Dimerization dynamics of carboxylic acids in helium nanodroplets

Andrew M Ellis, Julia A Davies, Ersin Yurtsever, Florent Calvo

Dimerization dynamics of carboxylic acids in helium nanodroplets

Andrew M. Ellis,[1] Julia A. Davies,[1,2] Ersin Yurtsever,[3] Florent Calvo[4,*]

[1] School of Chemistry, University of Leicester, University Road, Leicester, LE1 7RH, UK

[2] Present address: Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

[3] Department of Chemistry, Koç University, Rumelifeneri Yolu, Sariyer, 34450 Istanbul, Turkey

[4] Université Grenoble Alpes, CNRS, LiPhy, F38000, Grenoble, France

Email: florent.calvo@univ-grenoble-alpes.fr

Manuscript submitted to the *Journal of Chemical Physics*

Abstract

The dimerization of molecules in helium nanodroplets is known to preferentially yield structures of higher energy than the global energy minimum structure for a number of quite different monomers. Here we explore dimerization in this environment using an atomistic model within statistically converged molecular dynamics (MD) trajectories, treating the solvent implicitly through the use of a thermostat, or more explicitly by embedding one monomer in a He_{100} cluster. The focus is on the two simplest carboxylic acids, formic and acetic, both of which have been studied experimentally. While the global minimum structure, which comprises two CO···HO hydrogen bonds, is predicted to be the most abundant dimer in absence of the helium solvent, this is no longer the case once helium atoms are included. The simulations confirm the importance of kinetic trapping effects and also shed light on the occurrence of specific dynamical effects, leading to the occasional formation of high-energy structures away from minima, such as saddle configurations. Theoretically predicted infrared spectra, based on the MD statistics, are in good agreement with the experimental spectra.

1. Introduction

There is clear evidence for kinetic trapping in local minima when molecular clusters are formed in liquid helium nanodroplets. Prime examples include clusters of HCN,[1] the trimer of OCS,[2] the water hexamer,[3] and the dimers of formic and acetic acid.[4-6] For $(HCN)_{n>2}$ and $(OCS)_3$, linear structures result despite cyclic structures having a comparable or even significantly lower energy. These observations have been attributed to dipole-dipole attraction, which steers the individual molecules into head-to-tail chain orientations as they approach. Rearrangement would normally occur into lower energy cyclic structures but the rapid cooling and low temperature (*ca*. 0.4 K) provided within a helium nanodroplet quenches each cluster before it has chance to rearrange. The mechanism proposed for the preferential formation of a cyclic hexamer of water is different. The precursor to the hexamer is the water pentamer, which is cyclic in its global energy minimum. It is thought that a water molecule inserts into this pentamer ring to form a cyclic hexamer. Rearrangement into the lower energy cage structure is then prevented by rapid cooling by the surrounding helium.[7]

No mechanism has so far been proposed to fully explain how the dimerization of formic and acetic acid in superfluid helium leads to metastable structures, even though earlier investigations, based on Car-Parrinello molecular dynamics simulations in the gas phase,[8] have already emphasized the roles of temperature and kinetics on structural diversity under cryogenic conditions. The global energy minimum of these dimers has a cyclic hydrogen bonding arrangement comprising two strong (CO···OH) hydrogen bonds. However, while this is the overwhelmingly dominant species seen in experiments performed in the gas phase,[9] this isomer was not detected in experiments utilizing helium nanodroplets.[4-6] Instead, structures based on a single strong hydrogen bond and a much weaker secondary hydrogen bond are the only ones observed. Long-range dipole-dipole interactions have been proposed as having some role in the case of the formic acid dimer[4] but they cannot account for the

single specific structure observed in helium. Additional phenomena must therefore be operating, and these might also occur in the formation of clusters of other molecules in helium nanodroplets. As is the case for water clusters, the non-covalent nature of hydrogen bonding and the dual character of carboxylic acids, as both hydrogen-bond donor and acceptor, explain such a structural competition, which is already manifested in the different dimers exhibited in the crystalline,[10] liquid,[11-13] and gas[14,15] phases.

To uncover this behavior, we present a scheme based on atomistic modeling and within the framework of classical molecular dynamics (MD). Quantum descriptions based, for example, on the path-integral framework, are necessary to reliably model low-temperature properties, and especially the superfluid behavior of the bosonic helium environment around molecular dopants.[16] However they are also limited in terms of statistics or time scales, making it necessary to use approximations for addressing time-dependent processes.[17-19] Classical descriptions of nuclear motion but involving an explicit description of electronic structure[8] are also vulnerable to this sampling issue. Alternatively, quantum-corrected or even purely classical MD, can provide valuable insight into the effects of the solvent on the preferred molecular structures, as well as the formation dynamics itself.[20,21] The classical model presented here has the potential to be applied to other examples of cluster formation in superfluid helium.

Despite its simplicity, our model confirms the key role played by the solvent in stabilizing high-energy dimer configurations, with the predicted statistics correctly accounting for the experimentally observed infrared spectra. The methodological details are provided in the next section, which is followed by a presentation of the results and associated discussion.

2. Methods

Our computational approach to the dimerization of complex organic molecules in helium nanodroplets relies on atomistic modeling, at a level of approximation that allows us to account for statistical effects in a fully converged way. The main strategy is to describe the dimerization process as a collision between one isolated monomer and another monomer embedded in a helium nanodroplet. To cope with the large number of trajectories needed to cover the various possible dimerization pathways, some further approximations are necessary, the first one being the use of a flexible, nonpolarizable force field to describe the intramolecular and intermolecular contribution to the interaction. While such force fields are computationally efficient, they are not expected to be particularly reliable for predicting vibrational spectra and so for that we turn to quantum chemical methods, at the static level, to predict the individual infrared absorption spectra.

2.1 Classical molecular dynamics protocol

To simulate the effect of helium microsolvation, one of the monomers was located within a He_{100} nanodroplet prior to addition of a second monomer. A cluster of this size was chosen to provide sufficient helium atoms to act as a decent mimic of a helium droplet, while being sufficiently small to make calculations on a large number trajectories a viable proposition. The collision partners were initially prepared at 10 K using classical molecular dynamics and a Nosé-Hoover thermostat with relaxation time of 50 fs. Such a temperature is low, approaching that of a real helium nanodroplet, while still being high enough for the helium atoms to behave semiclassically. In particular, bosonic exchange statistics are not expected to play a significant role at 10 K. The remaining nuclear quantum effects will be ascribable to vibrational delocalization and, again, we do not anticipate these playing a major role for dimers of carboxylic acids, and hence their omission is justified. Note that a temperature of 10 K lies considerably above the experimentally temperature of 0.4 K for a helium

nanodroplet, but in the absence of zero-point delocalization the classical description keeps spontaneous evaporation limited until collisions take place.

For both formic and acetic acids, the two monomers that are collision partners were first simulated using thermostatted molecular dynamics, the two partners being prepared either in vacuum or embedded in the He_{100} cluster, before eventually meeting on another. Equilibrium samples were generated by performing MD trajectories in the canonical ensemble at 10 K, from which 10^4 phase space configurations (positions and velocities) were stored every 1 ps. A time step of 0.5 fs was used to integrate the equations of motion.

Dimerization trajectories were initiated from random configurations borrowed from these samples. Both configurations of collision partners were then randomly rotated in space with their centers of mass placed 50 Å away from each other along the main intermolecular axis. A random impact parameter *b* in the range of 0–30 Å was also chosen, shifting the partners along a perpendicular axis, and the relative velocities were given a prescribed value *v* of 10 m/s. Additional simulations at 1 m/s indicated no significant variation in the products statistics at low collision energies (vide infra). However, for the present relatively small helium droplets, higher velocities of 100 m/s or above lead to the fast evaporation of all helium atoms with no solvent remaining: only larger droplets would be able to accommodate such collision energies. The collision trajectories were propagated for 500 ps (v=10 m/s) or 5 ns (v=1 m/s), without a thermostat and employing a time step of 0.5 fs. They were stopped earlier than the maximum allowed duration if the distance between the centers of mass between the two organic monomers exceeds 150% of their initial distance, in which case dimerization was considered as not occurring. For each type of dimerization, 10 000 independent dimerization trajectories were performed and the resulting dimers, if connected, were further subject to local optimization after removing the helium, to eliminate the thermal noise caused by collisional heating and to identify the corresponding structure. Note that the

dimerization itself is sufficient to heat up the helium microsolvent and evaporate a number of atoms (such evaporation processes can be visualized in the movies provided as electronic supplementary material).

In addition to the classical MD simulations with one monomer in the helium solvent, and to assess the role of the solvent in more detail and separate its effects from the pure collisional effects, we also considered a simplified case where the solvent is now treated through its thermostatting effect only. Additional MD trajectories were thus conducted for the two monomers in the gas phase, but keeping a global Nosé-Hoover thermostat during the collisional trajectory, with the same relaxation time of 50 fs. The thermostat will redistribute excess energy from the collision, without which many dimers produced would be hot and therefore would transit between different energy minima with no well-defined structure.

2.2 Semiempirical potential and quantum chemical calculations

A classical force field of the Amber *ff99* type[22] was employed to model intramolecular interactions in the carboxylic acid monomers and dimers. Intermolecular forces of the pairwise type within this force field include repulsion-dispersion contributions (Lennard-Jones form) and multipolar electrostatics, assuming partial charges are distributed over the various atoms. These charges were obtained from a standard restricted electrostatic potential (RESP) fitting procedure using a quantum chemical calculation on the isolated monomers, here at the level of density-functional theory (LC-wPBE functional with aug-cc-pVTZ basis set). The equilibrium geometries eventually obtained with the force field, including the RESP charges, are provided as electronic supplementary material for both carboxylic acid monomers.

Despite being chemically quite simple, the force field predicts relative binding energies that are in good agreement with quantum chemistry calculations. More precisely, the systematic exploration provided by molecular dynamics trajectories yield six and eight different dimers for formic and acetic acid, respectively. These structures will be shown and discussed in more detail in section 3. Tables 1 and 2 give the energies of these different structures, as predicted by the Amber force field, in comparison with the results from density-functional theory (LC-wPBE functional with aug-cc-pVTZ basis set), or 2nd order perturbation theory (MP2, 6-311+G(2d,2p) basis set). The contributions of harmonic zero point energies to the DFT and MP2 energies are also indicated. Furthermore, we provided the Pearson correlation coefficients between pairs of methods, to quantify the level of similarity between their predictions, especially in terms of energy ordering. All quantum chemical calculations were carried out using the Gaussian09 set of programs.[24]

Table 1. Relative energies of the calculated equilibrium structures found for formic acid dimer, as obtained using the Amber force field (FF) or the MP2/6-311++G(2d,2p) and DFT/LC-wPBE/aug-cc-pVTZ quantum chemical methods, and also including zero-point energy corrections (ZPE) in the harmonic approximation. All energies are given in kJ/mol.

Isomer	Method				
	Force field	DFT	DFT+ZPE	MP2	MP2+ZPE
1	0.0	0.0	0.0	0.0	0.0
2	22.2	28.5	27.2	25.9	23.6
3	29.2	38.5	36.1	33.6	30.5
4	33.5	45.1	42.2	39.7	35.6
5	41.5	51.6	47.4	48.3	43.0
6	44.9	56.1	51.5	48.4	43.0
R	Pearson correlation coefficient, $R^{a)}$				

	(FF,DFT)=0.998	(DFT,MP2)=0.997

[a)] The Pearson correlation coefficients, *R*, between energy values at the FF and DFT level, or between DFT and MP2, are also reported.

Table 2. Relative energies of the calculated structures of acetic acid dimer, as obtained using the Amber force field (FF) or the MP2/6-311++G(2d,2p) and DFT/LC-wPBE/aug-cc-pVTZ quantum chemical methods, and also including zero-point energy corrections (ZPE) in the harmonic approximation. All energies are given in kJ/mol.

Isomer	Method				
	Force field	DFT	DFT+ZPE	MP2	MP2+ZPE
1	0.0	0.0	0.0	0.0	0.0
2	19.9	33.5	32.5	31.0	29.1
3	28.7	39.5	38.1	34.5	32.4
4	32.2	47.6	46.0	41.0	38.2
5	36.3	57.0	54.3	46.1	42.4
6	37.0	57.0	54.2	50.0	45.7
7	37.7	57.8	55.0	50.0	45.7
8	46.4	63.4	59.9	46.1	42.4
R	Pearson correlation coefficient, $R^{a)}$				
	(FF,DFT)=0.988		(DFT,MP2)=0.983		

[a)] The Pearson correlation coefficients R between energy values at the FF and DFT level, or between DFT and MP2, are also reported.

Although there are significant differences in the relative energies of the various isomers for all three methods, the force field approach maintains the same energy ordering and even agrees rather well with the MP2+ZPE method. For this reason we believe that the force field captures the essential features of the potential energy surface and so can be trusted to produce reliable statistics.

The interaction between the organic molecules and any surrounding helium atoms are assumed to be nonpolarizable for simplicity, and thus consists only of pairwise additive

repulsion-dispersion contributions, again taken as simple Lennard-Jones forms. The relevant LJ coefficients, σ and ε, were calculated from CCSD(T) calculations with an aug-cc-pvTZ basis for the X-He diatomics, with X=O, C, H, N. The values used in the present simulations are shown in Table 3. In addition to the intramolecular part, the force field naturally includes a pairwise He-He interaction, here taken as the accurate Janzen-Aziz potential.[23]

Table 3. Lennard-Jones parameters used to represent the atom-helium interactions in the force field.

Pair	O-He	C-He	H-He	N-He
σ (Å)	2.95	2.82	2.62	3.40
ε (kJ/mol/Å^6)	0.256	0.331	0.065	0.130

In addition to the intramolecular part, the force field naturally includes a pairwise He-He interaction, here taken as the accurate Janzen-Aziz potential.[23]

3. Results

3.1 Formic acid dimers

Figure 1 compares the statistical outcomes of 10^4 MD trajectories for formic acid. Data are shown for two scenarios, one in the gas phase under 10 K thermostatting and the other in a He_{100} cluster under microcanonical conditions. For the former conditions, the global energy minimum of the dimer, isomer 1, is found to be the most abundant species formed but there are also significant quantities of other isomers, where the numbering system employed is in order of increasing equilibrium energy. In a system at higher temperature and undergoing multiple collisions, there would be ample opportunity for these higher energy isomers to relax into the global energy minimum. This is the scenario that pertains in gas phase experiments, including spectra recorded under supersonic expansion conditions.[25,26] In these simulations

the thermostat acts at 10 K and has a relaxation time lower than picoseconds, hence the dimers produced this way are mainly driven by long-range forces and correspond to the quenching, upon contact, into the nearest relevant configuration at this low temperature. While such configurations are often close to local minima, and as will be discussed below in greater details, higher-energy structures are also occasionally produced. In any case it would take far too long at 10 K for the high-lying minima to rearrange into the doubly H-bonded isomer 1. The structural diversity found in these thermostatted simulations is entropic in nature, as it originates from the various ways the two monomers can establish contact as they approach each other, before the excess energy is transferred to the heat bath.

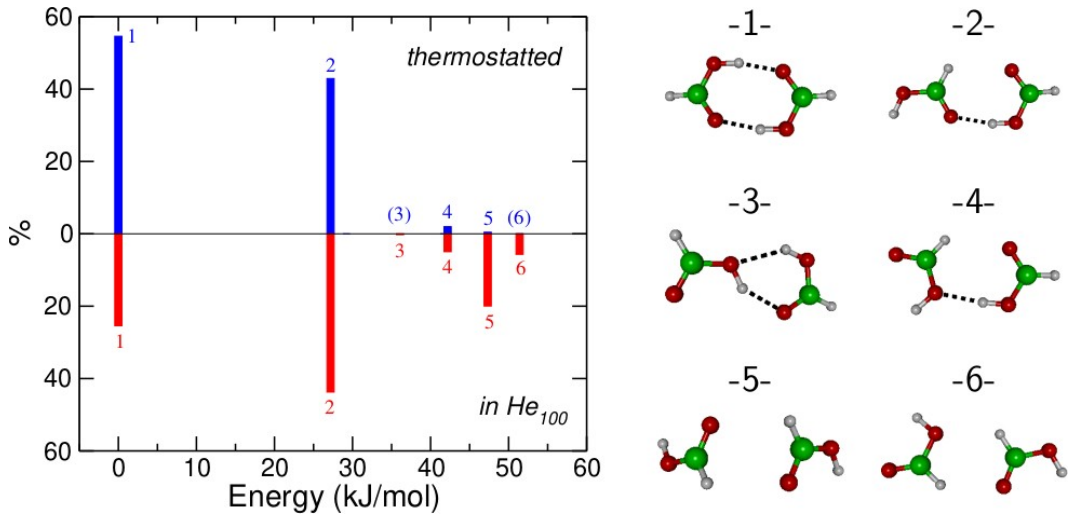

Figure 1. Statistical outcome of 10^4 MD trajectories for formic acid dimerization carried out in a thermostatted gas phase environment (upper panel) and in a He_{100} cluster (lower panel). The isomers of the resulting dimers are labelled in order of increasing energy for their equilibrium structures, with their energies being shown relative to the global energy minimum (isomer 1) on the horizontal axis. The vertical axis shows the percentage abundance of each isomer. The six structures of these dimers are depicted in the panel on the right.

However, the situation changes for microcanonical simulations in a He_{100} cluster. Here, isomer 2 becomes the most abundant species, although isomer 1 is still significant. That isomer 1 is no longer the most abundant structure in the presence of helium reflects the importance of the initial relative orientation of the colliding partners. In the gas phase, even if the two molecules do not orient themselves initially in such a way that the double hydrogen bond can be formed, they can rotate, eventually (but not always) leading to isomer 1. With one molecule embedded into helium, this reorientation process is slower and even hindered. This also explains why new dimer structures are obtained, such as those denoted as 3 and 6 (see later discussion), which are barely seen when dimerization takes place in vacuum. Isomer 5 is the most striking example of a dimer whose formation is negligible in vacuum but becomes very significant (20% abundance) in the helium cluster. Once formed, we would expect the isomer distribution to be 'locked' by the low temperature of the helium.

Table 4(a) demonstrates that the statistics of the various dimers obtained in the explicit presence of the microsolvent around one monomer are robust against collision velocities of both 1 and 10 m/s.

Table 4. Probability of forming the various structures of (a) formic acid dimers and (b) acetic acid dimers by collision of an isolated monomer in a He_{100} cluster containing an embedded monomer, at two different collision velocities. The results were obtained by accumulating the statistics on 10^4 independent MD trajectories.

(a) Formic acid dimers

Isomer	1	2	3	4	5	6
1 m/s	25.4%	43.8%	0.2%	5.0%	20.0%	5.7%

10 m/s	24.2%	46.4%	0.5%	5.5%	17.9%	5.6%

(b) Acetic acid dimers

Isomer	1	2	3	4	5	6	7	8
1 m/s	15.8%	46.5%	1.4%	11.4%	1.6%	13.9%	9.1%	0.3%
10 m/s	17.0%	46.2%	1.2%	10.4%	2.3%	13.2%	9.5%	0.3%

Fig. 2 shows the IR spectra in the OH stretching region calculated for individual dimers selected based on their occurrence probabilities. Also shown is a calculated IR spectrum obtained from a superposition of the IR spectra of individual isomers weighted by their relative abundances determined by the MD simulations. This simulated spectrum can be compared with the experimental spectrum, which is also shown in Figure 2. Despite the greater abundance of isomer 2 in the MD predictions, the dominant signal should come from isomer 1 because it has a larger transition dipole moment for the OH stretching vibration. However, what is not accounted for in the current simulations is the strong anharmonic coupling seen across the CH and OH stretching regions in experiments on isomer 1 for both formic and acetic acid dimers: in both cases this leads to extensive band broadening spanning 2600-3300 cm^{-1}.[11,27-30] In effect, the anharmonic coupling spreads the oscillator strength over a wide spectral range, thereby strongly reducing the absolute signal level for isomer 1 at any one wavelength. This will greatly diminish any contribution from isomer 1 and explains why it is not observed at a detectable level in the helium nanodroplet experiment.[4,5] In the present calculations, attempts to include anharmonic corrections through vibrational perturbation theory turned out to be inconclusive, which is perhaps unsurprising given the likely substantial magnitude of anharmonic coupling, and particularly in the global minimum.

Of the two bands that are seen in the experimental spectrum in Fig. 2, the one centred near 3230 cm^{-1} shows clear sub-structure. This spectral feature was previously attributed[4] to isomer 2, which has one strong CO···HO hydrogen bond and a much weaker CO···HC interaction. The band was assigned to stretching of the OH involved in the CO···HO hydrogen bond and the sub-structure was attributed to vibrational coupling to nearby vibrational dark states. The central position of this fairly broad band coincides quite closely with the strongest band shown for isomer 2 in Fig. 2, and is consistent with the aforementioned assignment. Note that isomer 2 also has an OH group that is not involved in hydrogen-bonding and this is calculated to weakly absorb at just below 3600 cm^{-1}. Again this is in good agreement with the experimental spectrum.

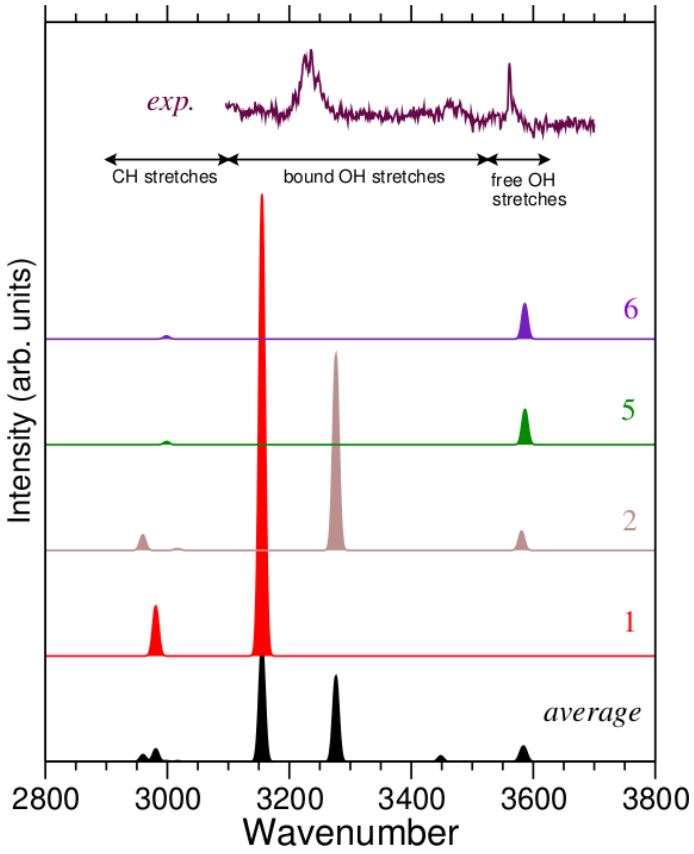

Figure 2. Calculated IR spectra for selected isomers of the dimer of formic acid together with an experimental spectrum recorded using helium nanodroplets.[5] Also shown is a spectrum calculated on the basis of the weighted averages of the isomers found from the MD simulations.

The next most abundant isomer, isomer 5, has no hydrogen bonding and so can be considered as a dimer held together purely via van der Waals forces and higher-order multipolar interactions. The OH stretching vibrations in the two monomer units will therefore be at virtually the same vibrational frequency as that of the free OH group in isomer 2, and so will be indistinguishable. Consequently, the MD simulations are consistent with the experimental finding in helium nanodroplet experiments, namely that a single metastable isomer, isomer 2, dominates the IR spectrum of formic acid dimers in helium nanodroplets.

3.2 Acetic acid dimers

The dimers of acetic acid pose a more challenging test of the simulations given that two distinct metastable isomers were observed in helium nanodroplets.[6] Fig. 3 shows the statistical outcome from the MD trajectories, both in the gas phase (thermostatted) and in He_{100}, along with the corresponding structures. The relative abundances of the isomers resemble the formic acid case, although the abundance of isomer 2 relative to isomer 1 is even more pronounced for acetic acid when formed in the helium cluster. Also notable is that isomer 4 has a three-fold greater abundance for acetic acid than was the case with formic acid. As with formic acid, the relative abundances obtained in He_{100} clusters are robust against collision velocities in the 1–10 m/s range [see Table 4(b)].

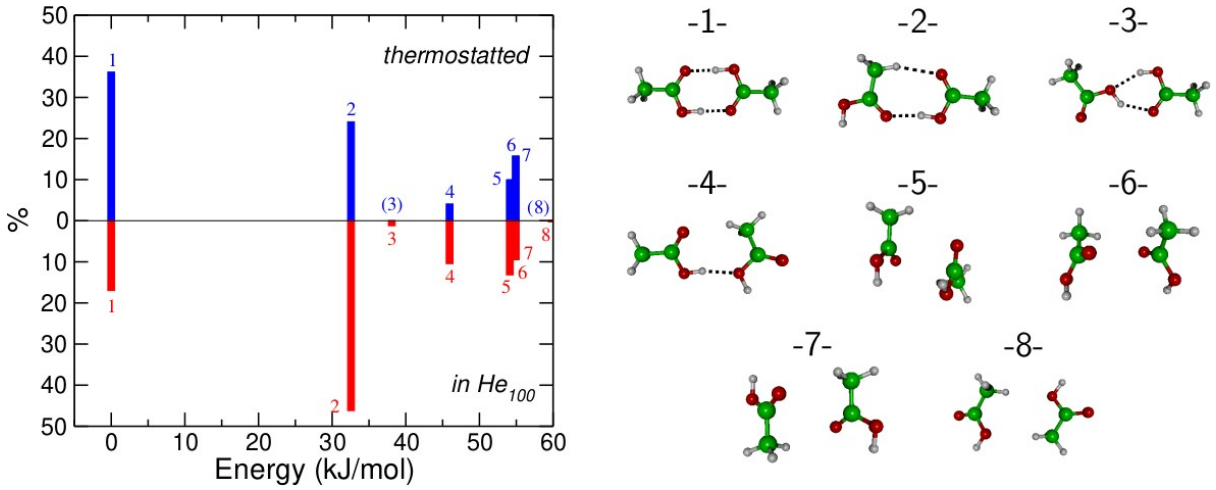

Figure 3. Statistical outcome of 10^4 MD trajectories for acetic acid carried out in the gas phase with a 10 K thermostat (upper panel) and in a He_{100} cluster (lower panel). The isomers of the resulting dimers are labelled in order of increasing energy for their equilibrium structures, with their energies being shown relative to the global energy minimum (isomer 1) on the horizontal axis. The vertical axis shows the percentage abundance of each isomer. The eight structures themselves are depicted in the panel on the right.

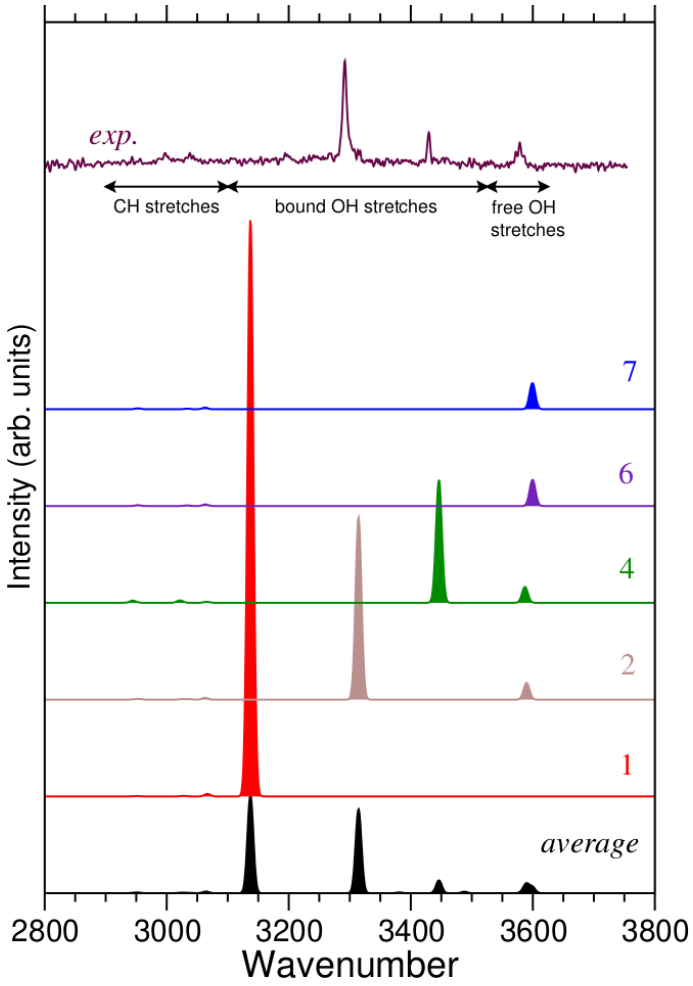

Figure 4. Calculated IR spectra for selected isomers of the dimer of acetic acid together with an experimental spectrum recorded using helium nanodroplets.[6] Also shown is a spectrum calculated on the basis of the weighted averages of the isomers found from the MD simulations.

These predictions agree very well with the experimental findings and their assignment, as can be seen from a comparison of calculated IR spectra with the experimental one (Fig. 4). Similar to the formic acid case, the contribution from isomer 1 will be negligible in practice because of extensive anharmonic coupling, and so we can focus on other isomers. The remainder of the statistically averaged spectrum now closely resembles the experimental spectrum, not only in terms of band positions but also relative intensities, which suggests that the MD calculations do well at predicting relative isomer abundances. Thus isomer 2, which comprises a strong CO···OH hydrogen bond and a much weaker CO···HC interaction, is the source of the intense band just at 3292 cm^{-1} and a much weaker band at ca. 3578 cm^{-1} (free OH stretch). The remaining experimental band, at 3429 cm^{-1}, aligns with the strongest band predicted for isomer 4. This isomer is held together primarily by an OH···OH hydrogen bond, which is intermediate in strength between a CO···HO and CO···HC hydrogen bond. Again this is consistent with the original experimental assignment.[6] Any contributions from other isomers will fall into the free OH stretching region and so are difficult to distinguish from the free OH band belonging to isomer 2.

4. Discussion

Despite its simplicity, the classical MD model used in the present study seems to capture the essential physics that leads to the experimental observations, including the formation of two

distinct and detectable metastable isomers for acetic acid dimer and only one for formic acid. The question therefore turns to what this reveals about the dimerization process.

In the gas phase, the propensity for the two collision partners to form specific dimers depends on their initial orientation and, more importantly, on the collision velocity. The latter determines the total energy available in the dimer, as well as its ability to rearrange and the associated time scale for such rearrangement. Owing to the diversity of impact parameters, the distribution of internal energies can be broad, even though it does not strictly produce a canonical ensemble in itself. Moreover, the orientation itself is not fixed, because the collision partners may also carry angular momentum. As they approach each other, the dominant long-range force (dipole-dipole interaction) comes into play and contributes to orienting the monomers at the expense of some internal energy redistribution. In the present gas phase simulations, the introduction of a thermostat removes the excess energy at a rate comparable to the intramolecular vibrational frequencies, strongly preventing subsequent intermolecular rearrangements. Comparison with experiment indicates that rapid thermostatting alone is insufficient to account for the observed dimerization behavior.

Once a monomer is explicitly embedded into a helium cluster, its angular momentum is strongly reduced but it is still exposed to random orientation from the incoming monomer. The excess energy from a collision is now released by evaporation of successive helium atoms over longer time scales than imposed by the Nosé-Hoover thermostat. While the dipole-dipole interaction remains, the rearrangement process is now further hindered due to the presence of the solvent and the dimer formed is prone to an even greater dependence on the relative orientations of the monomers, once contact is established.

This process can be quantified by comparing the structures of the dimers before and after local re-optimization, through the geometric distance separating them. Here we calculated the root-mean-square distance (RMSD) as the average geometric distance between

equivalent atoms in the two configurations, fixing their centers of mass at a common location but minimizing this quantity among all possible relative orientations using a Monte Carlo procedure. The distributions of RMSD values obtained for formic acid and acetic acid dimers are shown in Fig. 5 for the two sets of collision processes, namely in the thermostatted gas phase or where one one monomer is initially embedded in a He_{100} cluster.

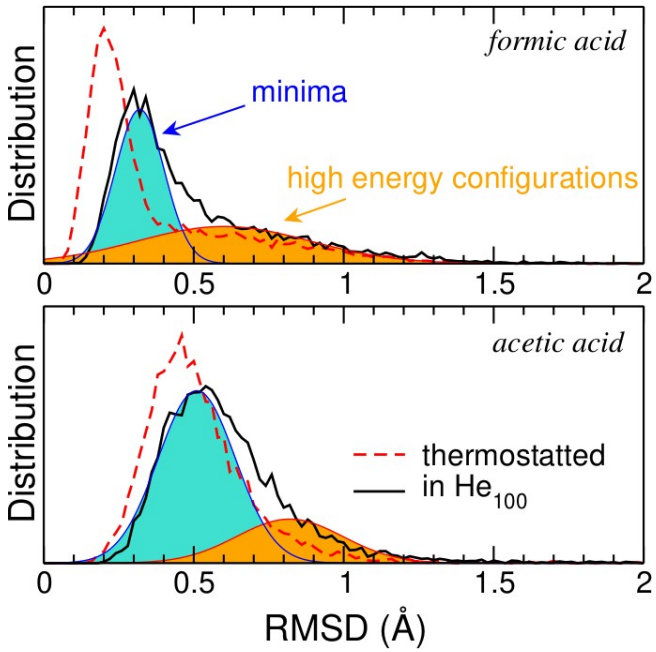

Figure 5. Distributions of root mean square distance between the connected dimer structures before and after local optimization, as obtained from simulations in the gas phase (thermostatted) and with one monomer initially embedded in He_{100}. These have been deconvoluted into two Gaussian functions for formic acid (upper panel) and acetic acid (lower panel) when in helium.

For both systems in helium clusters, the RMSD distributions exhibit a main peak at 0.3 Å for formic acid and 0.5 Å for acetic acid, and these correspond to situations in which

the dimer is already relatively close to a local minimum. However, the RMSD distributions also show a clear shoulder centered near 0.7–0.9 Å. For these dimers, the configurations lie further away from the nearest local minimum, and correspond to higher-energy configurations on the potential energy surface, often close to first-order stationary points (saddle points or transition states). Typical trajectories leading to such high-energy structures, and in particular saddle points, are provided as electronic supplementary material. That these saddle points are found even in the gas phase but under thermostatting, can be explained by the stabilizing effect of centrifugal forces originating from the finite impact parameter and orbital momentum of most collisional trajectories,[31] together with the fast redistribution of excess energy into the heat bath. However, their relative proportion is higher when one monomer is embedded in helium, showing that the helium has a stabilizing effect on less compact dimer configurations.

While we cannot determine using the present methodology how many of such high-energy configurations would be long-lived on experimental time scales, this relative stabilization away from true local minima suggests that the solvent participates actively in hindering the motion along the reaction coordinate, which we interpret as a favorable solvation energy when the dimer exposes a greater accessible area to the solvent in such an extended form. Trapping into a greater diversity of configurations than mere minima should also influence spectroscopic features, notably via an inhomogeneous broadening effect. Thus, the presence of a strongly cryogenic medium could produce two rather opposite effects on spectroscopic transitions, on the one hand causing narrowing of vibrational bands through the cooling of underlying rotational sub-structure, and on the other broadening them through configurations located away from true energy minima.

4. Concluding remarks

In summary, by means of classical molecular dynamics simulations, we have rationalized infrared spectroscopic measurements on formic and acetic acid dimers in helium droplets and confirmed that the signals observed can be attributed to metastable structures that lie higher in energy than the expected, doubly hydrogen bonded dimers. Some structural diversity is merely a consequence of low temperature, in which excess energy from collisions is dissipated into a heat bath. However, a more realistic model that treats the helium solvent explicitly, and allows energy redistribution through evaporation in the microcanonical ensemble, enlarges this diversity significantly and produces high-energy configurations that could further broaden the infrared spectral lines, depending on the time needed for relaxation into nearby local minima. More importantly, a comparison between the experimental spectra and reference IR spectra obtained from quantum chemical calculations indicates that the statistical abundances are correctly captured by our model when the helium environment is explicitly included.

Future efforts will be dedicated to extending the present computational modeling to other dimers and to larger oligomers, to other hydrogen-bonded complexes involving aromatic rings[32] as well as larger but more flexible organic molecules that are characterized by complex energy landscapes[33] and for which structural diversity has also been suggested by infrared spectroscopic measurements. The possible impact of nuclear quantum effects, which were ignored in the present work, deserve to be scrutinized and methods based on path-integral descriptions such as ring-polymer molecular dynamics[34] could be used to assess such effects further while still allowing for acceptable statistics.

Acknowledgements

JAD and AME would like to thank the Leverhulme Trust for financial support for this work (grant RPG-2019-044).

Supporting Information

Cartesian coordinates for all dimers (DFT level); static IR spectra from MP2 and DFT; typical animations for each dimer produced.

References

1 K. Nauta, and R. E. Miller, *Nonequilibrium self-assembly of long chains of polar molecules in superfluid helium*, Science **283**, 1895-1897 (1999).

2 J. D. Pickering, B. Shepperson, L. Christiansen, and H. Staplefeld, *Femtosecond laser induced Coulomb explosion imaging of aligned OCS oligomers inside helium nanodroplets*. J. Chem. Phys. **149**, 154302 (2018).

3 K. Nauta, and R.E. Miller, *Formation of cyclic water hexamer in liquid helium: the smallest piece of ice*. Science **287**, 293-295 (2000).

4 F. Madeja, M. Havenith, K. Nauta, R. E. Miller, J. Chocholoušová, and P. Hobza, *Polar isomer of formic acid dimers formed in helium nanodroplets*. J. Chem. Phys. **120**, 10554-10560 (2004).

5 K. A. E. Meyer, J. A. Davies, A. M. Ellis, *Shifting formic acid dimers into perspective: vibrational scrutiny in helium nanodroplets*. Phys. Chem. Chem. Phys. **22**, 9636-9646 (2020).

6 J. A. Davies, M. W. D. Hanson-Heine, N. A. Besley, A. Shirley, J. Trowers, S. Yang, and A. M. Ellis, *Dimers of acetic acid in helium nanodroplets.* Phys. Chem. Chem. Phys. **21**, 13950-13958 (2019).

7 K. Liu, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory, and D. C. Clary, *Characterization of a cage form of the water hexamer*. Nature (London) **381**, 501-503 (1996).

8 P. Rodziewicz and N. L. Doltsinis, *Formic Acid Dimerization: Evidence for Species Diversity from First Principles Simulations*. J. Phys. Chem. A **113**, 6266-6274 (2009).

9 Y. Maréchal, *IR spectra of carboxylic acids in the gas phase: A quantitative reinvestigation*, J. Chem. Phys. **87**, 6344-6353 (1987).

10 I. Nahringbauer, *Hydrogen bond studies. 39. Reinvestigation of crystal structures of acetic acid (at +5 degrees C and -190 degrees C)*. Acta Chem. Scand. **24**, 453-462 (1970).

11 I. Bertagnolli, *The structure of liquid acetic-acid – an interpretation of neutron-diffraction results by geometrical models*. Chem. Phys. Lett. **93**, 287-292 (1982)..

12 T. Nakabayashi, K. Kosugi, and N. Nishi, *Liquid structure of acetic acid studied by Raman spectroscopy and ab initio molecular orbital calculations*. J. Phys. Chem. A **103**, 8595-8603 (1999).

13 M. Lütgens, F. Friedriszik, and S. Lochbrunner, *Direct observation of the cyclic dimer in liquid acetic by probing the C=O vibration with ultrafast coherent Raman spectroscopy.* Phys. Chem. Chem. Phys. **16**, 18010-18016 (2014).

14 N. Lumbrosobader, C. Coupry, D. Baron, and D. H. Clague, *Dimerization of carboxylic acids: a vapor-phase NMR study.* J. Magn. Reson. **17**, 386-392 (1975).

15 O. Socha and M. Dracinsky, *Dimerization of acetic acid in the gas phase – NMR experiments and quantum-chemical calculations*. Molecules, **25**, 2150 (2020).

16 F. Uhl and D. Marx, *Helium Tagging of Protonated Methane in Messenger Spectroscopy: Does it Interfere with the Fluxionality of CH_5^+?* Angew. Chem. Int. Ed. **57**, 14792-14795 (2018).

17 A. Castillo-Garcia, A. W. Hauser, M. P. de Lara-Castells, and P. A. Villareal, *Path Integral Molecular Dynamics Simulation of a Harpoon-Type Redox Reaction in a Helium Nanodroplet*, Molecules **26**, 5783 (2021).

18 F. Calvo, E. Yurtsever, and Ö Birer, *Possible Formation of Metastable PAH Dimers upon Pickup by Helium Droplets*, J. Phys. Chem. A **120**, 1727-1736 (2016).

19 F. Calvo, and E. Yurtsever, *The metastable structures of anthracene-argon clusters inside helium nanodroplets*, Theor. Chem. Acc. **140**, 21 (2021).

20 C. Leidlmair, Y. Wang, P. Bartl, H. Schöbel, S. Denifl, M. Probst, M. Alcami, F.Martin, H. Zettergren, K. Hansen, O. Echt and P. Scheier, *Structures, Energetics, and Dynamics of Helium Adsorbed on Isolated Fullerene Ions, Phys. Rev. Lett.* **108**, 076101 (2012).

21 A. Volk, P. Thaler, D. Knez, A. W. Hauser, J. Steurer, W. Grogger, F. Hofer, and W. E. Ernst, *The impact of doping rates on the morphologies of silver and gold nanowires grown in helium nanodroplets*, Phys. Chem. Chem. Phys. **16**, 1451-1459 (2016).

22 J. Wang, P. Cieplak, and K. A. Kollman, *How well does a Restrained Electrostatic potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules?* J. Comput. Chem. **21**, 1049-1074 (2000).

23 A. R. Janzen, and R. A. Aziz, *An accurate potential energy curve for helium based on* ab initio *calculations*, J. Chem. Phys. **107**, 914-919 (1997).

24 Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, *et al.* Gaussian, Inc., Wallingford CT, 2009.

25 T. Häber, U. Schmitt, C. Emmeluth, and M. A. Suhm, *Ragout-jet FTIR Spectroscopy of Cluster Isomerism and Cluster Dynamics: From Carboxylic Acid Dimers to N_2O Nanoparticles*. Faraday Discuss. **118**, 331-359 (2001).

26 P. Zielke and M. A. Suhm, *Raman jet spectroscopy of formic acid dimers: low frequency vibrational dynamics and beyond.* Phys. Chem. Chem. Phys. **9**, 4528-4534 (2007).

27 G. M. Florio and T. S. Zwier, *Theoretical modeling of the OH stretch infrared spectrum of carboxylic acid dimers based on first-principles anharmonic couplings*. J. Chem. Phys. **118**, 1735-1746 (2003).

28 C. Emmeluth, M. A. Suhm and D. A. Luckhaus, *Monomers-in-dimers Model for Carboxylic Acid Dimers*, J. Chem. Phys. **118**, 2242-2255 (2003).

29 J. Dreyer, *Hydrogen-bonded Acetic Acid Dimers: Anharmonic Coupling and Linear Infrared Spectra Studied with Density-Functional Theory,* J. Chem. Phys. **122**, 184306 (2005).

30 P. Blaise, M. J. Wojcik, and O. Henri-Rousseau, *Theoretical Interpretation of the Line Shape of the Gaseous Acetic Acid Cyclic Dimer*. J. Chem. Phys. **122**, 064306 (2005).

31 F. Calvo and F. Spiegelmann, *Stabilization of cluster dimers by centrifugal effects.* Z. Phys. D: At. Mol. Clusters **41**, 195-203 (1997).

32 J. Khatri, T. K. Roy, K. Chatterjee, G. Schwaab, and M. Havenith, *Vibrational spectroscopy of Benzonitrile-(Water)$_{1–2}$ Clusters in Helium Droplets*, J. Phys. Chem. A **125**, 6954-6963 (2021).

33 D. A. Thomas, R. Chang, E. Mucha, M. Lettow, K. Greis, S. Gwinner, W. Schöllkopf, G. Meijer, and G. von Helden, *Probing the conformational landscape and thermochemistry of DNA dinucleotide anions via helium nanodroplet infrared action spectroscopy*, Phys. Chem. Chem. Phys. **22**, 18400-18413 (2020).

34 I. R. Craig and D. E. Manolopoulos, *Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics,* J. Chem. Phys. **121**, 3368 (2004).