation of Monolayer WS2 in Ambient Is a Photoinduced Process. *Nano Lett.* **2019**, *19* (8), 5205−5215.

(84) Theivendran, S.; Chang, L.; Mukherjee, A.; Sementa, L.; Stener, M.; Fortunelli, A.; Dass, A. Principles of optical spectroscopy of aromatic alloy nanomolecules: Au36−xAgx(SPhtBu)24. *J. Phys. Chem. C* **2018**, *122*, 4524−4531.

(85) Chang, L.; Baseggio, O.; Sementa, L.; Cheng, D.; Fronzoni, G.; Toffoli, D.; Aprà, E.; Stener, M.; Fortunelli, A. Individual Component Map of Rotatory Strength and Rotatory Strength Density Plots As Analysis Tools of Circular Dichroism Spectra of Complex Systems. *J. Chem. Theory Comput.* **2018**, *14*, 3703−3714.

(86) Medves, M.; Sementa, L.; Toffoli, D.; Fronzoni, G.; Fortunelli, A.; Stener, M. An efficient hybrid scheme for time dependent density functional theory. *J. Chem. Phys.* **2020**, *152*, 184104.

Supporting Information for:

"Theoretical Investigation of Photo-Induced Processes in Subnanometer Oxide-Supported Metal Catalysts"

Luca Sementa[1], Marta Monti[2], Daniele Toffoli[2], Alvaro Posada-Amarillas[3], Mauro Stener[2,*], Alessandro Fortunelli[1,*]

[1]*CNR-ICCOM & IPCF, Consiglio Nazionale delle Ricerche, ThC2-Lab and Molecular Modelling Team, 56124 Pisa, Italy; Alessandro Fortunelli – orcid.org/0000-0001-5337-4450*

[2]*Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste I-34127, Italy; Mauro Stener – orcid.org/0000-0003-3700-7903*

[3]*Departamento de Investigación en Física, Universidad de Sonora, 83000 Hermosillo, Sonora, Mexico; Alvaro Posada-Amarillas – orcid.org/0000-0001-7533-4820*
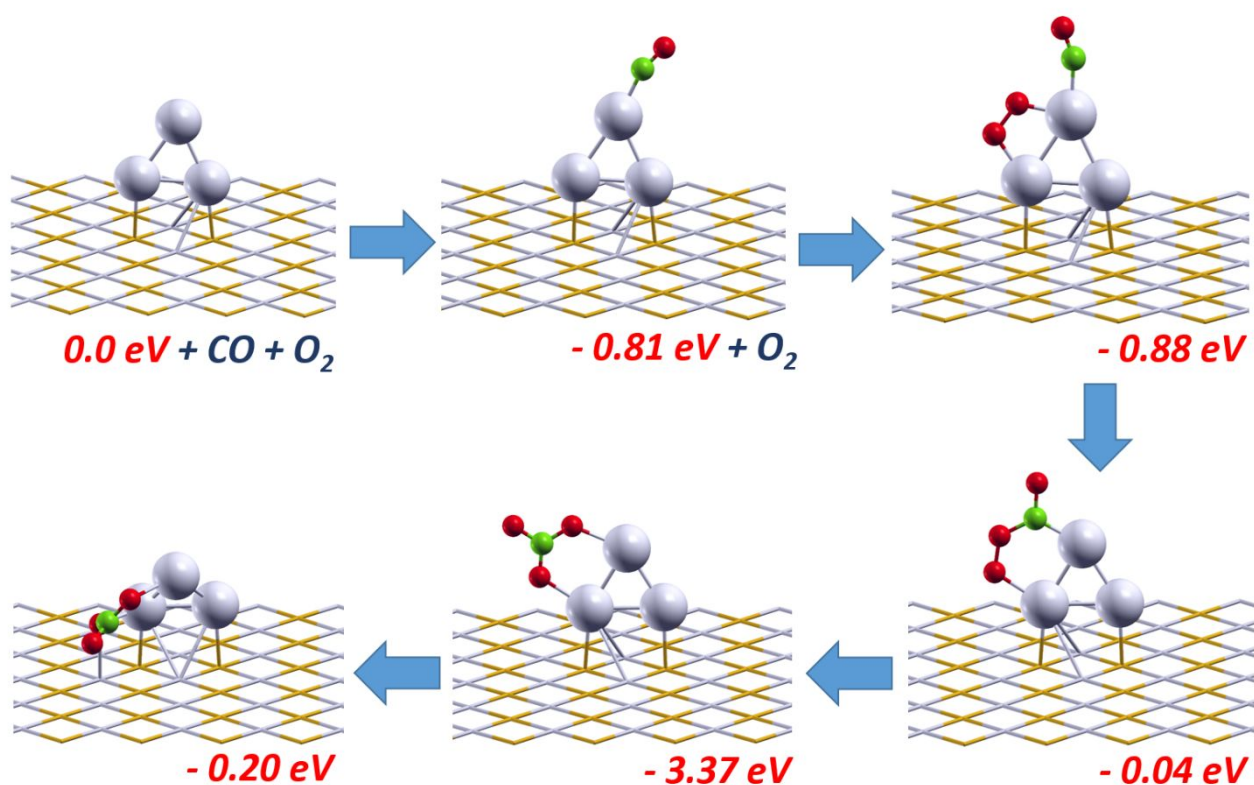
Figure S1 Pictorial illustration of the steps leading to the formation of the $Ag_3(CO_3)/MgO(100)$ ligand/cluster/support catalytic complex under COox conditions starting from the bare $Ag_3(CO_3)/MgO(100)$ complex, as discussed in Ref. [Negreiros, F. R.; Aprà, E.; Barcaro, G.; Sementa, L.; Vajda, S.; Fortunelli, A. A first principles theoretical approach to heterogeneous nanocatalysis Nanoscale, 2012, 4, 1208-1219, DOI: 10.1039/C1NR11051A].
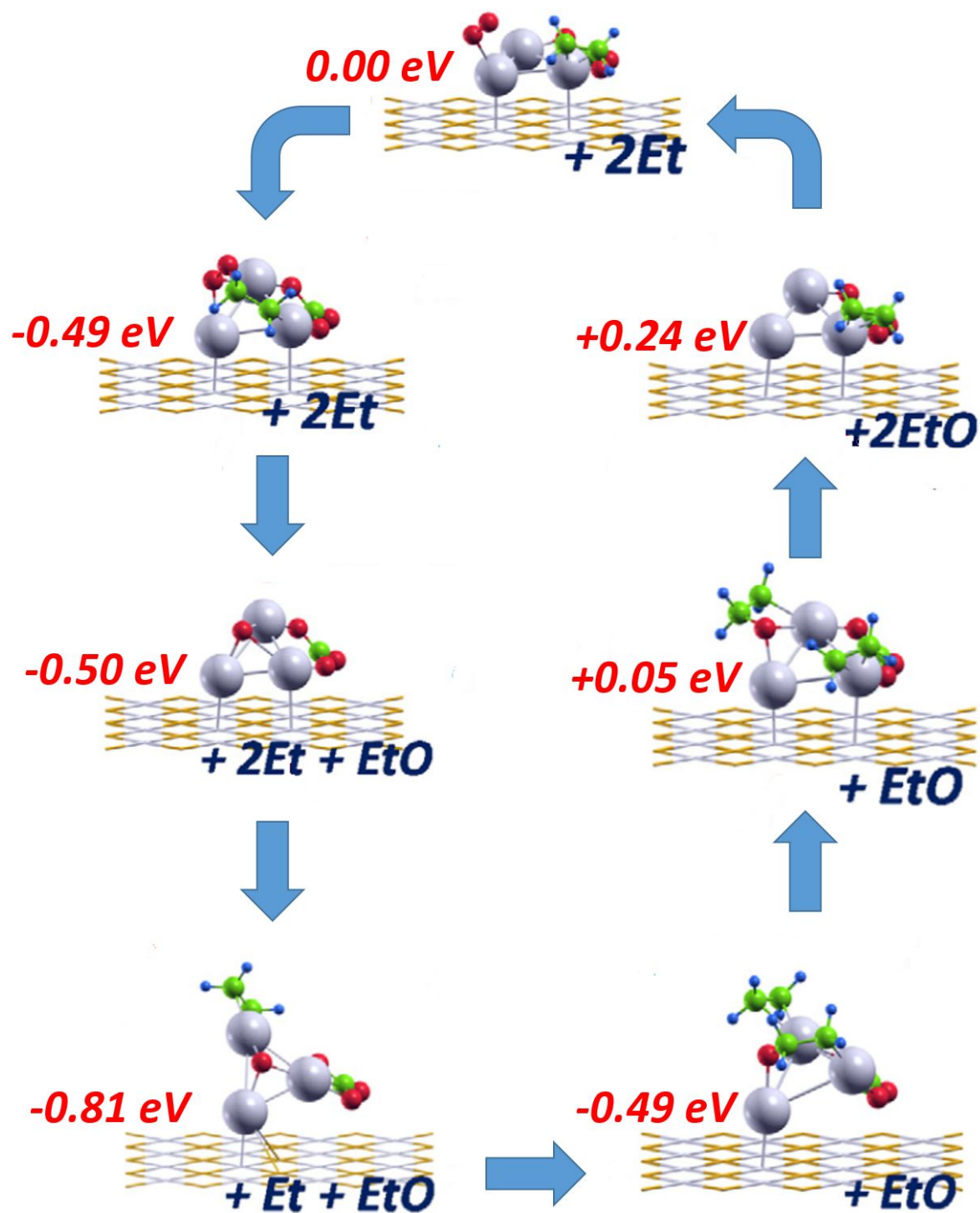
Figure S2 Pictorial illustration of the EtOx catalytic cycle of the $Ag_3(CO_3)(C_2H_4)$/MgO(100) ligand/cluster/support catalytic complex, as discussed in Ref. [Sementa, L.; Barcaro, G.; Negreiros, F. R.; Fortunelli, A. Ligand/cluster/support catalytic complexes in heterogeneous ultrananocatalysis: NO oxidation on Ag3/MgO(100). Phys. Chem. Chem. Phys. 2014, 16, 26570-26577, DOI: 10.1039/c4cp02135e].
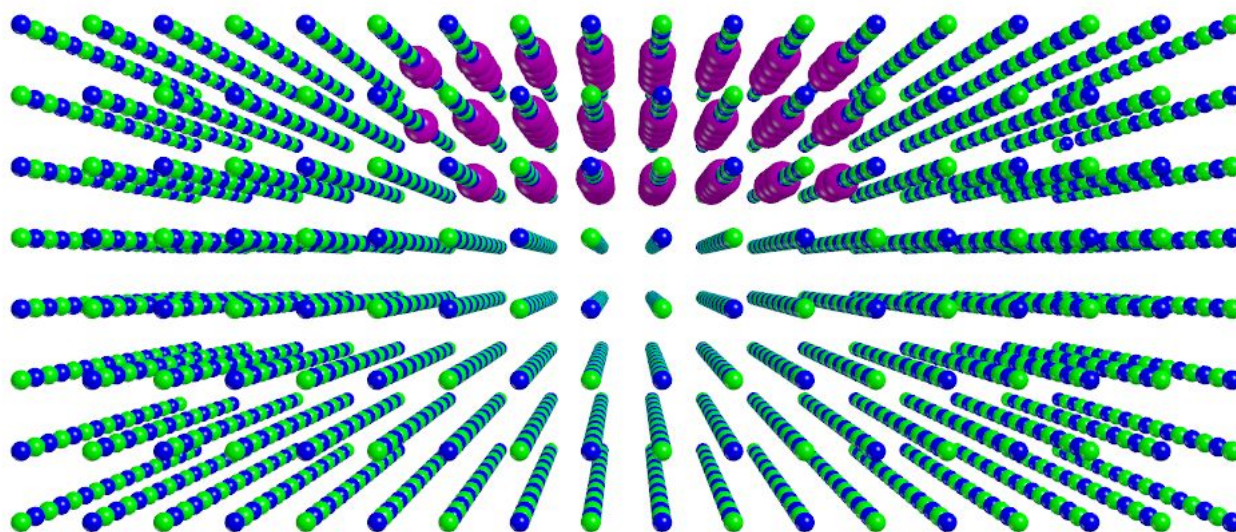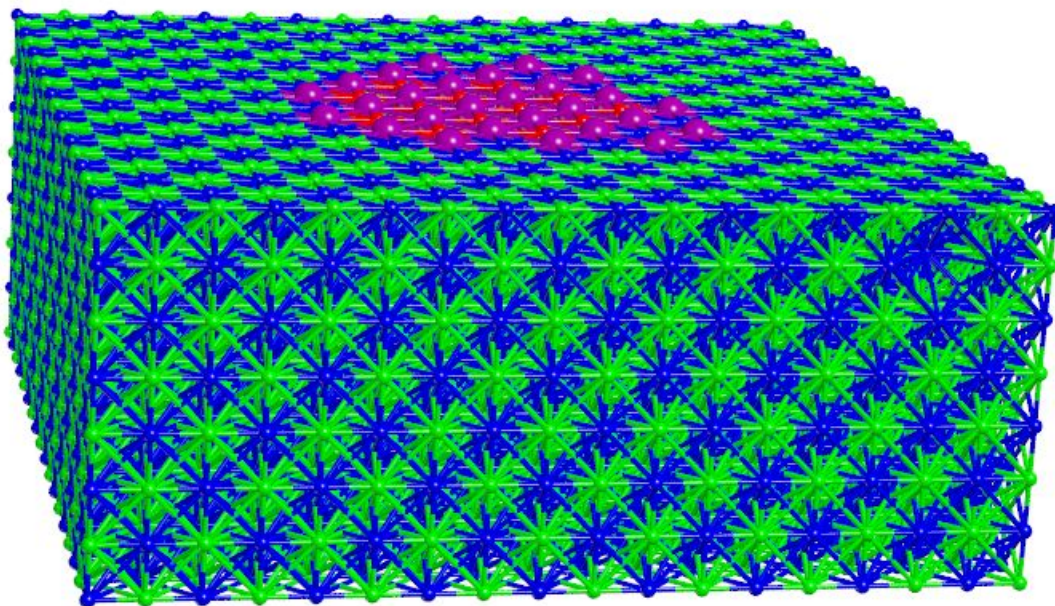
Figure S3 Top view (upper panel) and lateral view (lower panel) of the model for the MgO support. 'Real' atoms are indicated with larger spheres, small spheres indicates the arrays of point charges employed to reproduce the Madelung potential.
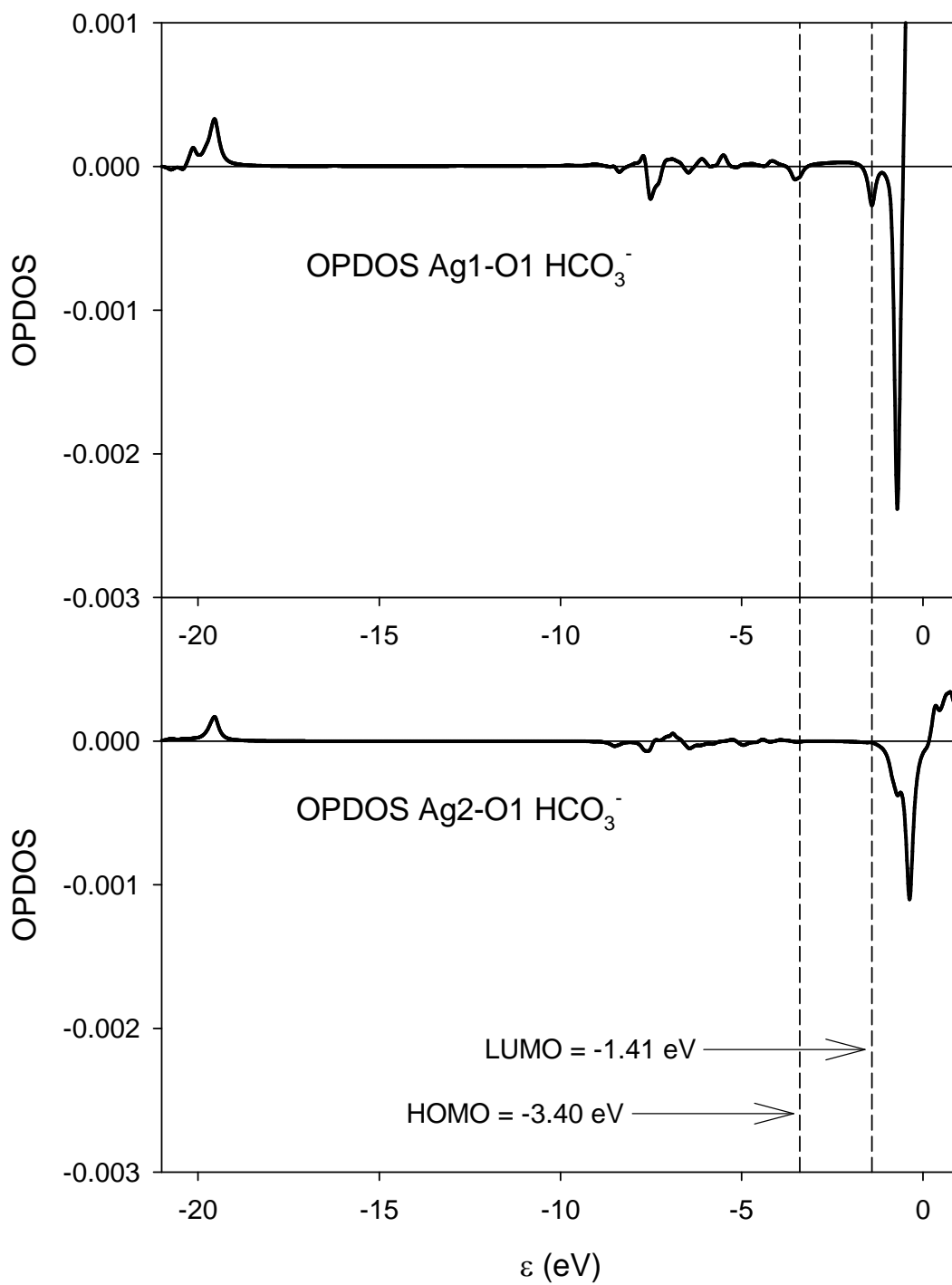
Figure S4. OP-DOS relative to the Ag1-O1 and Ag2-O1 atom pairs for the $Ag_3(HCO_3)(C_2H_4)_2(O)$ complex.
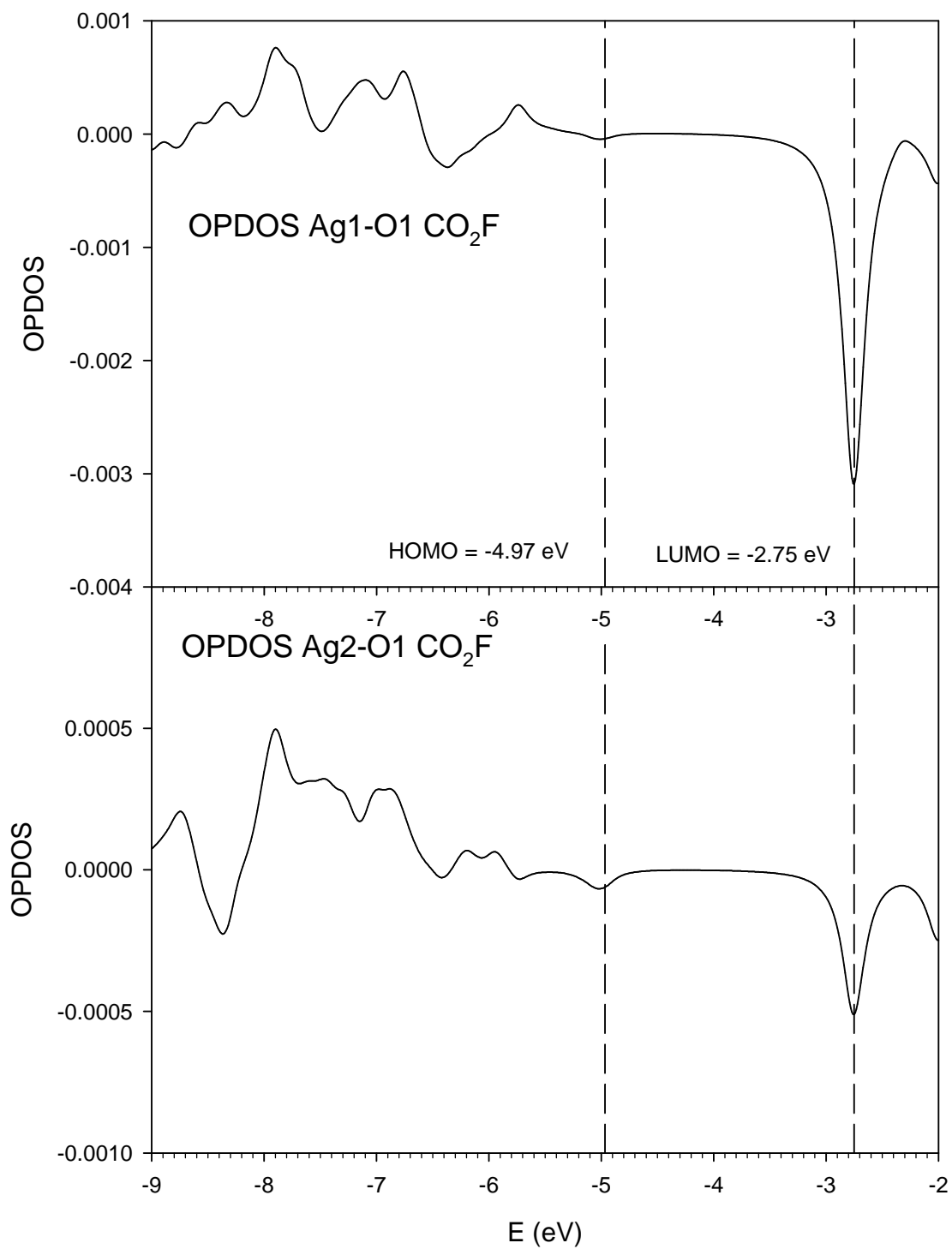
Figure S5. OP-DOS relative to the Ag1-O1 and Ag2-O1atom pairs for the $Ag_3(CO_2F)(C_2H_4)_2(O)$ complex.
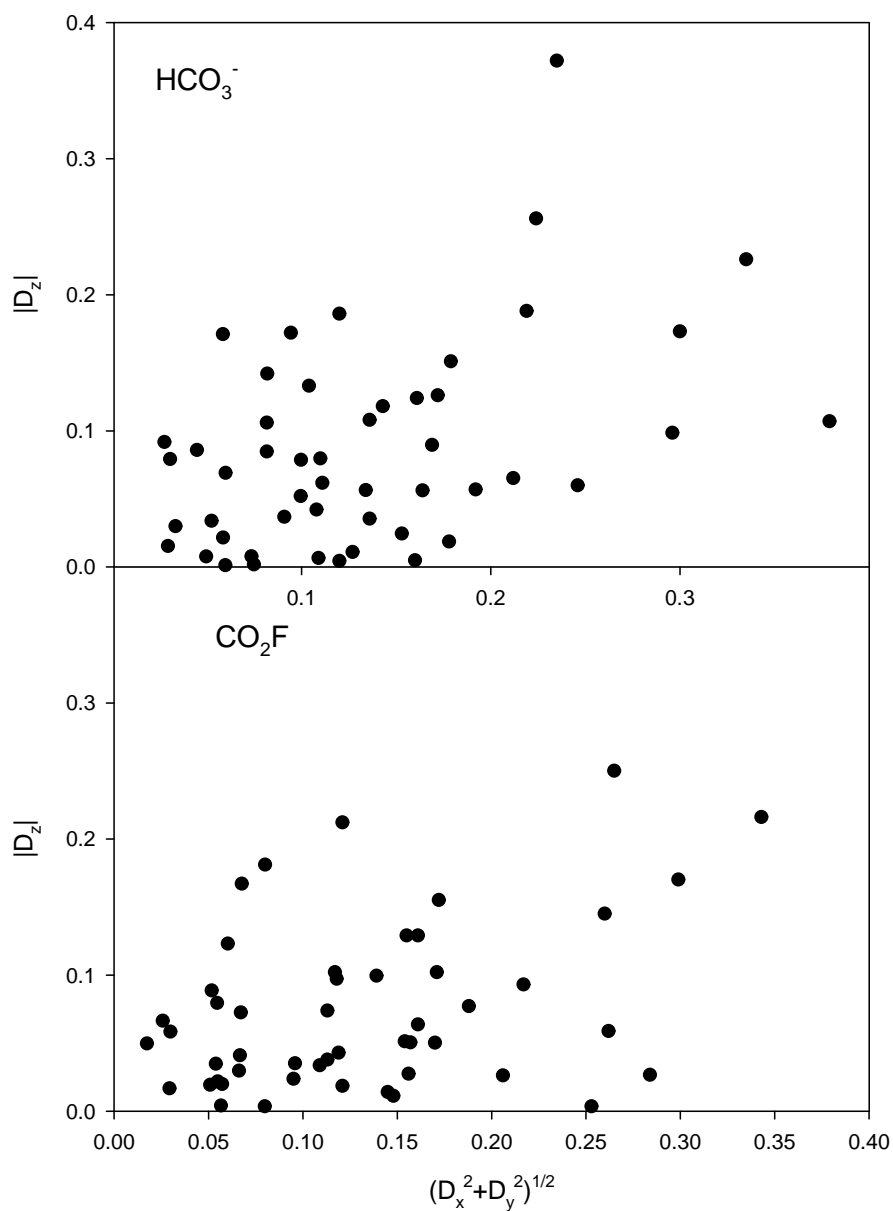
Figure S6. Distribution of the electric dipole transition moment according to their parallel or perpendicular component relative to the MgO surface.
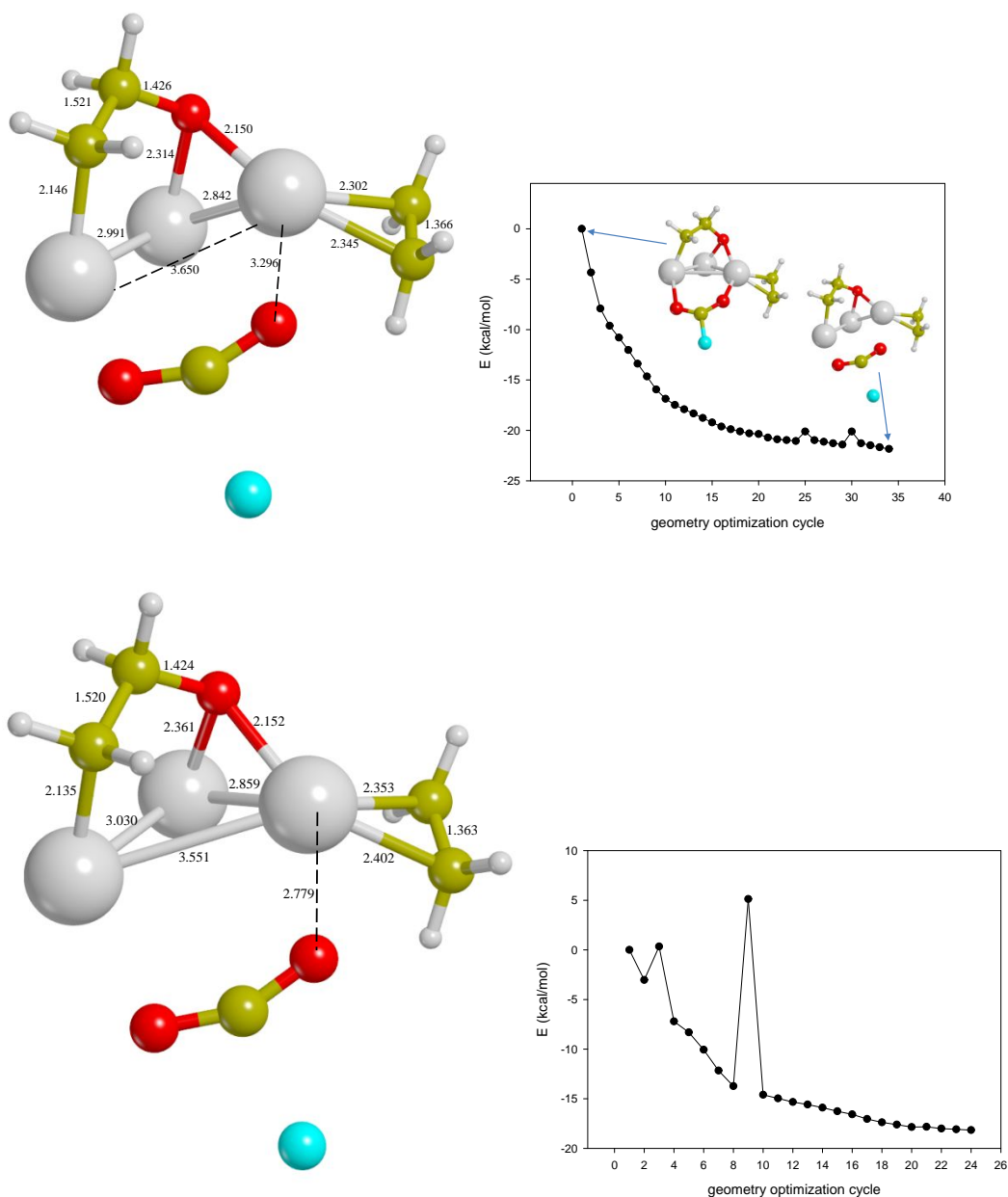
Figure S7. Photo-products obtained from the geometry optimizations of the excited states of Ag$_3$CO$_2$F(Et)$_2$O//MgO(100): state n. 1 (upper panels) and state n. 2 (lower panel). Structural details (left) and optimization path (right).

Table S1. Optimized Ag1-O1 and Ag2-O1 distances and relative energies with respect to optimized Ground-State (GS) for the $Ag_3(HCO_3)(C_2H_4)_2(O)$ complex.

	d(Ag1-O1) Å	d(Ag2-O1) Å	ΔE (eV)
GS	2.290	2.136	0.000
SS1	2.227	2.307	1.264
SS2	2.213	2.314	1.729
SS3	2.205	2.263	1.992
SS4	2.216	2.308	1.939
SS5	2.206	2.320	2.120
SS6*	2.162	2.235	1.463
SS7	2.207	2.311	2.209
SS8*	2.421	2.136	2.807
SS9	2.205	2.338	2.347
SS10	2.213	2.309	2.385
SS11	2.202	2.313	2.470
SS12	2.209	2.333	2.482
SS21**	2.159	2.356	2.837

* the HCO_3 fragment dissociates/decomposes during optimization

** the C_2H_4 fragment dissociates during optimization

Legend: GS initial optimized ground state, SS1 optimized excited state (Singlet-Singlet n. 1)

Table S2. Optimized Ag1-O1 and Ag2-O1 distances and relative energies with respect to optimized Ground-State (GS) for the $Ag_3(CO_2F)(C_2H_4)_2(O)$ complex.

	d(Ag1-O1) Å	d(Ag2-O1) Å	ΔE (eV)
GS	2.252	2.130	0.000
INT	3.527	2.047	0.233
SS1*	2.150	2.314	1.339
SS2*	2.152	2.362	1.645
SS3	2.188	2.224	2.130
SS4*	2.150	2.381	1.983
SS5	2.177	2.352	2.360
SS6*	2.126	2.262	1.403
SS7	2.167	2.389	2.293
SS8	2.168	2.341	2.309
SS9*	2.126	2.276	1.381
SS10	2.167	2.343	2.473
SS11	2.189	2.306	2.542
SS12	2.174	2.340	2.582
SS3_SS1*	2.125	2.331	1.371
INT_SS1	3.493	2.068	1.558

* the FCO_2 fragment dissociates/decomposes during optimization

Legend: GS initial optimized ground state, INT optimized intermediate, SS1 optimized excited state (Singlet-Singlet n. 1), SS3_SS1 optimized excited state n. 1 calculated from SS3 optimized geometry, INT_SS1 optimized excited state n. 1 calculated from INT optimized geometry.

Table S3. Excitation Energy (E), oscillator strength (f), electric dipole transition moments according to their parallel or perpendicular component relative to the MgO surface for the $Ag_3(HCO_3)(C_2H_4)_2(O)$ complex

n	E (eV)	f	$(D_x^2+D_y^2)^{1/2}$	$|D_z|$
1	2.0193	4.13E-04	6.00E-02	6.90E-02
2	2.3024	1.04E-03	1.10E-01	7.96E-02
3	2.3646	9.47E-03	3.35E-01	2.26E-01
4	2.4599	4.35E-04	3.06E-02	7.92E-02
5	2.5302	1.60E-03	1.60E-01	4.85E-03
6	2.5663	2.42E-03	9.44E-02	1.72E-01
7	2.6903	6.18E-04	4.48E-02	8.59E-02
8	2.7145	4.26E-03	2.46E-01	5.99E-02
9	2.729	3.67E-03	1.79E-01	1.51E-01
10	2.7485	2.45E-03	1.69E-01	8.96E-02
11	2.7859	2.36E-03	1.43E-01	1.18E-01
12	2.8535	1.25E-03	8.17E-02	1.06E-01
13	2.8784	1.42E-04	3.35E-02	2.99E-02
14	2.9078	2.86E-03	1.92E-01	5.69E-02
15	2.9142	6.95E-03	2.96E-01	9.85E-02
16	2.9249	8.62E-03	3.00E-01	1.73E-01
17	2.9931	3.62E-03	2.12E-01	6.52E-02
18	3.0154	1.87E-04	4.97E-02	7.62E-03
19	3.025	8.18E-05	2.95E-02	1.53E-02
20	3.0365	2.67E-04	5.98E-02	1.13E-03
21	3.0637	1.21E-03	1.11E-01	6.17E-02
22	3.0756	2.94E-04	5.86E-02	2.15E-02
23	3.0952	4.15E-04	7.36E-02	7.71E-03
24	3.1351	2.32E-03	1.36E-01	1.08E-01
25	3.1618	1.64E-03	1.34E-01	5.64E-02
26	3.1775	2.49E-03	1.78E-01	1.86E-02
27	3.2198	1.07E-03	1.08E-01	4.21E-02
28	3.2331	4.45E-04	7.49E-02	1.72E-03
29	3.2532	3.89E-03	1.20E-01	1.86E-01
30	3.2569	1.01E-03	9.96E-02	5.20E-02
31	3.2789	2.27E-03	1.04E-01	1.33E-01
32	3.3069	1.31E-03	9.98E-02	7.86E-02
33	3.3424	2.46E-03	1.64E-01	5.62E-02
34	3.3475	9.84E-04	1.09E-01	6.49E-03
35	3.3522	1.96E-03	1.53E-01	2.44E-02
36	3.3547	2.21E-03	8.19E-02	1.42E-01
37	3.3869	7.97E-04	9.09E-02	3.67E-02
38	3.3936	1.34E-03	1.27E-01	1.09E-02
39	3.4188	1.16E-03	8.17E-02	8.47E-02

40	3.4399	9.74E-03	2.24E-01	2.56E-01
41	3.4507	3.85E-03	1.72E-01	1.26E-01
42	3.4584	1.22E-03	1.20E-01	4.33E-03
43	3.4685	1.65E-02	2.35E-01	3.72E-01
44	3.481	3.32E-04	5.25E-02	3.38E-02
45	3.4909	7.15E-03	2.19E-01	1.88E-01
46	3.5014	2.80E-03	5.85E-02	1.71E-01
47	3.5108	1.34E-02	3.79E-01	1.07E-01
48	3.5526	1.72E-03	1.36E-01	3.54E-02
49	3.5566	3.59E-03	1.61E-01	1.24E-01
50	3.5596	8.00E-04	2.76E-02	9.17E-02

Table S4. Excitation Energy (E), oscillator strength (f), electric dipole transition moments according to their parallel or perpendicular component relative to the MgO surface for the $Ag_3(CO_2F)(C_2H_4)_2(O)$ complex

n	E (eV)	f	$(D_x^2+D_y^2)^{1/2}$	$|Dz|$
1	2.2728	5.43E-04	6.72E-02	7.24E-02
2	2.4566	1.76E-03	1.39E-01	9.94E-02
3	2.5288	1.02E-02	3.43E-01	2.16E-01
4	2.5967	1.76E-04	1.74E-02	4.96E-02
5	2.705	1.67E-03	1.56E-01	2.74E-02
6	2.7252	5.92E-03	2.60E-01	1.45E-01
7	2.7765	9.66E-04	1.13E-01	3.78E-02
8	2.7994	1.10E-03	1.19E-01	4.28E-02
9	2.8297	5.64E-03	2.84E-01	2.66E-02
10	2.8944	2.12E-03	1.61E-01	6.36E-02
11	2.9091	6.63E-04	5.47E-02	7.94E-02
12	2.9481	2.92E-03	1.55E-01	1.29E-01
13	2.9543	5.21E-03	2.62E-01	5.87E-02
14	2.9626	2.15E-04	5.09E-02	1.93E-02
15	3.0632	8.88E-03	2.99E-01	1.70E-01
16	3.0786	2.95E-03	8.00E-02	1.81E-01
17	3.0859	3.98E-04	6.62E-02	2.96E-02
18	3.0946	3.00E-03	1.71E-01	1.02E-01
19	3.1419	2.03E-03	1.54E-01	5.12E-02
20	3.1637	7.44E-04	9.51E-02	2.37E-02
21	3.1802	2.86E-04	5.73E-02	1.98E-02
22	3.1942	3.22E-04	5.39E-02	3.47E-02
23	3.2101	1.18E-03	1.21E-01	1.84E-02

24	3.2252	2.16E-03	1.57E-01	5.04E-02
25	3.2662	9.12E-05	2.94E-02	1.66E-02
26	3.271	1.76E-03	1.48E-01	1.11E-02
27	3.2894	8.40E-04	9.59E-02	3.51E-02
28	3.2957	3.46E-04	2.99E-02	5.83E-02
29	3.303	1.05E-03	1.09E-01	3.36E-02
30	3.3063	8.51E-04	5.17E-02	8.85E-02
31	3.3098	4.85E-03	1.21E-01	2.12E-01
32	3.35	1.09E-02	2.65E-01	2.50E-01
33	3.3747	4.43E-03	1.72E-01	1.55E-01
34	3.3843	3.51E-03	1.61E-01	1.29E-01
35	3.3916	4.64E-03	2.17E-01	9.29E-02
36	3.3958	2.90E-04	5.49E-02	2.17E-02
37	3.4014	3.60E-03	2.06E-01	2.61E-02
38	3.4112	5.12E-04	6.67E-02	4.09E-02
39	3.417	1.96E-03	1.18E-01	9.71E-02
40	3.4298	3.48E-03	1.88E-01	7.70E-02
41	3.4515	5.41E-04	7.99E-02	3.47E-03
42	3.4828	2.75E-04	5.67E-02	4.04E-03
43	3.4835	2.68E-03	1.70E-01	5.02E-02
44	3.4896	1.60E-03	6.03E-02	1.23E-01
45	3.4979	5.50E-03	2.53E-01	3.42E-03
46	3.5049	4.34E-04	2.58E-02	6.63E-02
47	3.5064	1.57E-03	1.13E-01	7.38E-02
48	3.5244	2.79E-03	6.77E-02	1.67E-01
49	3.5579	1.86E-03	1.45E-01	1.39E-02
50	3.5741	2.11E-03	1.17E-01	1.02E-01

4.2.2   Revealing the electronic properties of the BB bond: the bis-catecholato diboron molecule

In this work, the electronic properties of a diboron molecule, that is, (2-(1,3,2-benzodioxaborol-2-yl)-1,3,2-benzodioxaborole) (B_2Cat_2) were studied by combining experimental XPS and NEXAFS spectra with the corresponding DFT calculations. All the experimental investigations were carried out at the Elettra Syncrotron facility in Trieste by Prof. Albano Cossaro and co-workers. More in detail, the B 1s, C 1s, and O 1s K-edges were investigated for both the isolated gas-phase molecule and the molecule absorbed on the Au(111) surface. Theoretically, both the total and polarized components were studied with the DFT-TP approach, where half an electron is removed from the initial core orbital relaxing all the other ones until self-consistency is reached. All the DFT-TP calculations, as well as TDDFT ones (see below), were performed and analyzed by me.

A good agreement between experimental and calculated data was obtained, with the exception of the B K-edge gas-phase NEXAFS spectrum where significant discrepancies in terms of intensity distribution are found. Such mismatches are still retained switching to TDDFT calculations with different xc functionals. However, it was already demonstrated that the correct reproduction of B 1s NEXAFS features requires the inclusion of static correlation effects that go beyond those described in TDDFT with the coupling of singly-excited configurations. The description of the B 1s spectral features was significantly improved by adopting the ΔSCF-DFT procedure, which treat explicitly the relaxation changes among different excited-state configurations. Furthermore, the comparison of gas-phase and absorbed results revealed that the main features of the polarized NEXAFS spectra are not strongly affected by the presence of the substrate, thus suggesting weak $Au-B_2Cat_2$ interactions.

Check for updates

Revealing the electronic properties of the B–B bond: the bis-catecholato diboron molecule†

D. Toffoli, [ID] [a] C. Grazioli, [ID] [b] M. Monti, [ID] [a] M. Stener, [ID] [a] R. Totani, [ID] [c]
R. Richter, [ID] [d] L. Schio, [ID] [b] G. Fronzoni [ID] *[a] and A. Cossaro [ID] *[ab]

The electronic properties of a diboron molecule, namely bis(catecholato)diboron (2-(1,3,2-benzo-dioxaborol-2-yl)-1,3,2-benzodioxaborole) (B_2Cat_2), have been studied by comparing the results of photoemission (XPS) and near edge X-ray absorption spectroscopy (NEXAFS) experiments with the outcome of DFT calculations. The B 1s, C 1s and O 1s K-edges have been investigated for both the isolated gas phase molecule and the adsorbed one on the Au(111) surface. The main features of the polarized NEXAFS spectra at each of the three edges considered are not significantly affected by the presence of the substrate, with respect to the isolated molecule, indicating that the molecule–gold interaction is weak. Moreover, the comparison between the observed dichroism in the NEXAFS spectra of the adsorbed B_2Cat_2 and that in the NEXAFS spectra of the isolated molecule has confirmed the orbital symmetry assigned in the gas phase absorption spectra. The transitions to π(B–B) bonding and π*(B–B) anti-bonding final states represent the most relevant probe of the chemistry of the B_2Cat_2 molecule. We show that their theoretical description requires that the treatment of the relaxation changes among different excited state configurations, which we successfully implemented by using ΔSCF-DFT (ΔSCF) calculations.

Introduction

Since the last two decades organo-boron chemistry has been playing an emerging role in the development of novel green and sustainable organic-based technologies.[1–3] This promotes the research of convenient synthetic routes for the production of boron compounds. Among the investigated strategies, both metal-catalyzed and metal-free synthetic processes involving diboron(4) compounds as the reactants represent the most promising tool for the synthesis of molecules containing B–C bonds.[4] The properties of the B–B bond of these molecules is the key factor in the synthesis reactions. While being thermo-dynamically stable,[5] its Lewis acid character favors its activation upon interaction with bases.[4,6] The resulting adducts are precursors of the B–B bond cleavage for the formation of boron compounds.[7–11] Besides its employment in synthesis processes, due to this strong affinity with bases, diboron(4) compounds may be of potential interest as building blocks in the 2D supramolecular assembly on surfaces as well. The introduction

of functionalities on surfaces allows tailoring of intermolecular recognition processes and design of specific interaction schemes.[12–14] To our knowledge, and quite surprisingly, the use of diboron molecules to this aim and, in general, the study of the chemistry of the B–B bond on surfaces is mostly unexplored yet. We have recently reported on the on-surface synthesis of a 2D covalent framework supported by the Au(111) surface, formed by a network of B–B bonded boroxine groups,[15] which exhibits delocalized electronic states in the valence band, likely due to hybridization of the outer electronic states of the framework with the substrate. A proper description of the electronic properties of diboron systems on surfaces is mandatory for both exploring their use in synthesis processes catalyzed by the surface and describing the functionalities of 2D architectures they may be part of bis(catecholato)diboron (2-(1,3,2-benzodioxaborol-2-yl)-1,3,2-benzodioxaborole, B_2Cat_2 in the following) (see Fig. 1) is, together with bis(pinacolato)diboron, among the most frequently adopted diboron precursors in the synthesis of boronic compounds. It is a relatively simple molecule we selected as a prototype of diboron species and we present here a combined theoretical and experimental study of its electronic properties by comparing photoemission (XPS) and near edge X-ray absorption spectroscopy (NEXAFS) with DFT calculations. Experiments have been performed both for the isolated molecule in the gas phase and on a poorly reactive surface like Au(111) to evidence dichroic effects in the

[a] Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy. E-mail: fronzoni@units.it, acossaro@units.it

[b] IOM-CNR, Istituto Officina dei Materiali-CNR, S.S.14, Km 163.5, 34149 Trieste, Italy

[c] ISM-CNR, Istituto di Struttura della Materia, LD2 Unit, 34149 Trieste, Italy

[d] Elettra-Sincrotrone Trieste, 34149 Basovizza, Trieste, Italy

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cp03428f

This journal is © the Owner Societies 2021

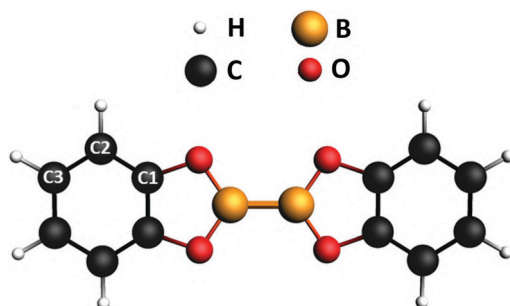*Phys. Chem. Chem. Phys.*, 2021, **23**, 23517–23525 | **23517**

Fig. 1 Chemical structure of B_2Cat_2. The labels indicate the nonequivalent carbon atoms. The molecular plane corresponds to the xy plane, with the direction of the B−B bond along the x axis.

absorption. Dichroic effects on the Au(111) surface have been studied by measuring NEXAFS spectra in p-polarization (p-pol), *i.e.* with the light electric field plane perpendicular to the metal surface, and in s-polarization (s-pol), *i.e.* with the light electric field plane parallel to the metal surface.

From a general perspective, the theoretical description of B 1s core excitations of boronic- and B–B containing systems is a challenging task, as static correlation effects between different excited states need to be taken into account.[16] This can be done using wave function based approaches, which becomes however prohibitively expensive when the size of the system increases, as in the case of the B_2Cat_2 molecule. Therefore, it is necessary to search for an alternative, less expensive method to calculate the B 1s core excited states. In the ΔSCF-DFT (ΔSCF) method, based on the energy difference between two self-consistent-field (SCF) calculations, the core excited state is found by promoting an electron from the core orbital to one of the unoccupied orbitals in the ground state. The MOs are then re-optimized for each excited electron configuration and a different set of MOs for each excited state is obtained. Relaxation changes among different excited-state configurations relative to the same core hole are therefore taken into account at the ΔSCF level, while they are not considered at the DFT level employing transition potential (TP) approximation[17,18] (DFT-TP), which employs the same fixed set of orbitals obtained from a single TP calculation to describe the entire excited state manifold.

We show here how the ΔSCF calculations better reproduce the B K-edge NEXAFS experimental spectrum with respect to both DFT-TP and time-dependent DFT (TDDFT). We give a detailed description of the transitions corresponding to the peaks of the NEXAFS O, B and C K-edges and to the geometry of the related molecular orbitals.

Theoretical methods and computational details

The ground-state equilibrium geometry of B_2Cat_2 was computed at the density functional theory (DFT)[19] level, employing the hybrid B3LYP xc functional and a basis set of Slater-type orbitals (STO) of triple-zeta quality (TZP) for all atoms. Calculations were

performed by using the Amsterdam Density Functional (ADF) quantum chemistry code.[20] The geometrical parameters (bond lengths and bond angles) of the optimized geometry are presented in Fig. S1 and Table S1 of the ESI† and show a close agreement with the experimental data derived from the crystallographic structure.[21] The molecule is planar (D_{2h} point group symmetry) and the molecular plane corresponds to the xy plane in the molecular frame, with the B–B bond along the x direction (see Fig. 1).

The XPS and NEXAFS spectra at the B 1s, C 1s and O 1s edges were computed at the DFT level employing the PW86xPerdew (PW86) generalized gradient approximation (GGA) for the xc functional[22,23] as implemented in the ADF program.[20,24]

In order to accurately describe higher excitations close to the ionization threshold, an even tempered Quadruple-Zeta basis set with three polarization and three diffuse functions (designated as ET-QZ3P-3DIFFUSE in the ADF database) has been employed for the core-excited atom (B, C or O, respectively), while a TZP (Triple Zeta + polarization) basis set was used for the remaining atoms. In particular, a frozen core TZP.1s basis set has been employed for the B, C and O non-excited atoms, to ensure the correct localization of the half core hole.

For the XP spectrum simulation, the B 1s, O 1s and C 1s core Ionization Potentials (IPs) have been calculated at the ΔSCF Kohn–Sham level allowing a full relaxation of the ionized core hole. The energy of each $1s^{-1}$ ionic state has been obtained through a SCF unrestricted calculation. The relative intensity among different primary lines is simply assumed to be proportional to the number of equivalent atoms of the same type. This is a good approximation since the XPS spectra were measured with a photon energy well above the considered ionization thresholds.

The B, O and C K-edge NEXAFS spectra at each nonequivalent atomic center have been computed with the DFT-TP method in which half an electron is removed from the initial core orbital and all the orbitals are relaxed until self-consistency is obtained. Relaxation effects upon the core-hole formation are usually adequately included in the DFT-TP approach.[25] The performance of the hybrid B3LYP potential[26–28] has also been tested considering the B 1s core excitation spectrum (see Fig. S5 of the ESI†). No significant differences emerged from the comparison between the PW86 and B3LYP calculated spectra.

For each of the three non-equivalent C sites of the B_2Cat_2 molecule (see Fig. 1), a separate computation of the C 1s excitation spectrum has been performed and the total spectrum has been obtained by summing up the partial contributions weighted for the number of equivalent atoms.

Within the DFT-TP scheme, the excitation energies correspond to the differences between the eigenvalues of the virtual orbitals and the 1s core orbitals referred to the TP configuration:

$$\Delta E_{i \to f} = \varepsilon_f^{TP} - \varepsilon_i^{TP}, \qquad (1)$$

while the transition intensities are expressed as oscillator strengths $f_{i \to f}$. For a free molecule $f_{i \to f}$ corresponds to

$$f_{i \to f} = \frac{2}{3} n_i \Delta E_{i \to f} \left| \left\langle \varphi_f^{TP} \middle| \boldsymbol{\mu} \middle| \varphi_i^{TP} \right\rangle \right|^2 \qquad (2)$$

23518 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 23517–23525

This journal is © the Owner Societies 2021

and involves matrix elements of the electric dipole operator between initial and final TP MOs, where n_i denotes the occupation number of the core orbital in the ground state. Since the DFT-TP approach leads generally to a less attractive potential and consequently the excitation energies are generally too large, they have been shifted with respect to the ΔIP value by an amount given by the energy difference $\varepsilon_{1s}^{TP} - \Delta SCF(1s)$. The energy of the ΔSCF $1s^{-1}$ ionic state corresponds to that previously calculated for the XP spectra.

Due to the discrepancies between experiments and the theoretical DFT-TP results of the B 1s NEXAFS spectrum, the B 1s core excitations of the gas-phase B_2Cat_2 have been also computed at the TDDFT level, which formally includes the coupling between the single excited configurations, employing different xc functionals with the ground state electron configuration. The TDDFT computational details and the results are reported in Section S4 of the ESI,† which show that no improvement is obtained at the TDDFT level compared to the simpler DFT-TP scheme.

To better describe electronic relaxation effects, the B 1s core excitations and oscillator strengths have been also calculated at the ΔSCF level employing the procedure described in the following. In the ΔSCF method, the initial (Ψ_i) and final (Ψ_f) N-electron wave functions entering the dipole matrix element (computed in the length gauge of the dipole operator),

$$\mu_{i \to f} = \langle \Psi_f | \hat{\mu} | \Psi_i \rangle, \tag{3}$$

are Slater determinants constructed from Kohn–Sham molecular orbitals (MOs) obtained with the SCF procedure relative to the ground state (GS) and excited state occupation numbers, respectively. The GS MOs are obtained from a spin-restricted calculation, while the excited-state MOs are calculated within a spin polarized scheme with $N_\alpha - N_\beta = 0$, where N_α and N_β denote the numbers of spin-up and spin-down electrons, respectively. In the specification of the occupation numbers, we removed a β electron from the B 1s core. Singlet core-excitation energies are obtained according to the spin-purification formula.[29] Since the two sets of MOs, $\{\varphi_\mu^i\}$ and $\{\varphi_\lambda^f\}$, used to construct Ψ_i and Ψ_f, respectively, are non-orthogonal, the N-electron matrix elements in eqn (3) are evaluated with the general rules for non-orthogonal spin orbitals derived by Löwdin.[30] Denoting with S_{fi} the overlap matrix between the two sets of occupied MOs, $(S_{fi})_{\lambda\mu} = \langle \varphi_\lambda^f | \varphi_\mu^i \rangle$, $\mu_{i \to f}$ of eqn (3) can be written as

$$\langle \Psi_f | \hat{\mu} | \Psi_i \rangle = \sum_{\lambda\mu} \langle \varphi_\lambda^f | \hat{\mu} | \varphi_\mu^i \rangle \mathrm{adj}(S_{fi})_{\lambda\mu}, \tag{4}$$

in terms of dipole matrix elements between the two sets of MOs and the adjugate of S_{fi} (i.e. the transpose of its cofactor matrix). When $\langle \Psi_f | \Psi_i \rangle = \det(S_{fi}) \neq 0$, eqn (4) reduces to

$$\langle \Psi_f | \hat{\mu} | \Psi_i \rangle = \det(S_{fi}) \sum_{\lambda\mu} \langle \varphi_\lambda^f | \hat{\mu} | \varphi_\mu^i \rangle (S_{fi}^{-1})_{\lambda\mu}. \tag{5}$$

The origin independence of the transition matrix elements is enforced by adding the dipole of the nuclear charges, weighted by the overlap $\langle \Psi_f | \Psi_i \rangle$. Results of a recent study[31] indicate that this correction gives results of accuracy comparable to that obtained by enforcing exact orthogonality of Ψ_i and Ψ_f. In any case, when exact orthogonality is not enforced by symmetry, we checked that the overlap between the initial and final states is rather small ($\sim 10^{-3}$ or smaller).

In order to compare the theoretical results with the NEXAFS experimental measurements recorded for the B_2Cat_2 molecule adsorbed on the Au(111) surface (in s- and p-polarization), we have considered the molecule fixed in space and have calculated the NEXAFS p- and s-pol intensities† employing the following formula for the oscillator strength:

$$f_{i \to f} = 2\Delta E_{i \to f} \left| \langle \Psi_f | \bar{e} \cdot \bar{\mu} | \Psi_i \rangle \right|^2 \tag{6}$$

where Ψ_i and Ψ_f are the initial and final TP MOs in the DFT-TP calculations and Slater determinants with the GS and excited state occupation numbers, respectively, in the ΔSCF calculations.

The ΔSCF B 1s, C 1s and O 1s IPs are presented in the figure of the corresponding NEXAFS spectrum and mark a separation of the whole spectral region into two sub-regions with different physical meanings. Only transitions that occur below the ionization threshold IP can be accurately described by the employed computational protocol. Above the IP only qualitative information can be extracted since the electronic continuum wave function cannot be properly described with standard basis sets of quantum chemistry programs.

Results and discussion

In the first part of the discussion the experimental XPS and NEXAFS spectra at the B 1s, O 1s and C 1s edges of the gas phase B_2Cat_2 will be compared with the theoretical results in order to achieve an assignment of the spectral features. In the second part, the polarization dependent NEXAFS measurements of the B_2Cat_2 adsorbed on the Au(111) surface will be analysed with the support of the theoretical results obtained for the fixed-in-space molecule. The experimental data have been acquired at the Gas Phase beamline[32] and at the ANCHOR-SUNDYN endstation[33] at the Elettra Synchrotron. Experimental details are reported in the ESI.†

Gas phase results

B 1s, C 1s and O 1s photoelectron spectra. The experimental XPS spectra of the B 1s, C 1s and O 1s core levels of B_2Cat_2 together with the calculated ΔSCF ionization potentials (IPs) for the B, C and O atoms are reported in the ESI† (Fig. S2, S3 and S4, respectively). Here we briefly comment on the main features observed. The experimental XP spectrum of the B 1s core level (Fig. S2, ESI†) shows a sharp intense peak at 196.89 eV which is close to the main B 1s peak at 197.53 eV of gas-phase catecholborane.[34] The latter is the molecule whose structure corresponds to half B_2Cat_2, with a terminal H bonded to the B atom. The difference in the B 1s binding energy we reported has therefore to be ascribed to the different chemical environment given by the presence of the B–B bond. The two much weaker structures here observed at 202.13 and 202.71 eV are attributed

Paper

to satellite states of the B 1s core level like those at 200.77 and 203.25 eV of catecholborane. The many-body nature of these structures cannot be described by the present computational approach.

The C 1s photoelectron spectrum of B_2Cat_2 is presented in Fig. S3 of the ESI† together with the calculated ΔSCF IPs of the three nonequivalent carbon atoms of the molecule (C1, C2 and C3, see Fig. 1). The two experimental peaks are assigned to the C2, C3 sites (first peak at 290.49 eV) and to the C1 site (second peak, at 291.94 eV); their energy splitting (1.45 eV) is a consequence of the reduced shielding of the C1 site, which is adjacent to the O atom, with respect to C2 and C3 sites. Calculations provide a chemical shift of 131 meV between the C2 and C3 sites with the lower binding energy associated to C3. The experimental line shape shows a pronounced asymmetry, detected also in catecholborane[34] and attributed to vibrational effects. The spectrum was fitted with a function of three skew gaussians and the extracted peak positions are 291.95, 290.58 and 290.38 eV in accordance with the calculated IPs.

The O 1s photoelectron spectrum of B_2Cat_2 is shown in Fig. S4 of the ESI.† The peak at 539.03 eV represents the IP of the four equivalent oxygen atoms, as also indicated by the calculated ΔSCF IP line. The second intense peak around 541 eV is due to the O 1s signal of CO_2 that we dosed together with B_2Cat_2 in order to have a known feature in the spectra to properly calibrate the photon energy scale. Details about the calibration references we used are reported in the ESI.†

B 1s, C 1s and O 1s NEXAFS spectra. The B K-edge NEXAFS experimental spectrum is shown in Fig. 2 together with the theoretical results relative to the ΔSCF and DFT-TP calculations. An intense lower energy peak (peak A at 191.362 eV) dominates the experimental spectrum, followed by a weaker double-peaked feature between 194 and 195 eV and by a poorly structured signal towards a higher energy. The DFT-TP results (lower panel of Fig. 2) show significant discrepancies with the experiment; in particular, the calculations do not reproduce correctly the intensity distribution between the first peak and the feature around 195 eV for which the calculations overestimate the oscillator strengths for the transitions involved. The inability of TDDFT to correctly describe the B 1s core excitations emerges in the case of the B_2Cat_2 molecule, as shown by the results presented in Fig. S6 of the ESI:† the discrepancies with respect to the B 1s experimental spectrum persist in terms of both energy separation and intensity distribution irrespective of the exchange–correlation functional employed (LB94, B3LYP and CAM-B3LYP).

The origin of the incorrect intensity distribution calculated at the DFT-TP level for the B 1s transitions has been already discussed in previous studies on boronic acid derivatives.[16,35] We demonstrated that to correctly reproduce the B 1s NEXAFS spectrum one needs to include static correlation effects beyond those described by the coupling of singly-excited configurations, as in the TDDFT approach. In ref. 16 this was achieved by performing *ab initio* Multi-Configuration Self-Consistent-Field (MCSCF) calculations. We are led to assume that the inclusion of residual static correlation effects plays an important role in
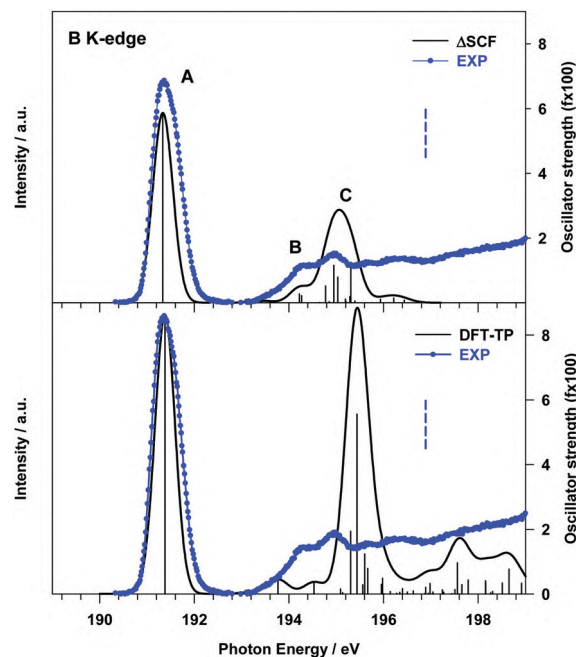

Fig. 2 B K-edge gas phase NEXAFS spectrum of B_2Cat_2. Upper panel: Comparison with the theoretical ΔSCF results; lower panel: comparison with the theoretical DFT-TP results. The theoretical spectra have been shifted by +0.45 eV (ΔSCF spectrum) and +0.17 eV (DFT-TP spectrum) in order to match the first experimental peak. The stick spectra are broadened by using a Gaussian line shape with FWHM = 0.5 eV. The experimental B 1s IP (196.89 eV) is indicated with a vertical dashed line.

the description of the B 1s photoexcitation of B_2Cat_2, as also evidenced by the shake-up structures of the XPS B 1s spectrum in Fig. S2 (ESI;† see the "B 1s, C 1s and O 1s photoelectron spectra" section).

The effect of a more accurate treatment of relaxation, which is the dominant electron correlation effect at play in core excitation processes, is apparent from the comparison between the B 1s experimental NEXAFS spectrum and the ΔSCF results, presented in the upper panel of Fig. 2.

A significant improvement of the description of the experimental spectrum is obtained compared to DFT-TP results, in particular as concerns the intensity distribution among the spectral features. The ΔSCF spectrum has been obtained by calculating, for each dipole-allowed final symmetry, the first six excitation energies and transition dipole moments. This choice has been supported by the analysis of the DFT-TP results, which allows the identification of the final MOs involved in the description of the main spectral features below the edge. The ΔSCF B 1s core excitation energies and oscillator strengths are presented in Table 1, together with the attribution of the spectral features and the experimental data. The first calculated peak A arises from the B 1s transition to the LUMO orbital ($7b_1$) which has a predominant π(B–B) bonding character (out of plane (oop) B 2p atomic components) with minor π*(B–O) antibonding contributions (the 3D plot of the $7b_1$ final MO is presented in Fig. S7 of the ESI†).

The two main transitions of feature C (at 194.95 and 195.31 eV) involve final MOs with mainly valence π*(B–B) (out of plane)

23520 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 23517–23525

This journal is © the Owner Societies 2021

Table 1 B 1s ΔSCF excitation energies (eV) and oscillator strengths f of the main transitions of B_2Cat_2

Peak	ΔSCF transition[a]	E^b (eV)	$f \times 100^c$	Assignment, main character of the final MO
A	B 1s → 7b$_1$	191.33	5.87	$\pi_{\text{out of plane}}$(B–B) + $\pi^{\star}$(B–O)
B	B 1s → 29a$_1$	194.22	0.27	Rydberg (in plane AO components)
	B 1s → 9b$_1$	194.27	0.21	$\pi^{\star}$(C–C)
C	B 1s → 23b$_2$	194.78	0.52	Mixed valence $\pi_{\text{in plane}}$(B–B)/Rydberg
	B 1s → 10b$_1$	194.95	1.15	Mixed valence $\pi^{\star}_{\text{out of plane}}$ (B–B)/Rydberg
	B 1s → 24b$_2$	195.03	0.79	Mixed valence $\pi_{\text{in plane}}$(B–B)/Rydberg
	B 1s → 11b$_1$	195.31	1.25	Mixed valence $\pi^{\star}_{\text{out of plane}}$ (B–B)/Rydberg

[a] When the core hole is localized on the B site the molecular symmetry lowers from D_{2h} to C_{2v} and the molecular plane is yz with the B–B bond along the z axis. [b] Calculated excitation energies shifted by 0.45 eV to match the first experimental peak. [c] Only calculated transitions with $f \times 100 \geq 0.20$ are reported.

antibonding character mixed with diffuse atomic components, while MOs with π(B–B) (in plane) bonding character, mixed with diffuse components, characterize the other two transitions at 194.78 and 195.03 eV (see Fig. S7 of the ESI†). The lower energy tail of C, labelled B in Fig. 2, arises from two low intensity transitions, one of diffuse nature (at 194.22 eV) and the other of mainly $\pi^{\star}$(C=C) antibonding character (at 194.27 eV).

The analysis of the contributions to the transition dipole moment (see eqn (4) and (5)) for the ΔSCF 9a$_1$ (B 1s^{-1}) → 7b$_1$ transition reveals that the intensity is mainly contributed by the $\left\langle \varphi^f_{7b_1} \left| \hat{\mu} \right| \varphi^i_{9a_1} \right\rangle$ dipole matrix element between the initial 9a$_1$ MO of the GS and the 7b$_1$ MO of the ΔSCF excited state. The value of the cofactor, which multiplies the dipole integral, is close to 1 (0.918), and thus it does not influence the calculated intensity of this transition. On this ground, the relatively large oscillator strength computed for the ΔSCF 9a$_1$ → 7b$_1$ transition can be related to the composition of the 7b$_1$ final MO, notably the significant participation of the B 2p atomic components, which carries most of the dipole transition moment. Considering the transitions of peak C, this kind of analysis still points out that the intensity of each transition is dominated by the contribution of the dipole matrix element between the initial 9a$_1$ MO (B 1s^{-1}) of the GS and the final MO of the ΔSCF transition considered.

Also the relative intensity of these transitions, compared to the first peak, can be correlated to the atomic composition of the final MO, in particular the B 2p/np atomic contributions, which turn out to be reduced with respect to the final 7b$_1$ MO of the first transition.

The significant decrease of the oscillator strength obtained at the ΔSCF level compared to the DFT-TP ones for the transitions of peak C can be mainly ascribed to the decreased value of the dipole integral of the corresponding transitions and therefore to the relaxation changes among different excited state configurations neglected in the DFT-TP approach, which treats the relaxation effects in an average way. We underline that the value of the cofactor multiplying the dominant dipole integral of the transitions of peak C is lower (around 0.8) than the one corresponding to the 9a$_1$ (B 1s^{-1}) → 7b$_1$ transition and the calculated intensity is further reduced compared to the DFT-TP results.

The intensity distribution as well as the energy separation between the π(B–B) bonding and the $\pi^{\star}$(B–B) anti-bonding final

states we described here represents a sensitive monitor of the chemistry the molecule is involved in, recalling what we discussed in the introduction. The Lewis acid character of the molecule is expected to be related to these orbitals, as they are localized on the B atoms.

The gas phase experimental C K-edge NEXAFS spectrum of B_2Cat_2 is shown in Fig. 3 together with the theoretical DFT-TP results. The calculated excitation energies and oscillator strengths are presented in Table 2, along with their assignment to single-particle transitions. Two main experimental peaks characterize the low energy region (around 285 and 287 eV, respectively), while several less intense and broader structures are present in going towards the ionization threshold. The spectrum resembles the one measured for catecholborane,[34] as expected due to the same chemical configuration for the carbon atoms in the two cases.

The comparison with the theory points out a good agreement for the groups of excited states which contribute to the first structures (A, B and C, with C describing the higher energy tail of the second experimental band). Conversely, a poorer agreement is observed at a higher energy, in particular in the energy region around 290 eV with an overestimation of the intensity provided by the calculations. Indeed, the close
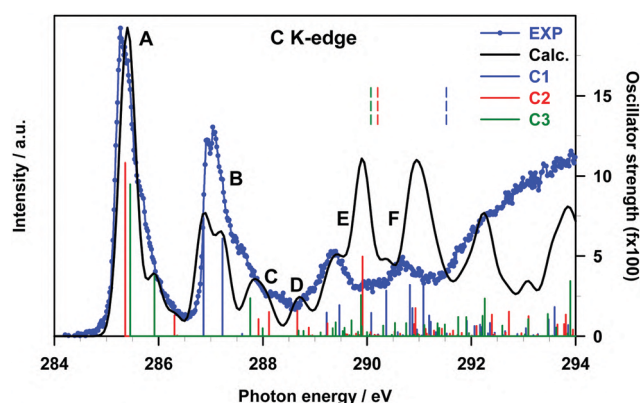
Fig. 3 C K-edge gas phase NEXAFS spectra of B_2Cat_2. The DFT-TP theoretical spectrum has been rigidly shifted by −0.37 eV on the experimental energy scale. The main calculated structures are labelled with capital letters. Colored solid lines represent the partial Ci contributions. ΔSCF C 1s IPs (C1: 291.89 eV; C2: 290.58 eV; C3: 290.45 eV) are indicated with vertical dashed lines. The stick spectrum is broadened by using a Gaussian line shape with FWHM = 0.35 eV.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 23517–23525 | **23521**

Table 2 DFT-TP C 1s excitation energies (eV) and oscillator strengths f of B_2Cat_2

Peak	E^a (eV)	f^b	Site	Assignment
A	285.36	10.81	C2	$\pi^*(C–C) + \pi(B–B)$ (LUMO)
	285.46	9.49	C3	$\pi^*(C–C) + \pi(B–B)$ (LUMO)
	285.92	3.79	C3	$\pi^*(C–C) + \pi(B–B)$ (LUMO+1)
	286.30	1.30	C2	
B	286.86	7.38	C1	$\pi^*(C–C) + \pi(B–B)$ (LUMO)
	287.22	6.08	C1	$\pi^*(C–C) + \pi(B–B)$ (LUMO+1)
C	287.76	2.38	C3	Rydberg (in plane components)
	288.12	1.51	C2	
D	288.66	1.60	C2	
E	289.23	1.47	C1	
	289.47	1.94	C1	
	289.89	2.57	C3	Mixed $\pi^*(C–C)$/Rydberg
	289.91	4.95	C2	
	290.08	1.45	C1	Rydberg (in plane components)
	290.37	2.85	C1	Mixed $\pi(B–B)$/Rydberg (in plane components)
F	290.82	3.21	C1	Mixed valence $\pi^*(C–C)$/Rydberg
	291.08	3.33	C1	

^{a} Calculated excitation energies shifted by -0.37 eV to match the first experimental peak. ^{b} Only the main transitions contributing to the theoretical features are presented.

Fig. 4 O K-edge gas phase NEXAFS spectra of B_2Cat_2. The DFT-TP theoretical spectrum has been shifted by 0.10 eV in order to match the first experimental peak. The main calculated structures are labelled with capital letters. The ΔSCF O 1s IP (539.03 eV) is indicated with a vertical dashed line. The stick spectrum is broadened by using a Gaussian line shape with FWHM = 0.8 eV.

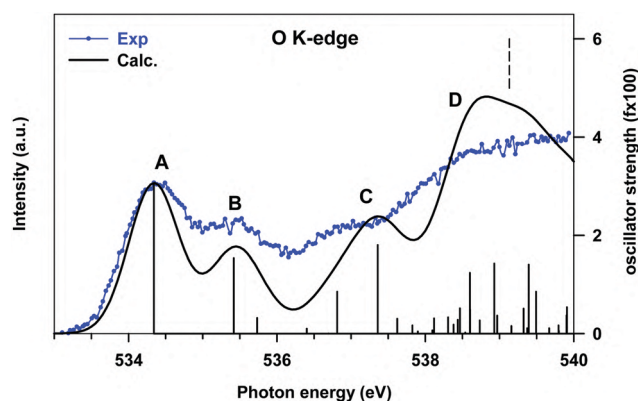**Table 3** DFT-TP O 1s excitation energies (eV) and oscillator strengths f of B_2Cat_2

Peak	E^a (eV)	$f \times 100^b$	Assignments
A	534.34	3.06	$\pi_{oop}(B–B) + \pi^*(B–O)$
B	535.42	1.54	$\pi^*(C{=}C) + \pi^*(C{=}O)$
C	536.81	0.85	Rydberg (in plane components)
	537.36	1.80	
D	538.60	1.24	Mixed valence $\pi^*(B–B) + \pi^*(B–O)$/Rydberg
	538.93	1.43	Rydberg (in plane components)

^{a} Calculated excitation energies shifted by 0.10 eV to match the first experimental peak. ^{b} Only the main transitions contributing to the theoretical features are reported.

proximity of the ionization region, above which the discretization of the non-resonant continuum affects the calculations, can explain the only qualitative description of the intensity distribution around and above 290 eV. Nonetheless, the attribution of the experimental features, which we will present in the following, can still be considered with confidence.

Band A arises from transitions from the C2 and C3 sites of the phenyl rings, namely the two C 1s → LUMO transitions, which contribute to the main peak, and the two C 1s → LUMO+1 transitions, which contribute to the higher energy tail of peak A (see Table 2). The same transitions from the C1 site, which is bonded to the O atom of the catechol moieties, are shifted to higher energies and give rise to peak B. The energy separation between the LUMO transitions of peaks A and B (about 1.4 eV) reflects the shift of IPs between C2/C3 and C1 sites (1.37 eV), denoting a predominance of initial state effects on these excitation energies. Both LUMO and LUMO+1 final orbitals are localized on the phenyl ring and present a $\pi^*(C–C)$ antibonding character with a minor contribution from the $\pi(B–B)$ bond between the B atoms connecting the two catechol moieties (a plot of the LUMO orbital with the core hole localized on the C2 site is presented in Fig. S8 of the ESI†).

The higher energy tail of peak B (denoted as C in Fig. 3) as well as feature D and the lower energy shoulder (around 289.5 eV) of peak E are contributed by several low intensity transitions towards final MOs of diffuse nature (see Table 2). In the energy region of 289.5–291 eV the agreement between the theory and experiment deteriorates, as previously commented: the calculations provide two more intense peaks (E and F) around the C 1s thresholds, while two broad features, slightly shifted to lower energies, are present in the experiment. The high intensity of E and F theoretical peaks mainly arises from transitions towards final MOs with a $\pi^*(C–C)$ antibonding character mixed with some diffuse components starting from all the C sites (peak E) and from the C1 site (peak F). The O

K-edge NEXAFS experimental spectrum of gas phase B_2Cat_2 is compared with the DFT-TP theoretical spectrum in Fig. 4. The assignment of the main peaks in terms of the calculated excitation energies and oscillator strengths is presented in Table 3. The experimental spectrum shows two main features in the energy range up to 536 eV, followed by a quite smooth signal of increasing intensity towards the ionization threshold. The good agreement between the calculated and measured spectra confirms the adequacy of the DFT-TP scheme to describe the core excitations of the O 1s channel in the B_2Cat_2 molecule.

The stronger calculated transition at 534.34 eV gives rise to peak A and involves the LUMO orbital which presents mainly a $\pi(B–B$, out of plane) bonding character with a minor $\pi^*(B–O)$ antibonding contribution (see Fig. S9 of the ESI†), similarly to the LUMO orbital obtained from the DFT-TP calculation with the core-hole localized on the B site.

Peak B mainly arises from the O 1s → LUMO+1 transition. The LUMO+1 MO is strongly localized on the phenyl ring ($\pi^*(C{=}C)$ character) with only a weak conjugation with the O $2p_z$ components, consistent with the low intensity calculated for this transition. Features C and D are characterized by several low intensity transitions toward MOs of mainly diffuse nature.

23522 | Phys. Chem. Chem. Phys., 2021, **23**, 23517–23525

This journal is © the Owner Societies 2021

It is worth noting that the transition at 538.60 eV (peak D) involves a final orbital which is the antibonding counterpart of the LUMO although several diffuse components are also present. The energy separation between the transitions towards the π(B–B) bonding and antibonding orbitals is about 4 eV, a value very close to that found for the corresponding transitions in the B 1s NEXAFS spectrum.

We underline the relatively low oscillator strengths provided by the DFT-TP calculations for the O 1s transitions, which stem from the small O 2p atomic contribution to the virtual orbitals compared to the C and B atomic contributions. This is a sign of the polarization of the C–O and B–O bonds and of the large involvement of O 2p atomic components in the occupied MOs.

B_2Cat_2 on Au(111) NEXAFS results. The NEXAFS experimental spectra of B_2Cat_2 on Au(111) have been recorded at two different polarization angles of the electric field with respect to the surface. p-pol and s-pol geometries refer to the light polarization vector being normal and parallel to the surface plane, respectively. The calculated p- and s-pol spectra are relative to a gas-phase molecule oriented with the molecular plane fixed in the xy plane of the molecular frame (see the Theoretical methods and computational details section). All B 1s, C 1s and O 1s spectra show a strong dichroism between s-pol and p-pol spectra, with the transitions to π^* orbitals having a maximum intensity in p-pol. This indicates a flat adsorption geometry for the B_2Cat_2 on the Au(111) surface.

The B 1s experimental and calculated p- and s-pol spectra are presented in Fig. 5, in the energy region below the threshold. The intensity contribution from the p-pol component is prevalent in all the spectral region below the edge. The most remarkable dichroic effects are present in the region of the first experimental peak (around 191 eV) and are correctly reproduced by the calculations. This behavior reflects the π symmetry of the final orbital ($7b_1$, LUMO) involved in the transition responsible for this peak. In the energy region above 194 eV

both p- and s-pol components are present. In particular, the s-pol component gives a significant contribution around 195 eV, which appears slightly underestimated by the theory and corresponds to the transitions encompassed by peak C in the gas phase spectrum, assigned to the excitations to the in-plane π(B–B) MOs.

In this energy region the p-pol experimental component is still stronger than the s-pol one, with a large feature with a maximum around 196 eV, while the theory provides one p-pol peak mainly derived from the two out of plane π^*(B–B) antibonding transitions.

Fig. 6 presents the C 1s experimental polarized spectra and their comparison with the calculated ones for the fixed in space B_2Cat_2 molecule. The p-pol component dominates in all the energy range below the threshold and the theory correctly describes the strong dichroic effects associated with the π out of plane nature of most final MOs. In particular, the first two peaks correspond to the A and B peaks of the gas-phase spectrum (Fig. 3) associated to π^*(C–C) antibonding transitions.

Around 288 eV the s-pol component starts to acquire some intensity with a feature corresponding to peak C of the gas phase spectrum due to the transitions towards diffuse MOs with in plane atomic components. Transitions of similar nature contribute to the experimental s-pol broad signal present in the energy range of 289–292 eV. The p-pol component contributes to this energy range with two broad features whose theoretical counterparts appear still overestimated and slightly shifted at a higher energy and correspond to the E and F peaks of the gas phase spectrum. This analysis of the p- and s-pol components of the C K-edge spectrum confirms the attribution of the spectral features proposed for the gas-phase experimental spectrum despite some discrepancies provided by the theory in the energy range around C 1s thresholds.

The results for the O K-edge p-pol and s-pol spectra are shown in Fig. 7. The dichroic effects are still relevant at a lower

Fig. 5 B K-edge NEXAFS spectra of B_2Cat_2@Au(111) at two different polarization angles. The ΔSCF calculated spectra of the fixed in space B_2Cat_2 have been shifted by 0.15 eV in order to match the first experimental peak. The ΔSCF B 1s IP is indicated with a vertical dashed line. The stick spectra are broadened by using a Gaussian line shape with FWHM = 0.80 eV.

Fig. 6 C K-edge NEXAFS spectra of B_2Cat_2@Au(111) at two different polarization angles. The DFT-TP calculated spectra of the fixed in space B_2Cat_2 have been reported without applying any energy shift. ΔSCF C 1s IPs are indicated with vertical dashed lines. The stick spectra are broadened by using a Gaussian line shape with FWHM = 0.75 eV.
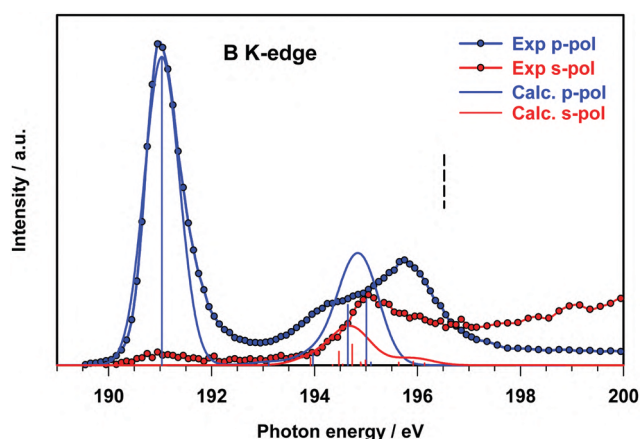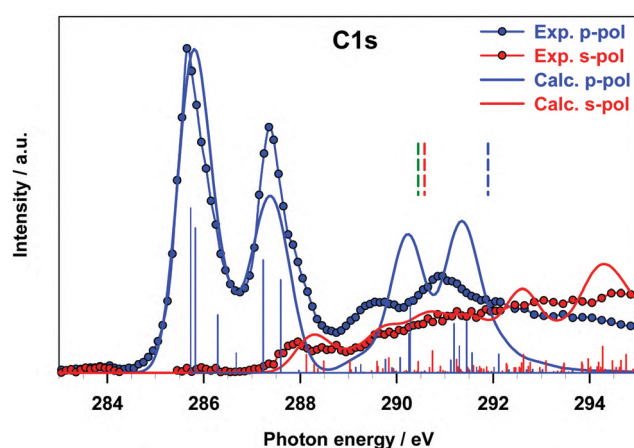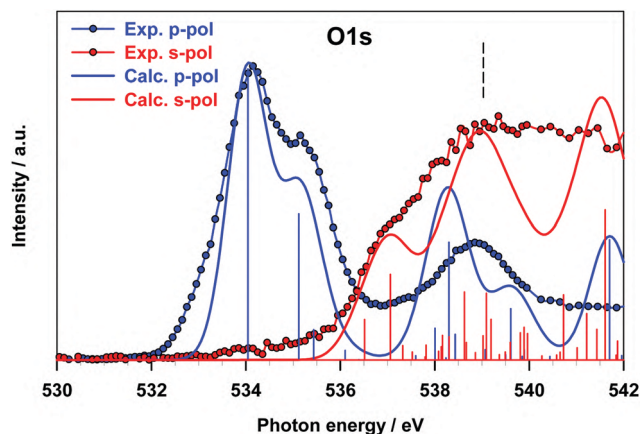
Fig. 7 O K-edge NEXAFS spectra of B$_2$Cat$_2$@Au(111) at two different polarization angles. The DFT-TP calculated spectra of the fixed in space B$_2$Cat$_2$ have been shifted by −0.20 eV in order to match the first experimental peak. ΔSCF O 1s IPs are indicated with vertical dashed lines. The stick spectra are broadened by using a Gaussian line shape with FWHM = 1.0 eV.

energy, while both components provide significant contributions to the features above 536 eV. The agreement between the experiment and theory is good for both components, and the strong dichroic effects present in the lower energy region are consistent with the π nature of the transitions relative to the A and B peaks of the gas phase spectrum.

The intensity of the s-pol component mainly arises from transitions of diffuse nature (peak C of the gas phase spectrum), while the intensity distribution between the s- and p-pol components near the threshold reflects the character of the transitions contributing to peak D of the gas phase spectrum (see Table 3). The good agreement found between theoretical calculations on the oriented molecule and experimental results for B$_2$Cat$_2$ adsorbed on Au(111) proves that the Au surface does not perturb the transitions of any of the B 1s, C 1s and O 1s channels confirming the weak interaction of the molecule with the surface.

Conclusions

DFT calculations have been used to assign the features observed in the B, C and O K-shell XPS and NEXAFS spectra of the B$_2$Cat$_2$ molecule, with the aim of obtaining detailed information on the B$_2$Cat$_2$ electronic structure. The assignment of the absorption peaks observed in the gas-phase NEXAFS spectra has been validated by a study of the corresponding NEXAFS spectra for the molecule adsorbed on the poorly reactive Au(111) surface, which allowed confirmation of the nature of the virtual electronic states involved in the core-excitation process. This represents a valuable reference to explore the surface chemistry of diboron compounds with other molecular species or with reactive metallic substrates. In particular the energy position and intensity of the B 1s core to the π(B–B) bonding and anti-bonding orbitals represent a sensitive probe of the chemistry of the molecule. In this regard, for

example, we outline that, in the B 1s edge spectra, a quite different position of the first transition (~191 eV) is found with respect to the case of the covalent framework of boroxine groups based on the B–B bond (~194 eV).[15] We suggest therefore that in the latter case a stronger interaction with the substrate was established and that the B–B centers were involved in a charge transfer process with a consequent change in their chemical environment.

Finally, our results point out that the ΔSCF method is required to describe the B 1s core excited states in order to reach a good agreement with the experimental data, with respect to both DFT-TP and linear response TDDFT approaches. This indicates the importance of explicitly treating the relaxation changes among different excited-state configurations relative to the B 1s core hole, unlike the C and O K-edges which can be conveniently described at the DFT-TP level with the relaxation effects included in an average way. This behavior can be traced to the less pronounced relaxation following the formation of the shallower B 1s core hole compared to the heavier first row atoms, such as C and O.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. L. Korich and P. M. Iovine, Boroxine chemistry and applications: A perspective, *Dalton Trans.*, 2010, **39**, 1423–1431.

2 E. Dimitrijević and M. S. Taylor, Organoboron Acids and Their Derivatives as Catalysts for Organic Synthesis, *ACS Catal.*, 2013, **3**, 945–962.

3 G. Duret, R. Quinlan, P. Bisseret, N. Blanchard, A. R. Martin, J.-J. Vasseur and M. Smietana, Boron chemistry in a new light, *Chem. Sci.*, 2013, **42**, 5684–5713.

4 E. C. Neeve, S. J. Geier, I. A. I. Mkhalid, S. A. Westcott and T. B. Marder, Diboron(4) Compounds: From Structural Curiosity to Synthetic Workhorse, *Chem. Rev.*, 2016, **116**, 9091–9161.

5 J. Wang, W. Zheng and Y. Zheng, Theoretical study on homolytic B-B cleavages of diboron(4) compounds, *RSC Adv.*, 2017, 7, 49251–49272.

6 M. Eck, S. Würtemberger-Pietsch, A. Eichhorn, J. H. J. Berthel, R. Bertermann, U. S. D. Paul, H. Schneider, A. Friedrich, C. Kleeberg, U. Radius and T. B. Marder, B-B bond activation and NHC ring-expansion reactions of diboron(4) compounds, and accurate molecular structures of B2(NMe2)4, B2eg2, B2neop2 and B2pin2, *Dalton Trans.*, 2017, **46**, 3661–3680.

23524 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 23517–23525

This journal is © the Owner Societies 2021

7 A. F. Eichhorn, L. Kuehn, T. B. Marder and U. Radius, Facile insertion of a cyclic alkyl(amino) carbene carbon into the B-B bond of diboron(4) reagents, *Chem. Commun.*, 2017, **53**, 11694–11696.

8 Y. Katsuma, H. Asakawa and M. Yamashita, Reactivity of highly Lewis acidic diborane(4) towards pyridine and isocyanide: Formation of boraalkene-pyridine complex and: Ortho -functionalized pyridine derivatives, *Chem. Sci.*, 2018, **9**, 1301–1310.

9 A. F. Pécharman, N. A. Rajabi, M. S. Hill, C. L. McMullin and M. F. Mahon, Diborane heterolysis and P(v) reduction by Ph₃P=O coordination to magnesium, *Chem. Commun.*, 2019, **55**, 9035–9038.

10 S. Pietsch, U. Paul, I. A. Cade, M. J. Ingleson, U. Radius and T. B. Marder, Room Temperature Ring Expansion of N-Heterocyclic Carbenes and B-B Bond Cleavage of Diboron(4) Compounds, *Chem. – Eur. J.*, 2015, **21**, 9018–9021.

11 T. Ishiyama, K. Ishida and N. Miyaura, Synthesis of pinacol arylboronates via cross-coupling reaction of bis(pinacolato)-diboron with chloroarenes catalyzed by palladium(0)-tricyclohexylphosphine complexes, *Tetrahedron*, 2001, **57**, 9813–9816.

12 A. Cossaro, M. Puppin, D. Cvetko, G. Kladnik, A. Verdini, M. Coreno, M. De Simone, L. Floreano and A. Morgante, Tailoring SAM-on-SAM formation, *J. Phys. Chem. Lett.*, 2011, **2**, 3124–3129.

13 Z. Feng, G. Kladnik, G. Comelli, C. Dri and A. Cossaro, Growth of regular nanometric molecular arrays on a functional 2D template based on a chemical guest–host approach, *Nanoscale*, 2018, **10**, 2067–2072.

14 J. V. Barth, Molecular architectonic on metal surfaces, *Annu. Rev. Phys. Chem.*, 2007, **58**, 375–407.

15 M. Stredansky, A. Sala, T. Fontanot, R. Costantini, C. Africh, G. Comelli, L. Floreano, A. Morgante and A. Cossaro, On-surface synthesis of a 2D boroxine framework: a route to a novel 2D material?, *Chem. Commun.*, 2018, **54**, 3971–3973.

16 D. Toffoli, A. Ponzi, E. Bernes, M. De Simone, C. Grazioli, M. Coreno, M. Stredansky, A. Cossaro and G. Fronzoni, Correlation effects in B 1s core-excited states of boronic-acid derivatives: An experimental and computational study, *J. Chem. Phys.*, 2019, **151**, 134306.

17 J. C. Slater, in *Advances in Quantum Chemistry*, ed. P.-O. Löwdin, Academic Press, 1972, vol. 6, pp. 1–92.

18 L. Triguero, L. G. M. Pettersson and H. Ågren, Calculations of near-edge X-ray-absorption spectra of gas-phase and chemisorbed molecules by means of density-functional and transition-potential theory, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 8097–8110.

19 W. Parr and R. G. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press, New York, 1989.

20 C. F. Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, Towards an order-N DFT method, *Theor. Chem. Acc.*, 1998, **99**, 391–403.

21 W. Clegg, M. R. J. Elsegood, F. J. Lawlor, N. C. Norman, N. L. Pickett, E. G. Robins, A. J. Scott, P. Nguyen, N. J. Taylor and T. B. Marder, Structural Studies of Bis-Catecholate, Bis-Dithiocatecholate, and Tetraalkoxy Diborane(4) Compounds, *Inorg. Chem.*, 1998, **37**, 5289–5293.

22 J. P. Perdew, Density-functional approximation for the correlation energy of the inhomogeneous electron gas, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **33**, 8822–8824.

23 J. P. Perdew, Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **34**, 7406.

24 E. J. Baerends, D. E. Ellis and P. Ros, Self-consistent molecular Hartree–Fock–Slater calculations I. The computational procedure, *Chem. Phys.*, 1973, **2**, 41–51.

25 P. Norman and A. Dreuw, Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules, *Chem. Rev.*, 2018, **118**, 7208–7248.

26 A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.*, 1993, **98**, 5648.

27 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

28 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

29 T. Ziegler, A. Rauk and E. J. Baerends, On the calculation of multiplet energies by the Hartree-Fock-slater method, *Theor. Chim. Acta*, 1977, **43**, 261–271.

30 P.-O. Löwdin, Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction, *Phys. Rev.*, 1955, **97**, 1474–1489.

31 S. Bourne Worster, O. Feighan and F. R. Manby, Reliable transition properties from excited-state mean-field calculations, *J. Chem. Phys.*, 2021, **154**, 124106.

32 K. C. Prince, R. R. Blyth, R. Delaunay, M. Zitnik, J. Krempasky, J. Slezak, R. Camilloni, L. Avaldi, M. Coreno, G. Stefani, C. Furlani, M. De Simone and S. Stranges, The gas-phase photoemission beamline at Elettra, *J. Synchrotron Radiat.*, 1998, **5**, 565–568.

33 R. Costantini, M. Stredansky, D. Cvetko, G. Kladnik, A. Verdini, P. Sigalotti, F. Cilento, F. Salvador, A. De Luisa, D. Benedetti, L. Floreano, A. Morgante, A. Cossaro and M. Dell'Angela, ANCHOR-SUNDYN: A novel endstation for time resolved spectroscopy at the ALOISA beamline, *J. Electron Spectrosc. Relat. Phenom.*, 2018, **229**, 7–12.

34 I. Ljubić, M. T. Cvitaš, C. Grazioli, M. Coreno, S. Kazazić and I. Novak, Vibrationally resolved valence and core photoionization and photoexcitation spectra of an electron-deficient trivalent boron compound: The case of catecholborane, *Phys. Chem. Chem. Phys.*, 2020, **22**, 25396–25407.

35 D. Toffoli, M. Stredansky, Z. Feng, G. Balducci, S. Furlan, M. Stener, H. Ustunel, D. Cvetko, G. Kladnik, A. Morgante, A. Verdini, C. Dri, G. Comelli, G. Fronzoni and A. Cossaro, Electronic properties of the boroxine–gold interface: evidence of ultra-fast charge delocalization, *Chem. Sci.*, 2017, **8**, 3789–3798.

Revealing the electronic properties of the B-B bond:

the bis-catecholato diboron molecule

D. Toffoli[1], C. Grazioli[2], M. Monti[1], M. Stener[1], R. Totani[3], R.Richter[4], L. Schio[2], G. Fronzoni[1,a)], A. Cossaro[1,2,a)]

AFFILIATIONS

[1] *Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy*
[2] *IOM-CNR, Istituto Officina dei Materiali-CNR, S.S.14, Km 163.5, 34149 Trieste, Italy*
[3] *ISM-CNR, Istituto di Struttura della Materia, LD2 Unit, 34149 Trieste, Italy*
[4] *Elettra-Sincrotrone Trieste, 34149 Basovizza, Trieste, Italy*

a) Authors to whom correspondence should be addressed: fronzoni@units.it; acosssaro@units.it

Electronic Supplementary Information

1. Geometry optimization of B_2Cat_2
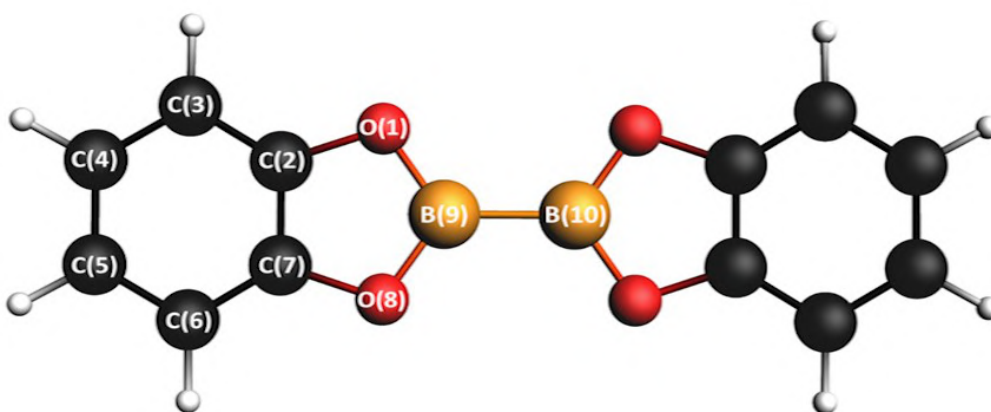
Figure S1. *Chemical structure of B_2Cat_2. Labels on the atoms refer to geometrical parameters listed in Table S1.*

Table S1. *Comparison between theoretical and experimental geometrical parameters for B2Cat2. Bond lengths are expressed in Å and bond angles in degrees (°).*

Bond lengths[a]	Calculated	Experimental[b]
B(9)-B(10)	1.685	1.678
O(8)-C(7)	1.382	1.387
O(1)-C(2)	1.382	1.386
C(7)-C(6)	1.381	1.372
C(6)-C(5)	1.398	1.391
C(5)-C(4)	1.398	1.393
C(4)-C(3)	1.398	1.393
C(3)-C(2)	1.381	1.376
C(2)-C(7)	1.390	1.384
O(8)-B(9)	1.391	1.382
O(1)-B(9)	1.390	1.394
Bond angles[a]	Calculated	Experimental[b]
B(10)-B(9)-O(8)	124.5	124.6
B(10)-B(9)-O(1)	124.5	124.2
B(9)-O(8)-C(7)	105.5	105.5
B(9)-O(1)-C(2)	105.5	105.1
O(8)-B(9)-O(1)	111.0	111.2
O(8)-C(7)-C(6)	128.9	128.7
O(1)-C(2)-C(3)	128.9	128.6
O(8)-C(7)-C(2)	109.0	108.9
O(1)-C(2)-C(7)	109.0	109.2
C(7)-C(2)-C(3)	122.1	122.1
C(2)-C(7)-C(6)	122.1	122.4

C(7)-C(6)-C(5)	116.5	116.2
C(6)-C(5)-C(4)	121.4	121.6
C(5)-C(4)-C(3)	121.4	121.5
C(4)-C(3)-C(2)	116.5	116.2

[a] Labels of the atoms refer to those reported in Figure S1.
[b] W. Clegg et Al. *Inorg. Chem.* **1994**, *33*, 4623-4624.

2. Experimental and theoretical XPS Spectra
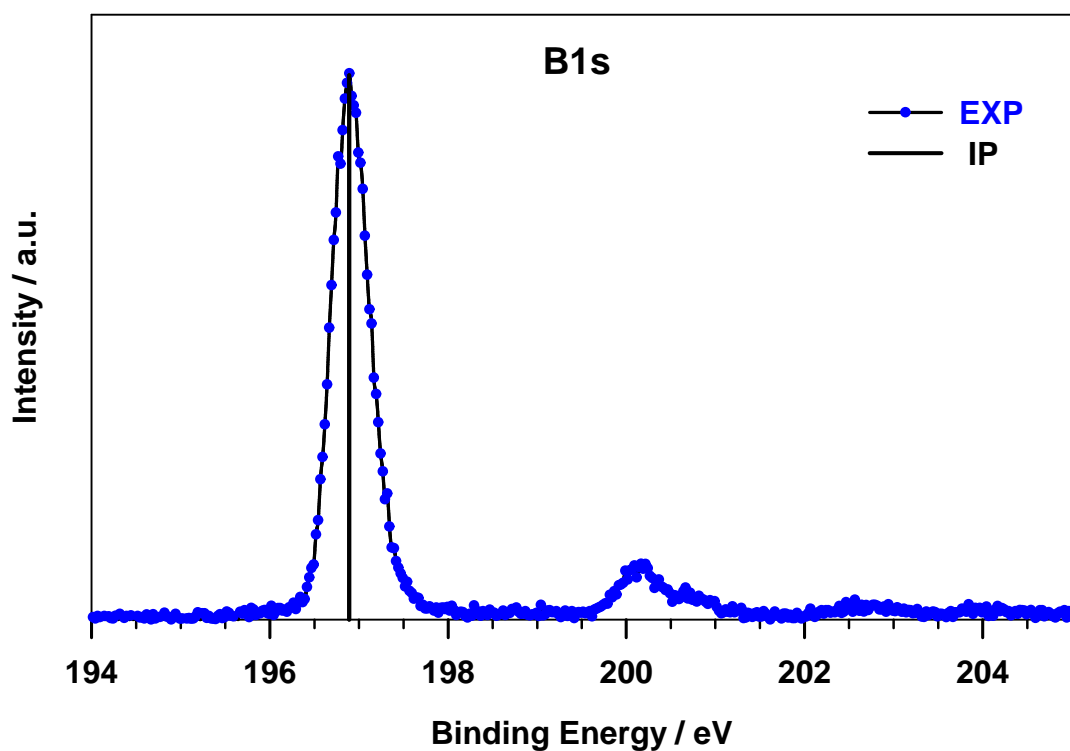
Figure S2. B1s photoelectron spectrum of B$_2$Cat$_2$. Vertical black line: B1s calculated ΔSCF IP (196.52 eV) shifted by +0.37 eV in order to match the experimental peak.
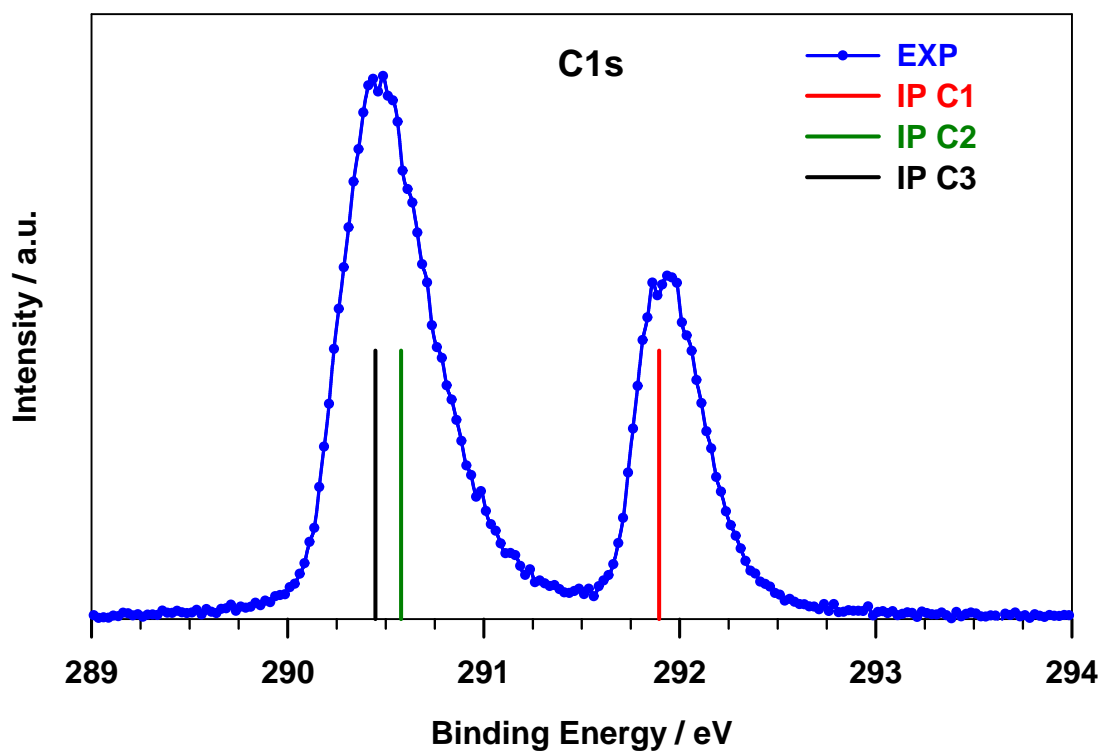
Figure S3. C1s photoelectron spectrum of B$_2$Cat$_2$. Vertical colored lines:C1s calculated ΔSCF IPs (C1 : 291.89 eV; C2: 290.58 eV; C3: 290.45 eV)
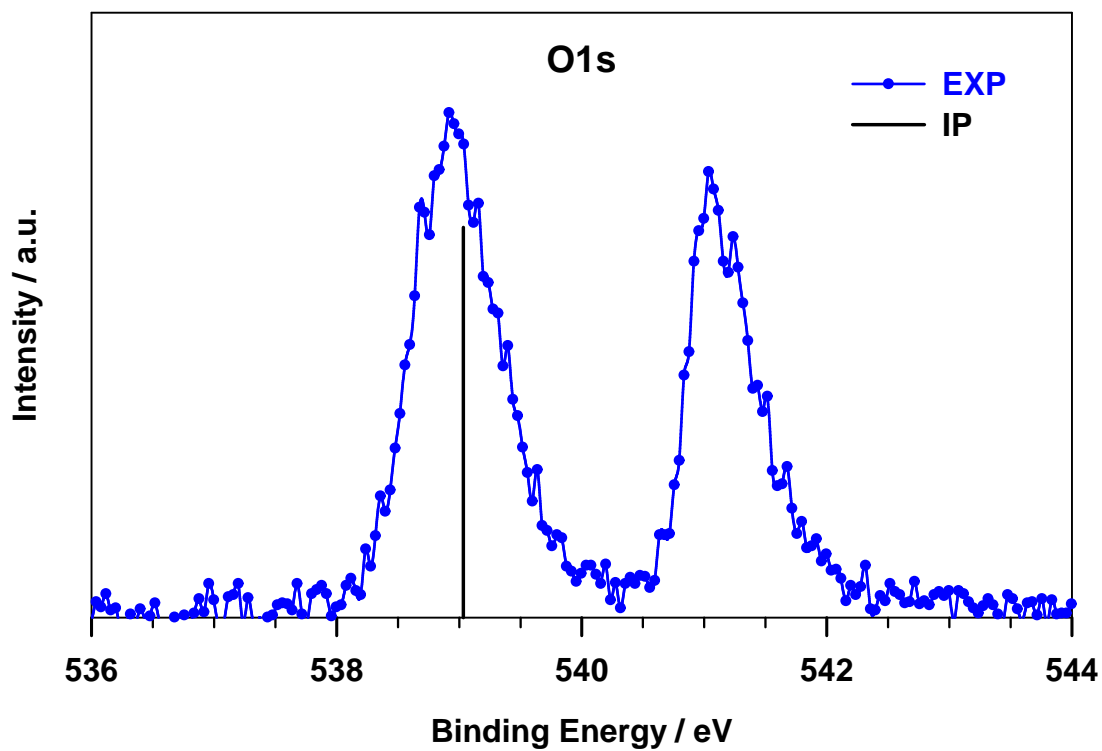
Figure S4. O1s photoelectron spectrum of B$_2$Cat$_2$. Vertical black line: O1s calculated ΔSCF IP (539.03 eV).

3. DFT-TP B K-Edge NEXAFS spectrum with different xc functionals
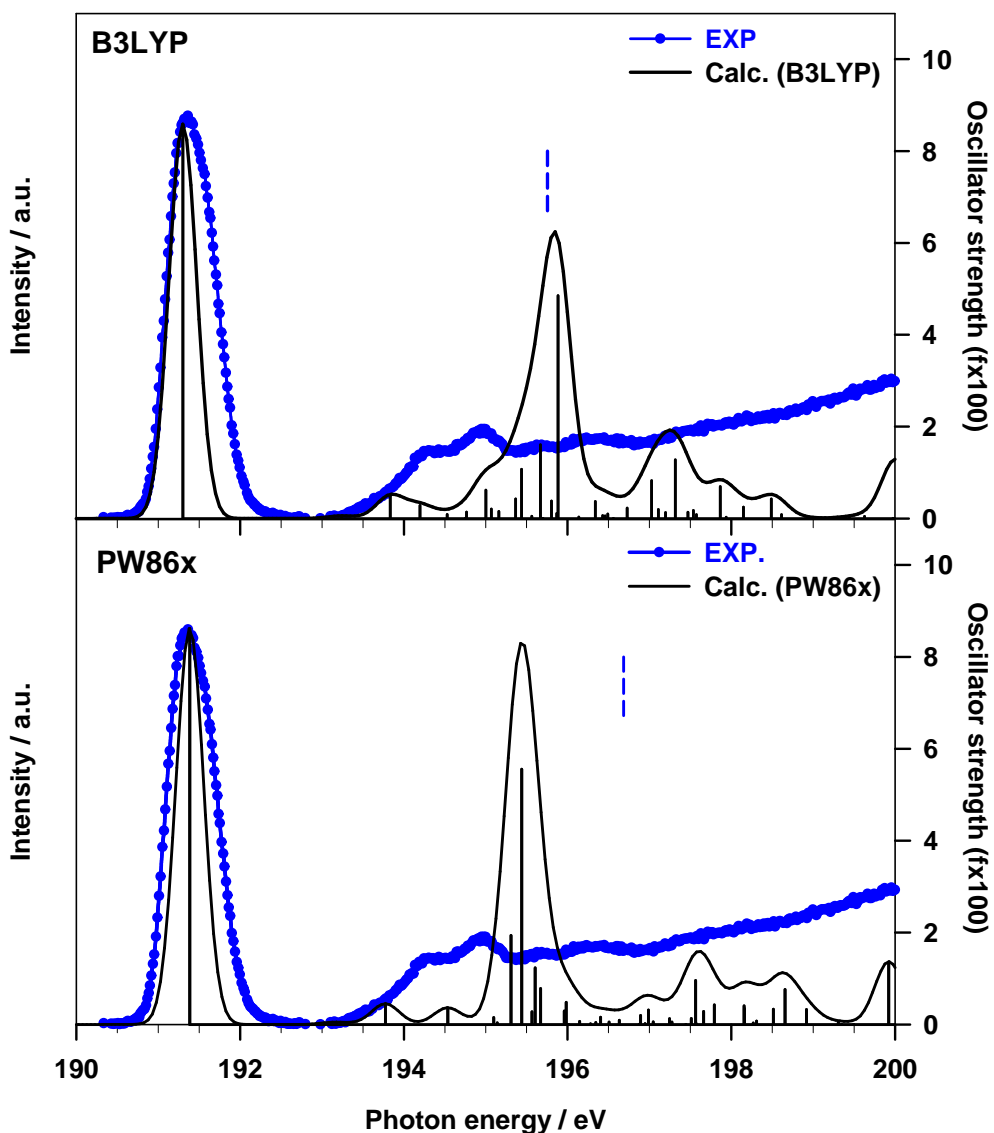
Figure S5. DFT-TP NEXAFS B1s spectra of B_2Cat_2. Upper panel: B3LYP; lower panel PW86. The experimental spectrum is reported for comparison The calculated spectra were rigidly shifted on the experimental energy scale (B3LYP: 0.76 eV; PW86 +0.17 eV). Vertical blue lines: IP experimental values. The stick spectra are broadened by using a Gaussian lineshape with FWHM=0.4 eV.

4. Details on the TDDFT calculations and Results

The B1s core excitations of B_2Cat_2 have also been calculated by employing the Time-Dependent DFT (TDDFT) method, which formally includes the coupling between single excited configurations. The core excited states can be efficiently computed by reducing the 1h-1p space to include only the configurations generated by the excitations from the subset of core orbitals[1]. In this way the coupling between core excitations and valence excitations to the continuum is neglected and the computational scheme does not need the treatment of the continuum states. The TDDFT scheme employed for the calculations of the B K-edge spectrum consists of the excitations from both the core orbitals of the two B atoms (coupled scheme).

In the TDDFT calculations, the adiabatic local density approximation (ALDA) to the exchange-correlation (xc) kernel has been employed[2]. For the xc potential applied in the self-consistent field (SCF) calculations, three different xc functionals with the ground state electron configuration have been considered: the LB94[3], chosen for its correct asymptotic behaviour, the hybrid B3LYP, and the range-separated hybrid CAM-B3LYP xc functionals. The TDDFT calculations have been performed by using a TZP basis set for all atoms.

The calculated B1s NEXAFS spectra are collected in Figure S6 and compared with the gas-phase experiment. The influence of the different xc functionals is also tested. The inspection of the figure reveals that the disagreement between theory and experiment persists also at TDDFT level, irrespective of the xc functional employed, in particular the incorrect intensity distribution between the first two peaks as well as the overestimate of their energy separation. Bearing in mind the DFT-TP results, no improvement is therefore obtained at the TDDFT level, suggesting that the configuration mixing introduced at the TDDFT level (within the ALDA approximation) is not adequate to properly describe the B1s core excitations, as already found in previous calculations on boronic-acid derivatives[4] . It is interesting to observe that the nature of the excitations at the TDDFT level is consistent with that obtained at the DFT-TP level, confirming the assignment of the experimental spectrum.
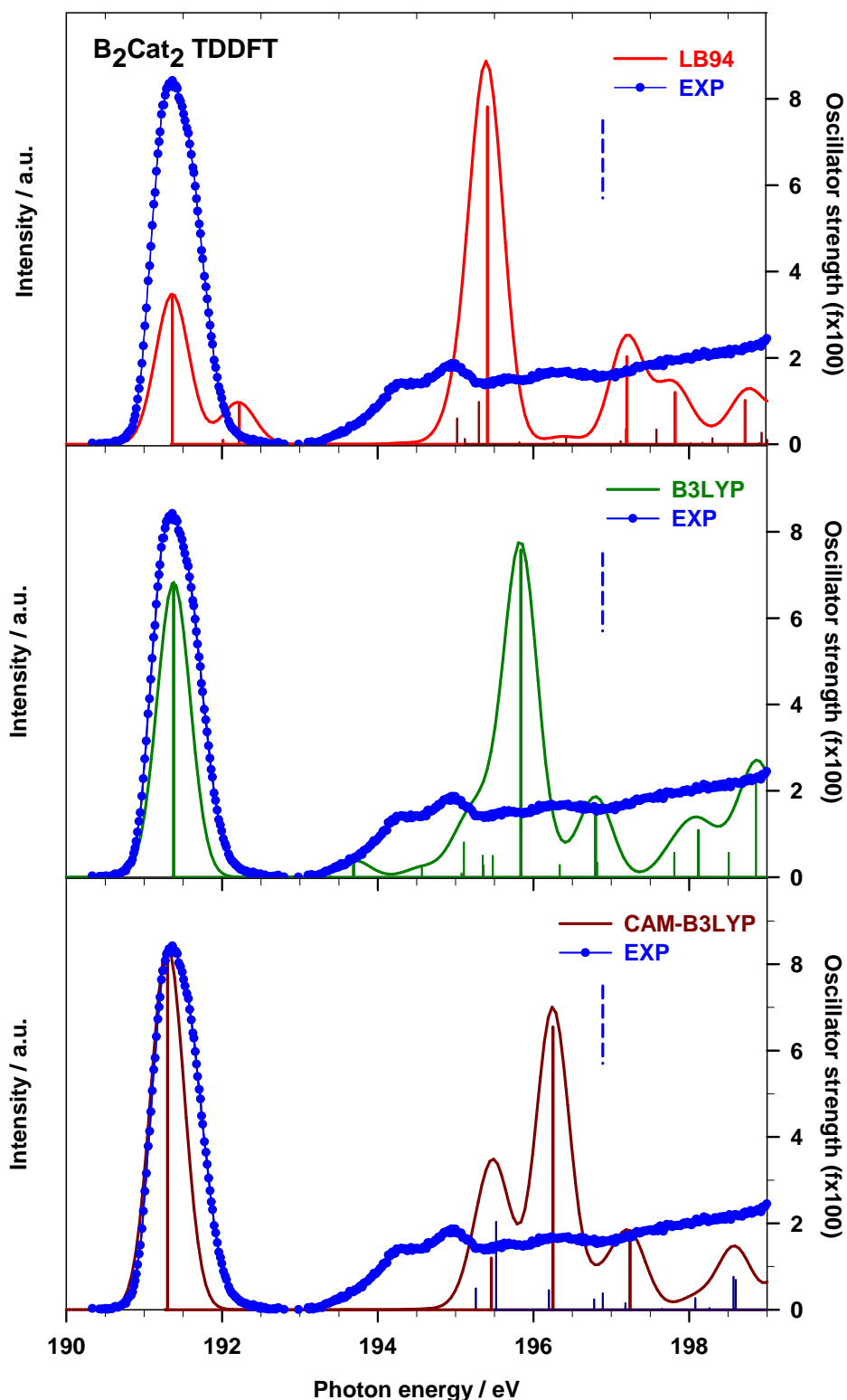
Figure S6. B1s NEXAFS spectra of B_2Cat_2 calculated at the TDDFT level by employing different xc potentials. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated B1s profiles to align the energy position of the first absorption peak to the experimental one (2.44 eV for LB94; 9.81 eV for B3LYP; 9.64 eV for CAM-B3LYP). Vertical blue lines: IP (B1s) experimental value. The stick spectra are broadened by using a Gaussian lineshape with FWHM=0.5 eV.

5. 3D plots of selected final MOs corresponding to the B1s, C1s and O1s core excitation calculations
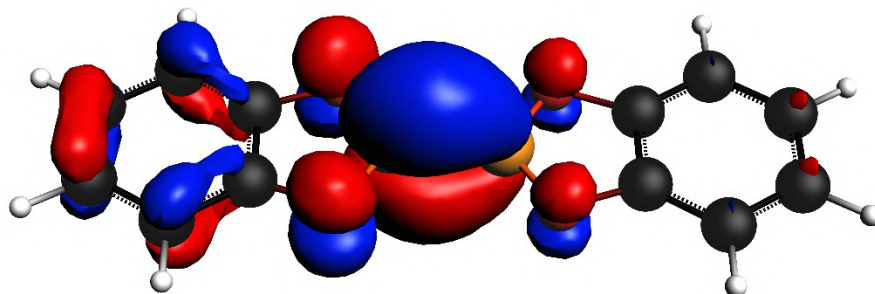
7 b1 (ΔSCF transition at 191.33 eV)
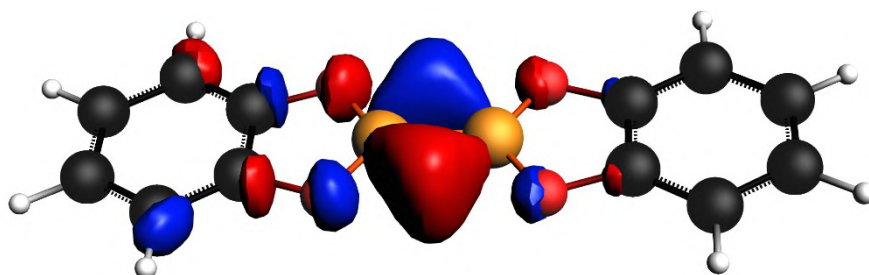
23 b$_2$ (ΔSCF transition at 194.78 eV)
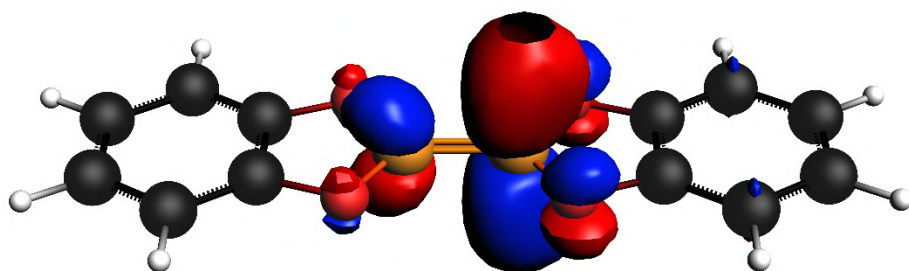
11 b$_1$ (ΔSCF transition at 195.31 eV)

Figure S7. 3D plots of final MOs (with a core-hole on the B site of B$_2$Cat$_2$) obtained from the ΔSCF calculations . Displayed isosurface corresponds to ±0.030 e$^{1/2}$a$_0^{-3/2}$ value.
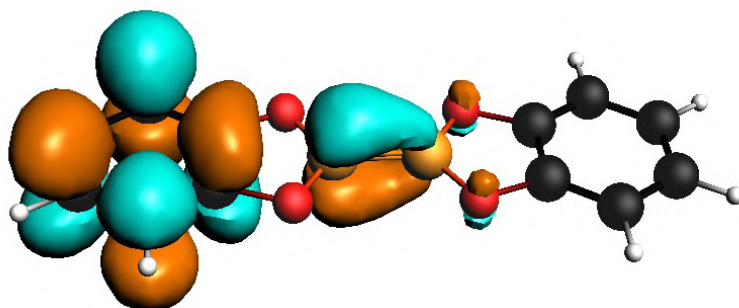
Figure S8. 3D plot of the LUMO MO of B_2Cat_2 (with a core-hole on C2) from the DFT-TP calculations. Displayed isosurface corresponds to $\pm 0.030 \ e^{1/2}a_0^{-3/2}$ value.
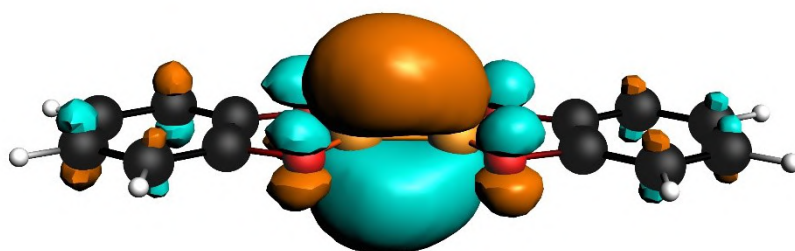
LUMO (transition at 534.34 eV)
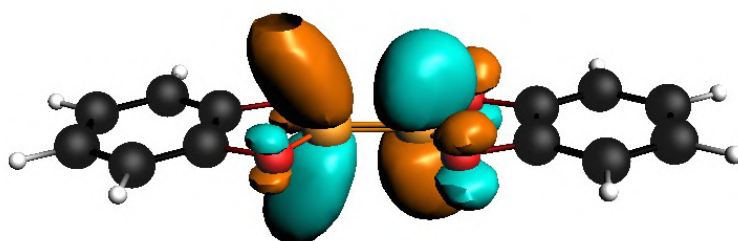
LUMO + 17 (transition at 538.60 eV)

Figure S9. 3D plots of LUMO and LUMO+17 MOs of B_2Cat_2 (with a core-hole on the O site) from the DFT-TP calculations. Displayed isosurface corresponds to $\pm 0.030\ e^{1/2}a_0^{-3/2}$ value.

6. Experimental details

The BCDB molecule, in the form of a powder, was purchased from TCI, purity >98%, and sublimated in vacuum by means of a Knudsen cell heated at 330 K. The molecule was studied in the isolated form at the Gas Phase beamline[5] at Elettra. The sublimation rate was optimized checking the mass spectrum by means of a home-made time-of-flight spectrometer (TOF). The powder was purified for one night and the spectra collection started after no trace of contamination in the photoemission signal was detected. The high-resolution N 1s, C 1s, and B 1s XPS and valence band spectra were recorded with a VG-220i hemispherical electron energy analyzer mounted in the plane perpendicular to the photon propagation direction at the magic angle (i.e. 54.7° with respect to the electric vector

of the light). In this geometry, photoemission measurements are insensitive to the photoelectron asymmetry β parameter. NEXAFS spectra at the N, C and B *K* edges were collected using the same TOF described before in a total-ion-yield configuration. The photon energies and resolutions used for the core excitation spectra are reported in Table S2. More details about the end station are available in Ref.[6]

Table S2. *Gas phase experimental parameters*

PES				
Spectral region	Photon en. [eV]	Overall resolution [meV]	Calibrant	Binding en. [eV]
B1s	262	130	SF6 (2p3/2)	180.21 eV [7,8]
C1s	390	150	CO2 (C1s)	297.7 eV[9]
O1s	628	240	CO2 (O1s)	541.08 eV [7]
VB	97	100	N2 (p_{BOND})	15.60 [10]
VB (ResPES)	189.3 191.3	360		
NEXAFS				
B K	190-210	20	SF6 ($2p_{1/2} \rightarrow t_{2g}$)	184.57[11]
C K	283-300	50	CO2 ($C1s \rightarrow \pi$)	290.77 [12]
O K	532-540	90	CO2 (O $1s \rightarrow \pi_u$)	535.4 [13]

The presented data refer to the saturation coverage of a monolayer. B_2Cat_2 was deposited on clean Au(111) single crystal surface by sublimation from a Knudsen cell heated at 330 K, with the substrate kept at room temperature. In these conditions, molecules beyond the first layer do not stick, that was confirmed spectroscopically by monitoring the intensity of the XPS peaks. Measurements were performed at the ANCHOR-SUNDYN endstation[14] of the ALOISA beamline at Elettra Synchrotron. XPS measurements were performed at photon energies of 650 eV for O1s and at 400 eV for C1s and B1s, with an overall resolution of 0.2 and 0.15 eV respectively. NEXAFS spectra were acquired in Auger yield, with photon resolution of 0.15 and 0.1 eV for O K-edge and B (C) K-edge respectively. In order to investigate dichroic effects the NEXAFS spectra were acquired at two different geometries: with the sample surface oriented either parallel to the light electric field plane (s-polarization) or closely normal to it (p-polarization).

References

1   M. Stener, G. Fronzoni and M. de Simone, Time dependent density functional theory of core electrons excitations, *Chem. Phys. Lett.*, 2003, **373**, 115–123.

2   E. K. U. Gross and W. Kohn, in *Density Functional Theory of Many-Fermion Systems*, ed. P.-O. Löwdin, Academic Press, 1990, vol. 21, pp. 255–291.

3   R. van Leeuwen and E. J. Baerends, Exchange-correlation potential with correct asymptotic behavior, *Phys. Rev. A*, 1994, **49**, 2421–2431.

4   D. Toffoli, A. Ponzi, E. Bernes, M. De Simone, C. Grazioli, M. Coreno, M. Stredansky, A. Cossaro and G. Fronzoni, Correlation effects in B1s core-excited states of boronic-acid derivatives: An experimental and computational study, *J. Chem. Phys.*, , DOI:10.1063/1.5120175.

5   K. C. Prince, R. R. Blyth, R. Delaunay, M. Zitnik, J. Krempasky, J. Slezak, R. Camilloni, L. Avaldi, M. Coreno, G. Stefani, C. Furlani, M. De Simone and S. Stranges, The gas-phase photoemission beamline at Elettra., *J. Synchrotron Radiat.*, 1998, **5**, 565–8.

6   O. Plekan, H. Sa'adeh, A. Ciavardini, C. Callegari, G. Cautero, C. Dri, M. Di Fraia, K. C. Prince, R. Richter, R. Sergo, L. Stebel, M. Devetta, D. Faccialà, C. Vozzi, L. Avaldi, P. Bolognesi, M. C. Castrovilli, D. Catone, M. Coreno, F. Zuccaro, E. Bernes, G. Fronzoni, D. Toffoli and A. Ponzi, Experimental and Theoretical Photoemission Study of Indole and Its Derivatives in the Gas Phase, *J. Phys. Chem. A*, 2020, **124**, 4115–4127.

7   L. Pettersson, J. Nordgren, L. Selander, C. Nordling, K. Siegbahn and H. Ågren, Core-electron binding energies in the soft X-ray range obtained in X-ray emission, *J. Electron Spectros. Relat. Phenomena*, , DOI:10.1016/0368-2048(82)85050-0.

8   J. E. True, T. D. Thomas, R. W. Winter and G. L. Gard, Electronegativities from core-ionization energies: Electronegativities of SF5 and CF3, *Inorg. Chem.*, , DOI:10.1021/ic0343298.

9   V. Myrseth, J. . Bozek, E. Kukk, L. . Sæthre and T. . Thomas, Adiabatic and vertical carbon 1s ionization energies in representative small molecules, *J. Electron*

Spectros. Relat. Phenomena, 2002, **122**, 57–63.

10   H.-H. Perkampus, D. W. Turner: Molecular Photoelectron Spectroscopy. Unter Mitarbeit von C. Baker, A. D. Becker und C. R. Brundle, Wiley-Interscience, John Wiley & Sons Ltd., London, New York, Sydney, Toronto 1970. 386 Seiten. Preis: 140 sh., *Berichte der Bunsengesellschaft für Phys. Chemie*, 1971, **75**, 181–181.

11   E. Hudson, D. A. Shirley, M. Domke, G. Remmers, A. Puschmann, T. Mandel, C. Xue and G. Kaindl, High-resolution measurements of near-edge resonances in the core-level photoionization spectra of SF6, *Phys. Rev. A*, 1993, **47**, 361–373.

12   M. Tronc, G. C. King and F. H. Read, Carbon K-shell excitation in small molecules by high-resolution electron impact, *J. Phys. B At. Mol. Phys.*, 1979, **12**, 137–157.

13   G. R. Wight and C. E. Brion, K-Shell energy loss spectra of 2.5 keV electrons in CO2 and N2O, *J. Electron Spectros. Relat. Phenomena*, 1974, **3**, 191–205.

14   R. Costantini, M. Stredansky, D. Cvetko, G. Kladnik, A. Verdini, P. Sigalotti, F. Cilento, F. Salvador, A. De Luisa, D. Benedetti, L. Floreano, A. Morgante, A. Cossaro and M. Dell'Angela, ANCHOR-SUNDYN: A novel endstation for time resolved spectroscopy at the ALOISA beamline, *J. Electron Spectros. Relat. Phenomena*, 2018, **229**, 7–12.

Chapter 5

Conclusions

In conclusion, the main goal of this PhD project was to implement a computational procedure for accurate and economic calculations of the ECD spectra of complex systems. Such procedure was developed combining classical molecular dynamics simulations, with the essential dynamics analysis, and state-of-the-art time-dependent density functional theory calculations. More in detail, a trajectory of the system of interest (*e.g.*, solute, or solute-solvent clusters) is obtained first by running a classical MD simulation. Such trajectory is analyzed by means of the ED analysis, which allows us to reduce the complexity of the conformational space that we want to investigate. By applying this strategy, one can build a 2D free-energy landscape from which the most probable conformers are extracted. Such structures are then treated as representative of the whole conformational ensemble for the further quantum-chemical calculations. First, the selected conformers are optimized at the DFT level constraining the torsional angles. Afterwards, the ECD spectra are calculated by using TDDFT. Here, different algorithms can be employed according to the size and nature of the system under investigation. All the spectra are then statistically weighted and summed up to give a final averaged ECD, which can be compared with the experimental reference for a qualitative evaluation. Similarities between experimental and calculated spectra confirm the quality of the approach, allowing us to look into the role of conformational, solvent, as well as thermal effects on the chiroptical features.

The protocol was validated on two biomolecules, namely, the cationic GAG^+ tripeptide, and the neutral $(GVGVP)_2$ decapeptide. Both the systems are soluble in water, and their experimental ECD are available at different temperature values. Working at 30 °C and 90 °C for the tripeptide, while at 15 °C and 85 °C for the decapeptide, we were able to evaluate if the approach is able to properly capture all the three effects mentioned above. By comparing the experimental and calculated optical responses, a good qualitative agreement was observed in all the cases. Therefore, the computational procedure confirmed to be both reliable and affordable. It was found out that both the systems exist in two possible conformations (pPII and β-strand for GAG^+, and random coil and β-turn for $(GVGVP)_2$), whose equilibrium is influenced by the temperature of work. Moreover, a detailed investigation on the water role revealed that not only it influences the conformational equilibrium, through inter-molecular interactions, but it also perturbs the electronic states of the cromophore. Therefore, in order to

reproduce the correct experimental features, it is essential to treat the water molecules explicitly.

The validation work was followed by two projects focused on more complex systems, *i.e.* thiolate-protected gold nanoclusters. In the first application, the chiral $Au_{38}(PET)_{24}$, where $PET=SC_2H_4Ph$, in toluene was investigated. Here, 12 statistically relevant conformations were extracted and later employed to calculate the ECD response. In this case, the organic solvent was explicitly treated only in the MD simulation because of its limited influence on the optical features of the cromophore. The averaged ECD spectrum was compared with the available experimental one, showing a good match in terms of both energy and intensity distribution in the PET energy region. Some discrepancies were found instead in the optical region of the metallic core, probably due to the choice of the xc functional. It is known that the GGA LB94 functional, selected here because of the computational cost, does not reproduce the spectral features of metals properly. Such hypothesis was corroborated by calculating the ECD on the resolved X-ray structure with the hybrid B3LYP functional. However, this calculation, together with that on the same X-ray structure and the LB94 functional, showed that the ligands features are not well reproduced anymore considering only one conformation. This proved the need of including, at least, the most relevant conformers to capture the flexibility and dynamics of the ligands which profoundly affect the chiroptical response. To the best of our knowledge, this is the first work where the conformational effects are explicitly treated to calculate the ECD response of a thiolate-protected gold nanocluster.

In the second work on nanoclusters, we reproduced the ECD spectrum of $[Au_{25}(GSH)_{18}]^{-1}$ (GSH=glutathione) in water. Here, the most plausible solute-(solvent)$_N$ clusters were also extracted. Repeating the whole procedure, a final set of 13 $[Au_{25}(GSH)_{18}]^{-1}$-$(H_2O)_{126}$ clusters was considered for the ECD calculations. A good qualitative agreement between the experimental and calculated ECD was found, despite some discrepancies in the higher energy region of the spectrum. Such differences are expected to be associated with limitations of the classical MD simulations, which produce an imbalance in the statistical weights. Moreover, the ECD calculations were performed with some approximations due to the size (more than 1000 atoms) and complexity of this non-covalent system. The former hypothesis was corroborated lowering the number of statistically relevant conformations in the final ECD, obtaining a much better agreement with the experiment. Furthermore, this work provided us with the possibility to investigate the active role of water on the chiroptical features of protected gold nanoclusters. It was observed that: (i) the molecular orbitals of water actively contribute to the high-energy electronic transitions, (ii) the spectral features at high energy are completely suppressed removing the solvent. It was also found out that the water-shell itself assumes a chiral arrangement due to the inter-molecular interactions, producing a non-null ECD response. As far as we are concerned, this is the first work where explicit solvent molecules are included in the ECD calculation of a gold nanocluster.

This approach for computing ECD also revealed to be helpful to correlate the conformational behavior in solution of heterochiral dipeptides and their solid-state evolution. In this

last work, a set of phenylalanine-based dipeptides containing one aliphatic amino acid (valine, norvaline, or isoleucine) was studied. These systems are quite interesting because of their ability to self-assemble into gels and/or crystals. We started comparing the experimental and calculated spectra to assess the quality of the computational results. Afterwards, we proceeded investigating in detail the conformational structures extracted to calculate the spectra. It was observed that sequences with Phe at the C-terminus are characterized by only folded conformers in the lowest energy regions of their conformational landscapes. Instead, sequences with Phe at the N-terminus and a bulky aliphatic amino acid (*e.g.* Val, or Ile) are defined by both extended and folded low-energy conformers, with the first ones being predominant. Remarkably, the folded conformations closely resemble the corresponding crystal-unit structures. This result explains why systems characterized by mostly low-energy folded conformers tend to crystallize very quickly in the real experimental sample. Instead, systems characterized by both extended and folded low-energy conformers give the gel first, crystallizing over days. Since the extended conformers are predominant here, we can suggest that they contribute to the fiber formation, while over time the equilibrium is pushed towards the folded structure, leading to the crystallization. It is known from the literature that folded conformations are favored by enthalpy, while extended ones are entropically-driven. Therefore, several factors can be used to explain and justify why our sequences assume certain conformations, hence undergoing or not the gellation process. In order to validate this hypothesis, we played on thermal effects heating the samples of dipeptides that only crystallize at room temperature. Going up to 70 °C, the conformational distribution was clearly changed since only the formation of fibrils was observed at the optical microscope.

Together with the projects regarding the here proposed computational scheme, other three works are presented. All of them share the same goal of studying optical properties, spanning from valence-shell spectroscopies to core-shell ones. Different methodologies were adopted according to the spectral features of interest. TDDFT was employed to calculate valence-shell spectra, while DFT-TP was used for the core-shell ones. However, in the latter case, additional TDDFT calculations were performed for further investigations. Moreover, we worked both in the frequency-, and time-domain. On this subject, a new time-domain protocol for ECD calculations has been successfully implemented and tested along this PhD project. Different systems were investigated as well, ranging from subnanometer metal complexes supported on MgO(100) to an organo-boron system (B_2cat_2) and molecules suitable for protocol validation.

Finally, the procedure developed in this PhD project has shown to be a valid tool for the study of ECD features and their relation with conformational, temperature and solvent effects. The affordable computational cost makes this procedure suitable for complex chiral systems, spanning from biomolecules to nanoclusters, of great interest nowadays. For instance, the good compromise between accuracy and computational effort allowed us to calculate, for the first time, the ECD spectra of RS-AuNCs with an explicit inclusion of conformational and solvent effects.

Bibliography

[1]   H. Wang, Z. Feng, B. Xu, *Angew. Chem.* **2019**, *131*, 10532–10541.

[2]   R. C. Price, R. L. Whetten, *J. Phys. Chem. B* **2006**, *110*, 22166–22171.

[3]   N. Berova, L. Di Bari, G. Pescitelli, *Chem. Soc. Rev.* **2007**, *36*, 914–931.

[4]   G. Pescitelli, L. Di Bari, N. Berova, *Chem. Soc. Rev.* **2011**, *40*, 4603–4625.

[5]   G. Pescitelli, T. Bruhn, *Chirality* **2016**, *28*, 466–474.

[6]   S. Barlow, R. Raval, *Surf. Sci. Rep.* **2003**, *50*, 201–341.

[7]   G. H. Wagnière, *On chirality and the universal asymmetry: reflections on image and mirror image*, John Wiley & Sons, **2007**.

[8]   T. G. Solomons, C. B. Fryhle, *Organic chemistry*, John Wiley & Sons, **2008**.

[9]   J. J. Pelayo, I. Valencia, A. P. Garcıa, L. Chang, M. López, D. Toffoli, M. Stener, A. Fortunelli, I. L. Garzón, *Adv. Phys. X* **2018**, *3*, 1509727.

[10] N. Berova, K. Nakanishi, R. W. Woody, *Circular dichroism: principles and applications*, John Wiley & Sons, **2000**.

[11] P. J. Stephens, *J. Phys. Chem.* **1985**, *89*, 748–752.

[12] J. P. Riehl, F. S. Richardson, *Chem. Rev.* **1986**, *86*, 1–16.

[13] J. Kumar, K. G. Thomas, L. M. Liz-Marzán, *Chem. Commun.* **2016**, *52*, 12555–12569.

[14] T. Kimura, K. Hamase, Y. Miyoshi, R. Yamamoto, K. Yasuda, M. Mita, H. Rakugi, T. Hayashi, Y. Isaka, *Sci. Rep.* **2016**, *6*, 26137.

[15] L. E. Gray Jr, J. Ostby, J. Furr, M. Price, D. R. Veeramachaneni, L. Parks, *Toxicol. Sci.* **2000**, *58*, 350–365.

[16] J. H. Kim, A. R. Scialli, *Toxicol. Sci.* **2011**, *122*, 1–6.

[17] N. Vargesson, *Birth Defects Res. Part C: Embryo Today: Reviews* **2015**, *105*, 140–156.

[18] S. Saha, S. K. Alamsetti, C. Schneider, *Chem. Commun.* **2015**, *51*, 1461–1464.

[19] C. Müller, A. Bauer, T. Bach, *Angew. Chem.* **2009**, *121*, 6767–6769.

[20] N. Kameta, M. Masuda, T. Shimizu, *Chem. Commun.* **2015**, *51*, 11104–11107.

[21] W. A. Bonner, *Orig. Life Evol. Biosph.* **1991**, *21*, 59–111.

[22] M. M. Green, V. Jain, *Orig. Life Evol. Biosph.* **2010**, *40*, 111–118.

[23] D. Sheehan, *Physical biochemistry: principles and applications*, John Wiley & Sons, **2013**.

[24] T. Robinson, *Life sciences* **1976**, *19*, 1097–1102.

[25] J. J. Corrigan, *Science* **1969**, *164*, 142–149.

[26] J. J. Bastings, H. M. van Eijk, S. W. Olde Damink, S. S. Rensen, *Nutrients* **2019**, *11*, 2205.

[27] P. M. Masters, J. L. Bada, S. J. Zigler, *Nature* **1977**, *268*, 71–73.

[28] N. Fujii, T. Takata, N. Fujii, K. Aki, H. Sakaue, *Biochim. Biophys. Acta (BBA)-Proteins and Proteomics* **2018**, *1866*, 840–847.

[29] I. Budin, J. W. Szostak, *Annu. Rev. Biophys.* **2010**, *39*, 245–263.

[30] U. J. Meierhenrich, *Eur. Rev.* **2013**, *21*, 190–199.

[31] V. A Pavlov, E. I Klabunovskii, *Curr. Org. Chem.* **2014**, *18*, 93–114.

[32] F. Frank, *Biochim. Biophys. Acta* **1953**, *11*, 459–463.

[33] Y. Wang, J. Xu, Y. Wang, H. Chen, *Chem. Soc. Rev.* **2013**, *42*, 2930–2962.

[34] T. R. Jensen, M. D. Malinsky, C. L. Haynes, R. P. Van Duyne, *J. Phys. Chem. B* **2000**, *104*, 10549–10556.

[35] W. Wu, P. N. Njoki, H. Han, H. Zhao, E. A. Schiff, P. S. Lutz, L. Solomon, S. Matthews, M. M. Maye, *J. Phys. Chem. C* **2011**, *115*, 9933–9942.

[36] S. Gomez-Grana, C. Fernández-López, L. Polavarapu, J.-B. Salmon, J. Leng, I. Pastoriza-Santos, J. Perez-Juste, *Chem. Mater.* **2015**, *27*, 8310–8317.

[37] Y. Zhao, F. Zhao, X. Wang, C. Xu, Z. Zhang, G. Shi, L. Qu, *Angew. Chem. Int. Ed.* **2014**, *53*, 13934–13939.

[38] O. Gülseren, F. Ercolessi, E. Tosatti, *Phys. Rev. Lett.* **1998**, *80*, 3775–3778.

[39] D. Bochicchio, R. Ferrando, *Nano Lett.* **2010**, *10*, 4211–4216.

[40] M. Farrag, M. Tschurl, U. Heiz, *Chem. Mater.* **2013**, *25*, 862–870.

[41] T. Nakashima, Y. Kobayashi, T. Kawai, *J. Am. Chem. Soc.* **2009**, *131*, 10342–10343.

[42] D. V. Kosynkin, W. Lu, A. Sinitskii, G. Pera, Z. Sun, J. M. Tour, *ACS Nano* **2011**, *5*, 968–974.

[43] Y. Zheng, L. Lin, X. Ye, F. Guo, X. Wang, *Angew. Chem.* **2014**, *126*, 12120–12124.

[44] C. Gautier, T. Bürgi, *ChemPhysChem* **2009**, *10*, 483–492.

[45] S. Srivastava, A. Santos, K. Critchley, K.-S. Kim, P. Podsiadlo, K. Sun, J. Lee, C. Xu, G. D. Lilly, S. C. Glotzer, N. A. Kotov, *Science* **2010**, *327*, 1355–1359.

[46] P. Koivunen, S. Lee, C. G. Duncan, G. Lopez, G. Lu, S. Ramkissoon, J. A. Losman, P. Joensuu, U. Bergmann, S. Gross, et al., *Nature* **2012**, *483*, 484–488.

[47] A. Funston, *Nano Lett* **2009**, *9*, 1651–1658.

[48] T. G. Schaaff, G. Knight, M. N. Shafigullin, R. F. Borkman, R. L. Whetten, *J. Phys. Chem. B* **1998**, *102*, 10643–10646.

[49] Y. Li, D. Yu, L. Dai, A. Urbas, Q. Li, *Langmuir* **2011**, *27*, 98–103.

[50] V. Bonačić-Kouteckỳ, R. Mitrić, C. Bürgel, B. Schäfer-Bung, *Comput. Mater. Sci.* **2006**, *35*, 151–157.

[51] C. Gautier, T. Bürgi, *J. Am. Chem. Soc.* **2006**, *128*, 11079–11087.

[52] C. P. Joshi, M. S. Bootharaju, O. M. Bakr, *J. Phys. Chem. Lett.* **2015**, *6*, 3023–3035.

[53] R. Jin, *Nanoscale* **2010**, *2*, 343–362.

[54] L. Jin, D. S. Weinberger, M. Melaimi, C. E. Moore, A. L. Rheingold, G. Bertrand, *Angew. Chem. Int. Ed.* **2014**, *53*, 9059–9063.

[55] Y. Song, S. Wang, J. Zhang, X. Kang, S. Chen, P. Li, H. Sheng, M. Zhu, *J. Am. Chem. Soc.* **2014**, *136*, 2963–2965.

[56] P. D. Jadzinsky, G. Calero, C. J. Ackerson, D. A. Bushnell, R. D. Kornberg, *Science* **2007**, *318*, 430–433.

[57] H. Qian, W. T. Eckenhoff, Y. Zhu, T. Pintauer, R. Jin, *J. Am. Chem. Soc.* **2010**, *132*, 8280–8281.

[58] J. Akola, M. Walter, R. L. Whetten, H. Häkkinen, H. Grönbeck, *J. Am. Chem. Soc.* **2008**, *130*, 3756–3757.

[59] C. Zeng, R. Jin, *Chem. Asian J.* **2017**, *12*, 1839–1850.

[60] C. Noguez, I. L. Garzón, *Chem. Soc. Rev.* **2009**, *38*, 757–771.

[61] O. Lopez-Acevedo, H. Tsunoyama, T. Tsukuda, H. Häkkinen, C. M. Aikens, *J. Am. Chem. Soc.* **2010**, *132*, 8210–8218.

[62] X. Gu, S. Bulusu, X. Li, X. C. Zeng, J. Li, X. G. Gong, L.-S. Wang, *J. Phys. Chem. C* **2007**, *111*, 8228–8232.

[63] A. Lechtken, D. Schooss, J. R. Stairs, M. N. Blom, F. Furche, N. Morgner, O. Kostko, B. von Issendorff, M. M. Kappes, *Angew. Chem. Int. Ed.* **2007**, *46*, 2944–2948.

[64] I. E. Santizo, F. Hidalgo, L. A. Pérez, C. Noguez, I. L. Garzón, *J. Phys. Chem. C* **2008**, *112*, 17533–17539.

[65] M.-R. Goldsmith, C. B. George, G. Zuber, R. Naaman, D. H. Waldeck, P. Wipf, D. N. Beratan, *Phys. Chem. Chem. Phys.* **2006**, *8*, 63–67.

[66] J. Autschbach, T. Ziegler, *J. Chem. Phys.* **2002**, *116*, 891–896.

[67] A. D. Becke, *J. Chem. Phys* **1993**, *98*, 5648–5652.

[68] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[69] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[70] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[71] M. E. Casida in *Recent Advances In Density Functional Methods: (Part I)*, World Scientific, **1995**, pp. 155–192.

[72] E. R. Davidson, *J. Comput. Phys.* **1975**, *17*, 87–94.

[73] M. de Wergifosse, J. Seibert, S. Grimme, *J. Chem. Phys.* **2020**, *153*.

[74] F. Bohle, J. Seibert, S. Grimme, *J. Org. Chem.* **2021**, *86*, 15522–15531.

[75] T. Kurtán, R. Jia, Y. Li, G. Pescitelli, Y.-W. Guo, *Eur. J. Org. Chem.* **2012**, *2012*, 6722–6728.

[76] S. K. Avula, H. Hussain, R. Csuk, S. Sommerwerk, P. Liebing, M. Górecki, G. Pescitelli, A. Al-Rawahi, N. Ur Rehman, I. R. Green, A. Al-Harrasi, *Tetrahedron: Asymmetry* **2016**, *27*, 829–833.

[77] T. Dražić, M. Roje, M. Jurin, G. Pescitelli, *Eur. J. Org. Chem.* **2016**, *2016*, 4189–4199.

[78] G. Mancini, M. Fusè, F. Lazzari, B. Chandramouli, V. Barone, *J. Chem. Phys.* **2020**, *153*, 124110.

[79] F. Paesani, *Acc. Chem. Res.* **2016**, *49*, 1844–1851.

[80] T. Giovannini, F. Egidi, C. Cappelli, *Chem. Soc. Rev.* **2020**, *49*, 5664–5677.

[81] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999–3094.

[82] M. Losada, P. Nguyen, Y. Xu, *J. Phys. Chem. A* **2008**, *112*, 5621–5627.

[83] P. Mukhopadhyay, G. Zuber, P. Wipf, D. N. Beratan, *Angew. Chem. Int. Ed.* **2007**, *46*, 6450.

[84] B. Mennucci, *WIREs Comput. Mol. Sci.* **2012**, *2*, 386–404.

[85] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[86] U. N. Morzan, D. J. Alonso de Armino, N. O. Foglia, F. Ramirez, M. C. Gonzalez Lebrero, D. A. Scherlis, D. A. Estrin, *Chem. Rev.* **2018**, *118*, 4071–4113.

[87] T. Giovannini, F. Egidi, C. Cappelli, *Phys. Chem. Chem. Phys.* **2020**, *22*, 22864–22879.

[88] S. Del Galdo, M. Fusè, V. Barone, *J. Chem. Theory Comp.* **2020**, *16*, 3294–3306.

[89] R. D'Cunha, T. D. Crawford, *J. Phys. Chem. A* **2021**, *125*, 3095–3108.

[90] M. Zhu, C. M. Aikens, F. J. Hollander, G. C. Schatz, R. Jin, *J. Am. Chem. Soc.* **2008**, *130*, 5883–5885.

[91] A. Baghdasaryan, K. Martin, L. M. L. Daku, M. M. Talamo, N. Avarvari, T. Bürgi, *Nanoscale* **2020**, *12*, 18160–18170.

[92] F. Hidalgo, A. Sánchez-Castillo, C. Noguez, *Phys. Rev. B* **2009**, *79*, 075438.

[93] O. Baseggio, G. Fronzoni, M. Stener, *J. Chem. Phys.* **2015**, *143*, 024106.

[94] O. Baseggio, D. Toffoli, G. Fronzoni, M. Stener, L. Sementa, A. Fortunelli, *J. Phys. Chem. C* **2016**, *120*, 24335–24345.

[95] D. Toffoli, O. Baseggio, G. Fronzoni, M. Stener, A. Fortunelli, L. Sementa, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3585–3596.

[96] I. Dolamic, B. Varnholt, T. Bürgi, *Nature Commun.* **2015**, *6*, 7117.

[97] B. Nieto-Ortega, T. Burgi, *Acc. Chem. Res.* **2018**, *51*, 2811–2819.

[98] L. Riccardi, F. De Biasi, M. De Vivo, T. Bürgi, F. Rastrelli, G. Salassa, *ACS nano* **2019**, *13*, 7127–7134.

[99] D. M. Chevrier, L. Raich, C. Rovira, A. Das, Z. Luo, Q. Yao, A. Chatt, J. Xie, R. Jin, J. Akola, P. Zhang, *J. Am. Chem. Soc.* **2018**, *140*, 15430–15436.

[100] V. Rojas-Cervellera, C. Rovira, J. Akola, *J. Phys. Chem. Lett.* **2015**, *6*, 3859–3865.

[101] M. P. Allen et al., *Computational soft matter: from synthetic polymers to proteins* **2004**, *23*, 1–28.

[102] R. H. Fowler, *Statistical mechanics: the theory of the properties of matter in equilibrium*, Cambridge University Press, **1980**.

[103] D. Fincham, *Comput. Phys. Commun.* **1986**, *40*, 263–269.

[104] L. Verlet, *Phys. Rev.* **1967**, *159*, 98–103.

[105] R. W. Hockney, S. Goel, J. Eastwood, *J. Comput. Phys* **1974**, *14*, 148–158.

[106] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *J. Chem. Phys.* **1982**, *76*, 637–649.

[107] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701–1718.

[108] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781–1802.

[109] C. W. Gear, *Prentice-Hall series in automatic computation* **1971**.

[110] H. Yoshida, *Phys. Lett. A* **1990**, *150*, 262–268.

[111] L. Boltzmann, *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie:(vorgelegt in der Sitzung am 8. Februar 1866)*, Staatsdruckerei, **1866**.

[112] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[113] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*.

[114] H. C. Andersen, *J. Chem. Phys.* **1980**, *72*, 2384–2393.

[115] S. Nosé, *J. Chem. Phys.* **1984**, *81*, 511–519.

[116] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695–1697.

[117] G. J. Martyna, M. L. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635–2643.

[118] M. Parrinello, A. Rahman, *J. Appl. Phys.* **1981**, *52*, 7182–7190.

[119] P. M. Morse, *Phys. Rev.* **1929**, *34*, 57.

[120] P. P. Ewald, *Ann. Phys.* **1921**, *369*, 253–287.

[121]   T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[122]   U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.

[123]   W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[124]   B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. a. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187–217.

[125]   W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

[126]   G. Brancolini, D. Toroz, S. Corni, *Nanoscale* **2014**, *6*, 7903–7911.

[127]   E. Pohjolainen, X. Chen, S. Malola, G. Groenhof, H. Häkkinen, *J. Chem. Theo. Comput.* **2016**, *12*, 1342–1350.

[128]   W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.

[129]   H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, *Intermolecular Forces*, Reidel Dordrecht, **1981**.

[130]   H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **1987**, *91*, 6269–6271.

[131]   K. Toukan, A. Rahman, *Phys. Rev. B* **1985**, *31*, 2643–2648.

[132]   B. Brooks, M. Karplus, *Proc. Natl. Acad. Sci.* **1983**, *80*, 6571–6575.

[133]   M. Levitt, C. Sander, P. S. Stern, *Int. J. Quant. Chem.* **1983**, *24*, 181–199.

[134]   A. E. García, *Phys. Rev. Lett.* **1992**, *68*, 2696.

[135]   A. Amadei, A. B. M. Linssen, H. J. C. Berendsen, *Proteins: Struct. Funct. Genet.* **1993**, *17*, 412–425.

[136]   R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, *Biochemistry* **1975**, *14*, 5355–5373.

[137]   R. Elber, M. Karplus, *Science* **1987**, *235*, 318–321.

[138]   S. Hayward, A. Kitao, N. Gō, *Proteins: Struct. Funct. Genet.* **1995**, *23*, 177–186.

[139]   I. Daidone, A. Amadei, *WIREs Comput. Mol. Sci.* **2012**, *2*, 762–770.

[140]   M. E. Karpen, D. J. Tobias, C. L. I. Brooks, *Biochemistry* **1993**, *32*, 412–420.

[141]   D. van Aalten, D. Conn, B. de Groot, H. Berendsen, J. Findlay, A. Amadei, *Biophys. J.* **1997**, *73*, 2891–2896.

[142]   B. L. de Groot, X. Daura, A. E. Mark, H. Grubmüller, *J. Mol. Bio.* **2001**, *309*, 299–313.

[143]   O. F. Lange, H. Grubmüller, *J. Phys. Chem. B* **2006**, *110*, 22842–22852.

[144]   A. Amadei, M. A. Ceruso, A. Di Nola, *Proteins* **1999**, *36*, 419–424.

[145] U. H. Hansmann, *Chem. Phys. Lett.* **1997**, *281*, 140–150.

[146] S. Piana, A. Laio, *J. Phys. Chem. B* **2007**, *111*, 4553–4559.

[147] G. Torrie, J. Valleau, *J. Comput. Phys.* **1977**, *23*, 187–199.

[148] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562–12566.

[149] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, **1989**.

[150] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–B871.

[151] M. Levy, *Proc. Natl. Acad. Sci.* **1979**, *76*, 6062–6065.

[152] E. H. Lieb, *Int. J. Quantum Chem.* **1983**, *24*, 243–277.

[153] M. Born, R. Oppenheimer, *Ann. Phys. (Berlin)* **1927**, *389*, 457–484.

[154] L. H. Thomas, *Proc. Cambridge Phil. Soc.* **1927**, *23*, 542–548.

[155] E. Fermi, *Rend. Accad. Naz. Lincei* **1927**, *6*, 602–607.

[156] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.

[157] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931–967.

[158] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian~16 Revision C.01, Gaussian Inc. Wallingford CT, **2016**.

[159] J. C. Slater, *Phys. Rev.* **1951**, *81*, 385.

[160] D. M. Ceperley, B. J. Alder, *Phys. Rev. Lett.* **1980**, *45*, 566–569.

[161] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[162] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865.

[163] R. van Leeuwen, E. J. Baerends, *Phys. Rev. A* **1994**, *49*, 2421–2431.

[164] J. P. Perdew, S. Kurth, A. š. Zupan, P. Blaha, *Phys. Rev. Lett.* **1999**, *82*, 2544–2547.

[165] J. P. Perdew, V. N. Staroverov, J. Tao, G. E. Scuseria, *Phys. Rev. A* **2008**, *78*, 052513–052525.

[166] L. A. Constantin, J. M. Pitarke, J. F. Dobson, A. Garcia-Lekue, J. P. Perdew, *Phys. Rev. Lett.* **2008**, *100*, 036401.

[167] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[168] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[169] J. Heyd, G. E. Scuseria, M. Ernzerhof, *J. Chem. Phys.* **2003**, *118*, 8207–8215.

[170] H. Iikura, T. Tsuneda, T. Yanai, K. Hirao, *J. Chem. Phys.* **2001**, *115*, 3540–3544.

[171] M. Seth, T. Ziegler, *J. Chem. Theo. Comput.* **2012**, *8*, 901–907.

[172] E. Runge, E. K. U. Gross, *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

[173] R. van Leeuwen, *Phys. Rev. Lett.* **1999**, *82*, 3863–3866.

[174] I. Tamm, *J. Phys. USSR* **1945**, *9*, 449.

[175] S. M. Dancoff, *Phys. Rev.* **1950**, *78*, 382.

[176] E. K. U. Gross, W. Kohn, *Phys. Rev. Lett.* **1985**, *55*, 2850–2852.

[177] L. Rosenfeld, *Zeitschrift für Physik* **1929**, *52*, 161–174.

[178] P. Pyykkö, J. P. Desclaux, *Acc. Chem. Res.* **1979**, *12*, 276–281.

[179] J. P. Desclaux, P. Pyykkö, *Chem. Phys. Lett.* **1976**, *39*, 300–303.

[180] P. A. M. Dirac, *Proceedings of the Royal Society of London. Series A Containing Papers of a Mathematical and Physical Character* **1928**, *117*, 610–624.

[181] M. Douglas, N. M. Kroll, *Ann. Phys.* **1974**, *82*, 89–155.

[182] E. J. van Lenthe, E. Baerends, J. G. Snijders, *J. Chem. Phys.* **1993**, *99*, 4597–4610.

[183] B. Hess, H. Bekker, H. Berendsen, J. Fraaije, *J. Comput. Chem* **1998**, *18*, 1463–1472.

[184] T. Bruhn, A. Schaumlöffel, Y. Hemberger, G. Bringmann, *Chirality* **2013**, *25*, 243–249.