



**HAL**  
open science

# Infrared Spectra of Deprotonated Dicarboxylic Acids: IRMPD Spectroscopy and Empirical Valence-Bond Modeling

Florian Thauunay, Florent Calvo, Edith Nicol, Gilles Ohanessian, Carine  
Clavaguéra

► **To cite this version:**

Florian Thauunay, Florent Calvo, Edith Nicol, Gilles Ohanessian, Carine Clavaguéra. Infrared Spectra of Deprotonated Dicarboxylic Acids: IRMPD Spectroscopy and Empirical Valence-Bond Modeling. *ChemPhysChem*, 2019, 20 (6), pp.803-814. 10.1002/cphc.201800947 . hal-02161622

**HAL Id: hal-02161622**

**<https://hal.science/hal-02161622>**

Submitted on 21 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Infrared Spectra of deprotonated dicarboxylic acids: IRMPD spectroscopy and empirical valence bond modeling

Florian Thauunay,<sup>‡</sup> Florent Calvo,<sup>§</sup> Edith Nicol,<sup>‡</sup>  
Gilles Ohanessian,<sup>‡</sup> Carine Clavaguéra<sup>¶</sup>

## Abstract

Experimental infrared multiple photon dissociation (IRMPD) spectra recorded for a series of deprotonated dicarboxylic acids,  $\text{HO}_2(\text{CH}_2)_n\text{CO}_2^-$  ( $n = 2-4$ ) and, are interpreted using a variety of computational methods. The broad bands centered near  $1600\text{ cm}^{-1}$  can be reproduced neither by static vibrational calculations based on quantum chemistry nor by a dynamical description of individual structures using the many-body polarizable AMOEBA force field, strongly suggesting that these molecules experience dynamical proton sharing between the two carboxylic ends. To confirm this assumption, AMOEBA was combined with a two-state empirical valence bond (EVB) model to allow for proton transfer in classical molecular dynamics simulations. Upon suitable parametrization based on *ab initio* reference data, the EVB-AMOEBA model satisfactorily reproduces the experimental

---

<sup>‡</sup>LCM, CNRS, Ecole Polytechnique, 91128 Palaiseau Cedex, France

<sup>§</sup>LIPhy, Université Grenoble Alpes and CNRS UMR 5588, 140 Avenue de la Physique, 38402 St Martin d'Hères, France

<sup>¶</sup>Laboratoire de Chimie Physique, CNRS - Université Paris Sud, Université Paris-Saclay, 15 avenue Jean Perrin, 91405, Orsay Cedex, France.

<sup>||</sup>corresponding author: gilles.ohanessian@polytechnique.edu, carine.clavaguera@u-psud.fr

infrared spectra, and the finite temperature dynamics reveals a significant amount of proton sharing in such systems.

## 1 Introduction

Proton transfer is an ubiquitous elementary reaction step in organic chemistry and biochemistry. Understanding its energetics and dynamical mechanisms has thus been a considerable endeavor for chemists over several decades. Proton transfer readily occurs in neutral, cationic or anionic species. The associated energetic profiles are quite diverse, especially since hydrogen bonds are often significantly stronger in charged systems relative to neutral ones. In the present work we focus on the proton dynamics possibly occurring between two carboxylate groups, namely  $-\text{CO}_2^- \cdots \text{H}^+ \cdots ^-\text{O}_2\text{C}-$  as an overall singly charged anion. Such a situation commonly occurs for carboxylic acids dissolved at sufficient concentration and mild pH conditions, and the above motif is then intermolecular. It can also occur in an intramolecular fashion when two carboxyl groups are present in a single molecule and are not too far from one another. This is most often realized in peptides and proteins carrying multiple aspartate (Asp) or glutamate (Glu) residues; it is already present in bare Asp and Glu. Proton sharing between two carboxylates has already been characterized by infrared (IR) spectroscopy for proteins in the condensed phase [1, 2]. Organic dicarboxylic acids are other common cases exhibiting proton transfer, particularly  $\text{HO}_2\text{C}-(\text{CH}_2)_4-\text{CO}_2\text{H}$  (adipic acid) that is used in a number of industrial processes such as the fabrication of Nylon 6,6 [3].

With the aim of understanding the intrinsic interactions between the two carboxylates and the bridging proton at a fundamental level, the present work focuses on isolated species in the gas phase. The coupling of mass spectrometric selectivity and sensitivity to infrared multiple photon dissociation (IRMPD) has provided a route to structural data for carefully mass-selected gaseous molecular ions [4–8]. IRMPD spectroscopy of deprotonated gaseous amino acids shows a qualitatively different behavior of Asp and Glu relative to all other cases studied [9]. The specificity of these dicarboxylic acids was

argued to originate from a proton that would be shared between the carboxylate groups rather than being covalently bound to one specific group and hydrogen-bound to the other group. This interpretation was supported by our own work on Asp and Glu [10]. IRMPD spectra suggest that similar phenomena could also occur in small peptides containing Asp or Glu residues [11].

Over the last few years there has been of surge of interest about the dynamics of proton sharing in deprotonated dicarboxylic acids of the simple organic family  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2\text{H}$  ( $n = 0-10$ ). When one of the carboxyl groups is deprotonated, the hydrogen bonding interaction between the two ends may be strong enough to offset the entropic penalty associated with folding into a ring. Both photoelectron spectra recorded for  $n = 1-10$  [12] and the infrared photodissociation (IRPD) spectra obtained for  $n = 0$  (Ref. [13]) and 10 (Ref. [14]) suggest that this is the case. The photoelectron spectra measured by Woo and coworkers [12] also indicate that adiabatic electron detachment energies, and thus the strength of these hydrogen bonds vary with size in a 0.2 eV range, suggesting that they are rather sensitive to the strain associated with cycle formation. This family of anionic molecules may thus display contrasted proton sharing propensities for different sizes  $n$ .

From the theoretical point of view, proton sharing is a highly anharmonic process causing the failure of static harmonic or even perturbative anharmonic approaches. The dynamics of proton sharing and its influence on vibrational spectroscopy have been investigated by classical and quantum dynamical simulations using reactive potential energy surfaces (PES's) and a more realistic account of anharmonicity [15-28].

First-principles molecular dynamics (MD) methods based on the Born-Oppenheimer (BOMD) or Car-Parrinello (CPMD) schemes have been used to model proton transfer in organic molecules in the gas phase, e.g. for protonated dialanine [18] or the  $(\text{HCO}_3^-)_2$  dimer [19]. In the latter case, nuclear quantum effects were also accounted for within the path-integral framework, owing to their possible importance for describing proton delocalization, especially at low temperatures. Even more rigorous quantum approaches such

as the multiconfiguration time-dependent Hartree (MCTDH) method were applied to model intramolecular proton transfer in malonaldehyde in relation with its vibrational signature [23, 24].

To extend the size and time scales inherent to quantum approaches, semi-empirical simulations have been performed to study proton transfer reactions in water clusters [29, 30], and QM/MM approaches have been used to model an excess proton in water and enzymatic activity [31, 32]. The need of a multistate description in such complex environments has led to extensions of the empirical valence-bond (EVB) model pioneered by Warshel and co-workers [33, 34] notably to address proton transport in water [15, 16]. A multistate EVB (MS-EVB) model developed by Voth *et al.* was used to assign the infrared signature of strong hydrogen bonds in small protonated water clusters based on classical MD simulations [17]. Nuclear quantum effects were further incorporated in MS-EVB simulations to explain how proton transfer is assisted by multiple weak hydrogen bonds [27, 28].

Focusing on proton sharing between two carboxylate groups in the gas phase, the low temperature IR spectrum of oxalate (deprotonated dicarboxylic acid with  $n = 0$ ) has recently become an interesting test bed for a number of theoretical approaches, as it displays a highly diffuse feature in the high frequency portion (ca. 2500–3200  $\text{cm}^{-1}$ ) corresponding to the stretching motion of the shared proton, even at 30 K [13]. While variational configuration interaction in an extended Fermi resonance formulation was found to account for the low frequency skeletal and OH bending motions, it was found necessary to rely on an adiabatic treatment of the variational wavefunction to describe the couplings of such motions with the high-frequency OH stretch [13]. In another theoretical study of the same spectrum [28], MD was used in a number of flavors to account for both parts of the spectrum. While semi-empirical MD using a density-functional based tight-binding approach did not reproduce the IR spectrum successfully, using the DFT B3LYP/6-31G PES yielded much better results, especially near room temperature. In an approach more closely related to the present work, MD was performed on a PES computed using a reactive force field allowing for proton transfer. The latter includes both stretching and

bending motions of the transferring proton, and takes variations of the other modes into account through a linear combination of energy functional forms for the two structures corresponding to the initial and final proton locations. Quantum nuclear effects were also considered at the level of ring-polymer MD. The very encouraging results from this study on this rather difficult case support the general validity of the MD framework relying on reactive force fields to provide a theoretically sound and computationally efficient approach to understand the dynamics of proton transfer. The force field-based EVB methodology used in the present study, although different in its treatment of bond transfer (vide infra), follows the same general aim.

We have recently extended the polarizable and anharmonic AMOEBA force field [35–37] to a two-state EVB description of proton sharing in deprotonated aspartic acid [10], successfully explaining its spectroscopic signature. In the present contribution we further address the case of deprotonated dicarboxylic acids using a similar methodology, but for a greater variety of molecules differing in the chain length  $n$ . While experimental spectra are available up to size  $n = 4$ , we have tested the transferability of the EVB-AMOEBA model to larger molecules by predicting their IR spectra. The overall strategy for designing the EVB-AMOEBA model relies on the determination of several ingredients, namely multipole expansions, polarizabilities and the specific EVB coupling term. The latter contains parameters that are adjusted to reproduce dedicated *ab initio* calculations.

The article is organized as follows. In Section 2 the experimental IRMPD spectra of several deprotonated dicarboxylic acids  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$  ( $n = 2-4$ ) are presented together with static spectra obtained by means of quantum chemistry methods. The failure of the static approach leads us to address the explicit dynamics of these systems by MD simulations using the AMOEBA force field in Sec. 3. Proton transfer is then enabled by implementing the EVB-AMOEBA method described in Sec. 4. The IR spectra obtained at finite temperature using this approach are presented and discussed in this section, focusing on adipic acid ( $n = 4$ ) and extending the methodology to other sizes for which experimental spectra are available or not. By varying the chain length  $n$ , different dynamical behaviors are found,

and limitations of the present approach in terms of transferability can be pointed out. Some concluding remarks are finally given in Sec. 5.

## 2 IRMPD experiments and static quantum chemistry spectra

### 2.1 Experiments

IRMPD spectra were recorded for the deprotonated dicarboxylic acids  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$  ( $n = 2-4$ ) at the CLIO (Centre Laser Infrarouge d’Orsay, France) facility [38] using a free-electron laser (FEL) coupled to a 7T FT-ICR mass spectrometer. Samples were dissolved in  $\text{H}_2\text{O}/\text{CH}_3\text{CN}$  50:50 mixtures, except for glutaric acid which was dissolved in pure  $\text{CH}_3\text{CN}$ , at typical concentrations of  $10^{-5}$ – $10^{-4}$  M. The ions were formed by ESI in standard conditions and transferred to the ICR cell. Mass-selected ions were stored in the cell and submitted to irradiation by IR photons during 0.3–2 s depending upon laser power, most often in the 0.5–1 W range. Absorption-induced fragmentation was monitored as a function of laser wavelength in the 1250–1800  $\text{cm}^{-1}$  range. All ions were found to produce  $m/z$ -44 fragments resulting from  $\text{CO}_2$  elimination. Elimination of water was also observed in all cases. The joint elimination of both  $\text{H}_2\text{O}$  and  $\text{CO}_2$  was also observed for adipic acid. It is expected that both fragmentations are made easier through the formation of a cyclic product, which takes place without entropic penalty if the parent ions are folded with interactions between the two carboxylate/carboxylic ends.

The IR spectra were obtained by representing the fragmentation efficiency,  $F_{\text{eff}}$ , as a function of wavelength.  $F_{\text{eff}}$  is defined as  $F_{\text{eff}} = -\ln[I_p/(I_p + \sum I_{\text{frag}})]$ , where  $I_p$  and  $I_{\text{frag}}$  are the parent and fragment ion intensities, respectively. The experimentally recorded IRMPD spectra are shown in Fig. 1.

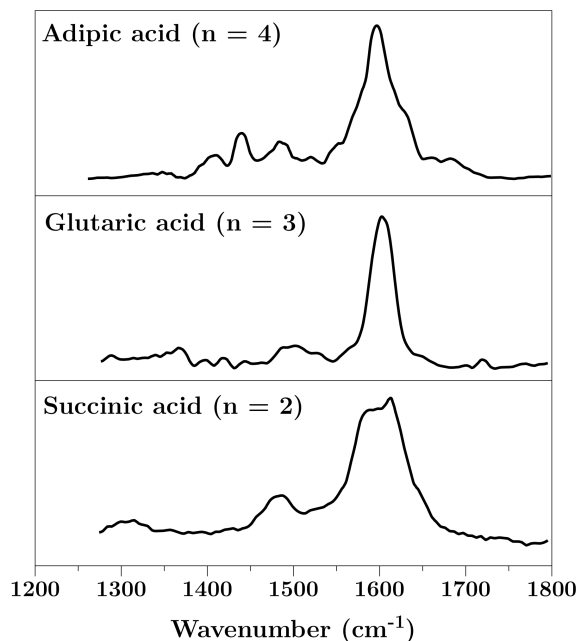


Figure 1: IRMPD spectra of several deprotonated dicarboxylic acids  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$ ,  $n = 2-4$ , recorded at room temperature at the CLIO setup.

For succinic ( $n = 2$ ), glutaric ( $n = 3$ ) and adipic ( $n = 4$ ) acids, the spectra do not display the superposition of typical bands expected for a carboxylic acid and a carboxylate group [39], even if frequency shifts due to hydrogen bonding occur [40]. The large band centered near  $1600 \text{ cm}^{-1}$ , with rather limited absorption beyond  $1650 \text{ cm}^{-1}$ , is an unusual feature. It has been assigned earlier to proton sharing in aspartate and glutamate [9, 10], whereby both  $\text{O} \cdots \text{H}$  interactions are intermediate between a covalent bond and a hydrogen bond. However, this band is not as broad here as observed previously for aspartate [10] and it is assumed that the symmetry of the present dicarboxylic acids favors stronger proton sharing. For succinic acid ( $n = 2$ ), the parent ion signal was not as stable as desirable, and the fragment mass to charge ratio ( $m/z=73$ ) lies near the low-end detection limit of the instrument. Small intensity variations, such as the shoulder near  $1600 \text{ cm}^{-1}$  (see Fig. 1) are thus not deemed significant.



## 2.2 Quantum chemistry calculations

The previous assumption is supported by quantum chemistry calculations that we first carried out at the MP2/cc-pVDZ level to identify the low-energy conformers. All quantum chemistry calculations were performed with the Gaussian 09 package [41]. The optimized folded conformations obtained for the smallest molecules ( $n < 4$ ) are depicted in Fig. 2. The shared proton is aligned with the two proton-bearing oxygens for both molecules.

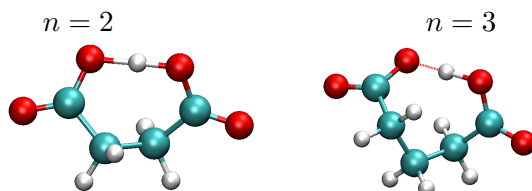


Figure 2: Bent conformations for  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$  with  $n = 2$  (left) and 3 (right), optimized at the MP2/cc-pVTZ level.

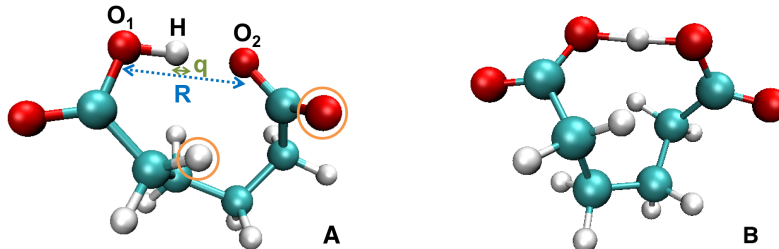


Figure 3: The two most stable isoenergetic conformations of  $\text{HO}_2\text{C}-(\text{CH}_2)_4-\text{CO}_2^-$  optimized at the MP2/cc-pVDZ level. (Left) conformer A with one  $\text{HCH}\cdots\text{OCO}$  hydrogen bond (orange circles); (right) conformer B without the hydrogen bond but with fully shared proton.  $R$  denotes the distance between  $\text{O}_1$  and  $\text{O}_2$ ,  $q$  the distance between the central proton and the center of the line between  $\text{O}_1$  and  $\text{O}_2$ .

For adipic acid, two low-energy conformations bearing a carboxylate-carboxylic acid hydrogen bond could be identified at the MP2/cc-pVDZ level of theory, they are shown in Fig. 3. Conformer A presents one intramolecular hydrogen bond of  $\text{HCH}\cdots\text{OCO}$  type while conformer B does not. At this

Table 1: Internal coordinates (distances in Å) involved in intramolecular proton transfer for the two conformers of adipic acid depicted in Fig. 3 at the MP2 and DFT/M06 (only for conformer A because of the higher energy of conformer B levels using different basis sets.

	conformer A				conformer B		
	M06	MP2			MP2		
	b1	b1	b2	b3	b1	b2	b3
$R_{OO}$	2.617	2.578	2.565	2.546	2.432	2.408	2.403
$R_{O\dots H}$	1.609	1.564	1.550	1.531	1.325	1.241	1.239
$r_{OH}$	1.011	1.018	1.021	1.021	1.109	1.167	1.164

b1: 6-31++G(d,p), b2: cc-pVDZ, b3: cc-pVTZ

level of theory, the two conformers are iso-energetic. Conformer B may favor proton sharing owing to fewer geometric constraints. However, when treated at the level of density functional theory (DFT) with the M06 functional and the 6-31++G(d,p) basis set, conformer B now lies more than 4 kJ/mol above A. Conversely, it becomes more stable than conformer A (by 2.2 kJ/mol) when the larger basis set cc-pVTZ is associated to the MP2 level of theory. These results highlight the difficulty in drawing firm and reliable conclusions from such static quantum chemistry calculations. The internal coordinates involved in proton transfer, which will be used for force field parametrization below, are provided in Table 1. For conformer A, the structures are similar whatever the quantum chemistry level. For conformer B, increasing the size of the basis set entails a shorter distance between  $O_1$  and  $O_2$  along with an elongation of the O-H covalent bond.

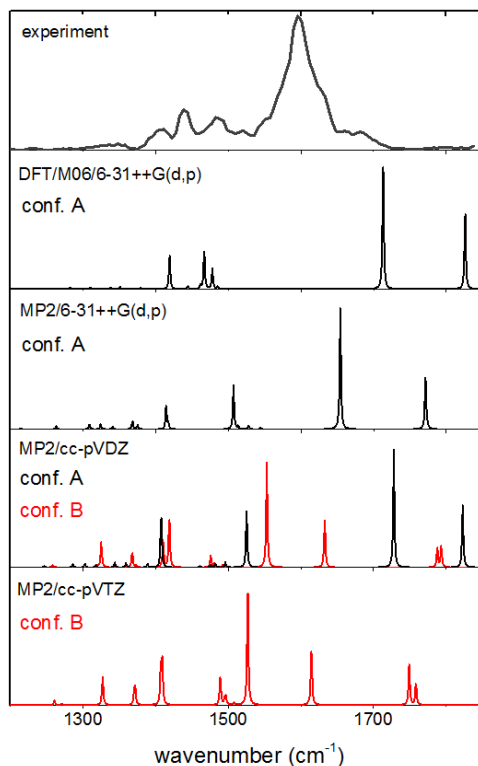


Figure 4: IR spectra of adipic acid. Upper panel: experimentally measured IRMPD spectrum. The lower panels show the harmonic spectra obtained from quantum chemistry calculations for conformers A (black line) and B (red line) using the DFT/M06/6-31++G(d,p), MP2/6-31++G(d,p), MP2/cc-pVDZ, and MP2/cc-pVTZ methods. No scaling factor was applied to the calculated spectra. The calculated band intensities were convoluted by assuming a Lorentzian profile with a full width at half maximum of  $1 \text{ cm}^{-1}$ .

Harmonic static spectra of the most stable conformers of adipic acid obtained at the DFT/M06 and MP2 levels and using various basis sets are represented in Fig. 4. At the MP2/cc-pVDZ level, two spectra are compared because conformers A and B are isoenergetic. The computed spectrum appears to depend strongly on the level of theory due to differences in the geometries. However, it is clear that none of them provides a satisfactory comparison with experiments, and the use of scaling factors or the

computation of anharmonic frequencies do not reduce the discrepancy significantly. For conformer A, the band located at 1825, 1771, or 1823  $\text{cm}^{-1}$  at the M06/6-31++G(d,p), MP2/6-31++G(d,p), and MP2/cc-pVDZ levels, respectively, is attributed to a combination of O=COH stretching and COH bending modes of the carboxylic acid. The second band in the blue region of the spectrum, located at 1712, 1654 and 1728  $\text{cm}^{-1}$  at the aforementioned levels, corresponds to the carboxylate O=C=O asymmetric stretching mode. Bands appearing in the 1300–1500  $\text{cm}^{-1}$  range arise mainly from the carboxylate O=C=O symmetric stretching mode in the red part, and the COH bending mode in the blue part, both associated to  $\text{CH}_2$  modes. The latter can be either HCH bending or  $\text{CH}_2$  wagging modes depending on the band as well as the specific quantum chemistry method.

For conformer B, the shared proton signature is visible with a double band at 1787 and 1793  $\text{cm}^{-1}$  at the MP2/cc-pVDZ level, and at 1749 and 1758  $\text{cm}^{-1}$  at the MP2/cc-pVTZ level. From this MP2 normal mode analysis, it is attributed to a combination of the two O=COH stretching modes. Furthermore, the bands at 1552 and 1633  $\text{cm}^{-1}$ , or at 1527 and 1614  $\text{cm}^{-1}$  for the two previous levels are assigned to the COH bending mode. Despite these signatures, using an equilibrium geometry associated to vibrational frequencies computed in the double harmonic approximation is not sufficient to reproduce the experimental spectrum in any satisfactory way, and dynamical effects at finite temperature should at least be included.

### 3 Dynamical modeling using the AMOEBA force field

#### 3.1 The AMOEBA polarizable force field

The failure of quantum chemistry methods in the static approximation encourages us to pursue the alternative approach of dynamical spectra based on an accurate force field. In this work, the AMOEBA model was chosen owing to its realistic description of nonbonded intermolecular terms through the inclusion of multipole distributions in the electrostatic treatment, to-

gether with an explicit polarization term in which the induced dipoles are treated self-consistently. AMOEBA has been successfully used to model the structural and energetic properties of biomolecules in the gas and condensed phases [35, 42, 43], as well as thermodynamic properties of solvated ions or protein-ligand complexes [44–47]. Its application to IR spectroscopy is more recent but allowed anharmonic and temperature effects to be correctly accounted for in gas phase molecules [48–50].

Several ingredients of the AMOEBA model require parametrization, in particular the atomic charges, dipoles and quadrupoles of the multipolar expansion. In the original spirit of this *ab initio* force field, these parameters were derived from dedicated quantum chemistry calculations, using the distributed multipole analysis (DMA) developed by Stone [51, 52]. It was notably shown that such a high quality electrostatic treatment enables an accurate description of hydrogen bonds both in strength and directionality [42, 48, 53]. Within AMOEBA, van der Waals (vdW) interactions are described by a buffered 14-7 pair potential [54] that reproduces *ab initio* results satisfactorily both in the gas and liquid phases:

$$E_{\text{vdW}}(R_{ij}) = \varepsilon_{ij} \left( \frac{1.07}{\rho_{ij} + 0.07} \right)^7 \left( \frac{1.12}{\rho_{ij}^7 + 0.12} - 2 \right) \quad (1)$$

with  $\rho_{ij} = R_{ij}/R_{ij}^0$ ,  $R_{ij}$  being the distance between atoms  $i$  and  $j$ ,  $R_{ij}^0$  the equilibrium distance, and  $\varepsilon_{ij}$  the well depth. In heterogeneous atom pairs, the following mixing rules are applied for these coefficients:

$$R_{ij}^0 = \frac{(R_{ii}^0)^3 + (R_{jj}^0)^3}{(R_{ii}^0)^2 + (R_{jj}^0)^2} \quad (2)$$

and

$$\varepsilon_{ij} = \frac{4\varepsilon_{ii}\varepsilon_{jj}}{(\sqrt{\varepsilon_{ii}} + \sqrt{\varepsilon_{jj}})^2} \quad (3)$$

Bonding terms for stretchings, angle bendings, torsions and stretch-bend couplings have the same functional forms than those used by the MM3 force field and notably include anharmonic components [55] that are particularly

important for vibrational spectroscopy.

Parameters were taken from the AMOEBA<sub>bio</sub>-09 set available in the Tinker 7 package [56], except those specifically developed for the present study. The additional parametrization was carried out for adipic acid, for which atomic multipolar distributions were obtained from the MP2/aug-cc-pVTZ density using the DMA analysis, followed by an adjustment of atomic dipoles and quadrupoles on the electrostatic potential while keeping the charges fixed. Parameters were adjusted on conformer A of adipic acid alone, leaving conformer B away from this analysis. This choice of not including conformer B is justified by the very specific nature of its strong hydrogen bond, making parameters for this structure expectedly less transferable to other conformations of the molecule, the vdW parameters being likely overestimated in B to describe the strong bond [57]. Furthermore, conformer B has a proton shared nearly evenly by the two carboxylate ends, preventing a clear distinction between the carboxylate and carboxylic groups that is necessary within the nonreactive AMOEBA force field.

### 3.2 Infrared spectra calculations

IR spectra for dicarboxylic acids were modeled using the AMOEBA force field, computing the spectrum from the Fourier transform of the dipole moment time autocorrelation function (DACF) as obtained from the MD trajectories. In the nonreactive approach pursued so far, the two interactions involving the shared proton are considered differently, the proton located between two carboxylate oxygens forming a covalent bond with one oxygen and a hydrogen bond with the other. These different interactions are taken into account as intramolecular and non-bonded terms, respectively. After optimizing the AMOEBA parameters on conformer A, the characteristic distances with the optimized AMOEBA geometry are  $R_{OO} = 2.73 \text{ \AA}$  and  $r_{OH} = 0.99 \text{ \AA}$ , whose comparison with the quantum chemistry data of Table 1 shows that the hydrogen bond is somewhat weaker with AMOEBA.

To calculate the finite temperature IR spectra with the AMOEBA force field and dipole moment surfaces, MD trajectories were first propagated at

constant temperature using Nosé-Hoover thermostats. These configurations and momenta were further propagated in absence of thermostat for 50 ps, the total dipole moment vector  $\vec{\mu}(t)$  being saved every time step of 0.1 fs. The final IR spectrum is averaged over the contributions from 10 independent trajectories. The spectrum obtained at 300 K over a total simulation time of 500 ps is represented in Fig. 5.

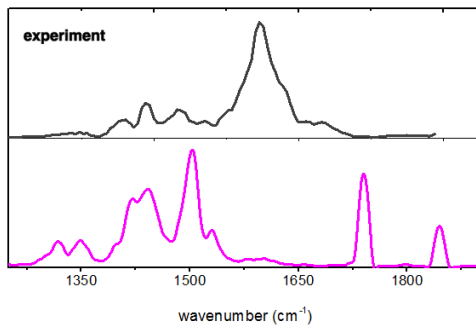


Figure 5: Non-reactive DACF-AMOEBA IR absorption spectrum calculated for adipic acid using MD simulations at 300 K (pink), compared with the IRMPD experimental spectrum (black).

As expected, a broadening of the bands is obtained that reflects anharmonicity at finite temperature, but their positions are similar to the static M06/6-31++G(d,p) frequencies. The non-reactive AMOEBA method is thus not able to provide a satisfactory description of the proton dynamics between the two carboxylate sites.

## 4 Effects of proton sharing: an EVB-AMOEBA model

### 4.1 Accounting for proton migration with AMOEBA

#### The EVB-AMOEBA model

Proton sharing is now accounted for in the framework of a two-state empirical valence-bond model built on AMOEBA for the nonreactive part. In the EVB methodology, the ground state energy of the dicarboxylic acid molecule

is written as a linear combination of two valence diabatic states  $\psi_1$  and  $\psi_2$  for which the proton is covalently bonded to either of the two oxygen atoms involved in the exchange, the wavefunction  $\Psi$  being a linear combination of these two states,

$$\Psi = g_1\psi_1 + g_2\psi_2, \quad (4)$$

where  $g_1$  and  $g_2$  are real numbers such that  $g_1^2 + g_2^2 = 1$ . The diabatic state potential energies  $V_1$  and  $V_2$  are described by the AMOEBA force field, in which the two oxygen atoms play symmetric roles depending on the state number. The EVB Hamiltonian can be written as a  $2 \times 2$ -matrix with a coupling term  $V_{12}$  allowing for proton exchange between the two diabatic states. The lowest eigenvalue of the Hamiltonian is explicitly given by:

$$\begin{aligned} V_{\text{EVB}} &= \frac{1}{2} \left[ V_1 + V_2 - \sqrt{(V_1 - V_2)^2 + 4V_{12}^2} \right] \\ &= \frac{1}{2}(V_1 + V_2 - D), \end{aligned} \quad (5)$$

with  $D = [(V_1 - V_2)^2 + 4V_{12}^2]^{1/2}$

and the eigenvector components  $(g_1, g_2)$  read

$$\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} [1 + (V_{\text{EVB}} - V_1)^2/V_{12}^2]^{-1/2} \\ (1 - g_1^2)^{1/2} \end{pmatrix}. \quad (6)$$

MD trajectories on the ground state EVB surface are propagated by solving the equations of motion as usual, the forces being determined owing to the Hellmann-Feynman theorem:

$$\frac{\partial V_{\text{EVB}}}{\partial x} = g_1^2 \frac{\partial V_1}{\partial x} + g_2^2 \frac{\partial V_2}{\partial x} + 2g_1g_2 \frac{\partial V_{12}}{\partial x}. \quad (7)$$

The dipole moment vector  $\vec{\mu}$  required to compute IR spectra is obtained by linear combination of its values  $\vec{\mu}_1$  and  $\vec{\mu}_2$  for the diabatic states:

$$\vec{\mu}_{\text{EVB}} = g_1^2 \vec{\mu}_1 + g_2^2 \vec{\mu}_2. \quad (8)$$



Proton migration between the two carboxylate groups is enabled by the off-diagonal coupling term  $V_{12}$ , which mixes the two diabatic states. For this term we chose one well-established type of relatively simple expression involving only the distance  $R_{OO}$  between the two protonation sites [15, 58, 59] and the distance  $q$  between the proton and the midpoint separating these sites [58, 59] (see Fig. 3),

$$q = \left\| \vec{r}_{\text{H}} - \frac{1}{2}(\vec{R}_{\text{O}_1} + \vec{R}_{\text{O}_2}) \right\|. \quad (9)$$

The form for the coupling potential was adapted from Vuilleumier and Borgis [58] with a smoother decrease with  $q$  than in this original work where it was also exponential:

$$V_{12} = A \frac{\exp(-\alpha R)}{1 + \gamma q^2}, \quad (10)$$

$A$ ,  $\alpha$  and  $\gamma$  being three parameters that need to be fitted appropriately. In addition, much better adjustment to quantum chemistry data was achieved by introducing some dependence of  $A$  with  $R$  as

$$A(R) = A_0 + A_1 R + A_2 R^2 \quad (11)$$

with three parameters  $A_0$ ,  $A_1$  and  $A_2$  for  $A$ .

This EVB model was implemented in Tinker software [56]. To avoid I/O limitations, data are stored in memory at the beginning of the simulation to compute  $V_1$  and  $V_2$  energies and the corresponding forces. A supplementary cost is due to the parameter attribution on each atom. As two force fields are used for the two states, the order of the two isomers is reversed at each step of the simulation to limit the computational cost in parameter attribution. For the size of the systems under study here, an EVB-AMOEBA simulation is approximately 1.6 times longer than the standard corresponding AMOEBA simulation.

### Adjustment of the EVB-AMOEBA parameters

For the present molecules the two diabatic states are symmetric and can be described by the same force field, except for some topological reassignment of atoms or bonds. This high symmetry reduces the overall number of parameters and this facilitates parametrization especially on the atomic multipole set. Aiming to produce a transferable EVB potential for other dicarboxylic acids, we choose again adipic acid for parametrizing the coupling terms, as this molecule is of intermediate length, offering greater flexibility and in particular other hydrogen bonding possibilities like  $\text{HCH}\cdots\text{OCO}$  involving  $\text{CH}_2$  group of the carbon chain [12]. In addition, the IRMPD spectrum for adipic acid is well resolved with a clear proton sharing signature as seen in Fig. 1.

Parametrization of the coupling term requires sampling of the proton migration across the two carboxylate oxygens. Consistently with the non-reactive part of the potential, this was achieved by performing *ab initio* calculations at the MP2/cc-pVTZ level of theory. A 2-dimensional grid was thus constructed by varying  $R$  and  $q$ , using appropriate ranges for these quantities as  $2.4 \text{ \AA} < R < 2.9 \text{ \AA}$  and  $0.8 \text{ \AA} < q < 2.1 \text{ \AA}$  scanned by steps of  $0.1 \text{ \AA}$  for both. The  $\text{O}_2\widehat{\text{O}}_1\text{H}$  angle would ideally be kept to 0 to ensure linearity, however this was not possible with the optimizer used and this angle had to be kept to  $2^\circ$ . At each grid point, the remaining degrees of freedom were fully relaxed, yielding a potential energy surface (PES) that is symmetric with respect to proton exchange, as shown in Fig. 6.

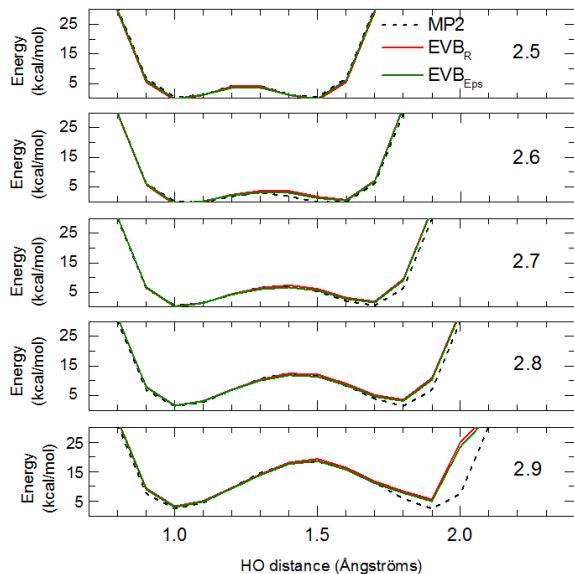


Figure 6: Potential energy surface for proton migration across the two carboxylate oxygens of adipic acid, as obtained from MP2/cc-pVTZ calculations and compared with two EVB-AMOEBA model calculations denoted as  $\text{EVB}_R$  and  $\text{EVB}_{\text{Eps}}$  and defined in the main text. The energy is represented as a function of the H-O distance and for different values of the interoxygen distance  $R$  (both in  $\text{\AA}$ ).

Along the  $R$  coordinate, the energies of both local minima increase in average by  $0.75 \text{ kcal.mol}^{-1}$  by each step of  $0.1 \text{ \AA}$ . In this range of  $R$ , the partially constrained relaxed geometries obtained from the MP2 calculations are similar to conformer A, with a  $\text{HCH} \cdots \text{OCO}$  hydrogen bond forming a 8-membered ring owing to the rotation of one carboxylate group. At  $R = 2.4 \text{ \AA}$  and below, a competition between conformers of types A and B occurs, and the PES calculation at the MP2 level becomes difficult. Therefore we did not include quantum chemistry data in our training set for such low distances.

In adjusting the parameters of the EVB-AMOEBA model, we found impossible to reproduce the MP2 data by only varying the coupling parameters. Some AMOEBA parameters of the diabatic states involved in proton transfer had to be considered as well because the force field was not able to deal with very short interactions. Therefore we decided to modify the vdW parame-

Table 2: Internal coordinates (in Å) of adipic acid specific to the proton transfer at the MP2/cc-pVDZ level and with EVB-AMOEBAs for different vdW parameter sets in which the equilibrium distances (EVB<sub>R</sub>) or well depths (EVB<sub>Eps</sub>) involved in hydrogen bonds have been modified.

	MP2	EVB <sub>R</sub>	EVB <sub>Eps</sub>
$R_{\text{OO}}$	2.565	2.551	2.556
$R_{\text{O}\cdots\text{H}}$	1.550	1.555	1.559
$r_{\text{OH}}$	1.021	0.996	0.998

ters of the atoms involved in the proton transfer. More specifically, two sets of EVB-AMOEBAs were produced by allowing changes either in the equilibrium distances  $R_{\text{HH}}^0$ ,  $R_{\text{OO}}^0$  or in the well depths  $\varepsilon_{\text{HH}}$  and  $\varepsilon_{\text{OO}}$ . In practice, the changes amounted to a 9% decrease in the equilibrium distances and to a 44% decrease in the well depths. For both vdW sets, that we respectively denote as EVB<sub>R</sub> and EVB<sub>Eps</sub>, a single set of EVB coupling parameters could be optimized to yield  $A(R) = 13345 - 10240R + 2000R^2$  kcal/mol (with  $R$  expressed in Å),  $\alpha = 0.745 \text{ \AA}^{-1}$ , and  $\gamma = 5.35 \text{ \AA}^{-2}$ . The potential energy surfaces predicted by our two EVB-AMOEBAs, superimposed in Fig. 6, show very satisfactory variations relative to the MP2 data.

The main geometric features related to proton migration at the equilibrium structure of conformer A, given in Table 2, also indicate the good accuracy of the presently parametrized reactive potentials. In addition, the error on the relative energies of the minima is less than 1 kcal/mol.

For comparison, the original AMOEBAs potential employed for the diabatic states predicts a much longer hydrogen bond than when reactivity is incorporated, with  $R_{\text{OO}} = 2.73 \text{ \AA}$  to be compared with  $2.57 \text{ \AA}$  in the MP2 reference calculation.

## 4.2 Results for adipic acid

MD simulations based on the reactive EVB-AMOEBAs force field were first performed at thermal equilibrium in the NVT ensemble. Fig. 7 shows the time evolution of the EVB weights  $g_1$  and  $g_2$ , the distances  $R$  and  $q$  for a typical 50 ps trajectory at 100 K, all properties being short-time averaged

over 50 fs windows to improve clarity.

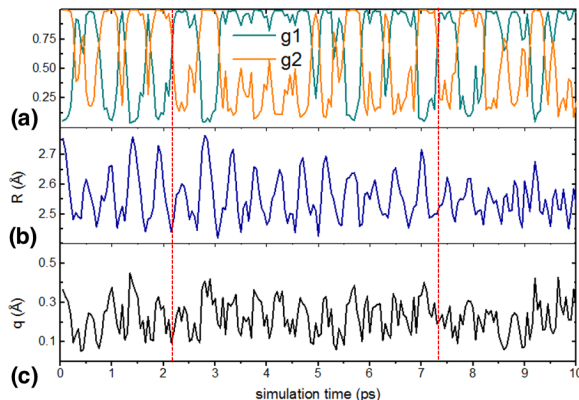


Figure 7: Time-resolved properties obtained for a typical EVB-AMOEBA MD simulation of adipic acid at 100 K. (a) EVB weights  $g_1$  (green) and  $g_2$  (orange); (b) interoxygen distance  $R$ ; (c) Distance  $q$  of proton to oxygen midpoint. All distance are in Å, and properties were short-time averaged over 50 fs to improve quality. Two vertical dashed red lines highlight examples of proton exchange events.

Proton exchange is proven by the variations shown by  $g_1$  and  $g_2$ , which both cover the entire range 0–1. The shared position ( $g_1 = g_2 = 1/\sqrt{2}$ ) is never seen over an extensive time period, the proton spending far longer time on either of the two carboxylate sites. Proton transfer occurs at least once every 500 fs in average, and transfers occur back and forth with an equal probability, which is consistent with the symmetry of the diabatic states. Most proton transfer events appear for small values of  $R$  in the range  $2.45 \text{ \AA} \leq R \leq 2.50 \text{ \AA}$ , which corresponds to distances a little shorter than the MP2 equilibrium value. Variations of the distance  $q$  mimic those of  $R$ , with smaller values corresponding to the most symmetrical structures that allow also smaller values of  $R$ .

These results indicate that proton sharing occurs by means of frequent transfer between the two carboxylate groups, rather than proton trapping at intermediate distances between the two sites. Although this sharing mode is dynamical, it is made possible by the rather compact structure sampled,

as illustrated by the statistics described below.

The distributions of the distances  $R$  and  $q$  accumulated over 10 such trajectories are given in Fig. 8 at the two temperatures of 50 and 300 K. We further compare here the predictions of the two EVB models with parameter sets  $\text{EVB}_R$  and  $\text{EVB}_{\text{Eps}}$ .

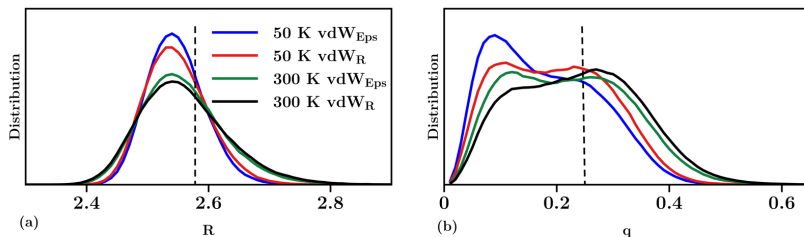


Figure 8: Equilibrium distributions of distances  $R$  (a) and  $q$  (b) obtained from EVB-AMOEBA MD trajectories of adipic acid at 50 and 300 K, for the two parameter sets  $\text{EVB}_R$  and  $\text{EVB}_{\text{Eps}}$ . In each panel, the vertical black dashed lines locate the corresponding values in the MP2 equilibrium structure.

The distribution of interoxygen distance is barely affected by the choice of the vdW parametrization in the EVB model,  $R$  being spread out over and below the MP2 equilibrium value. As expected, the distributions become broader with increasing temperature for both quantities. In contrast to  $R$ , the distance of the proton to the oxygen midpoint appears more sensitive to the EVB model details, and notably favors trapped proton sharing ( $q < 0.1 \text{ \AA}$ ) at low temperature.

IR absorption spectra were then computed with EVB-AMOEBA from microcanonical MD simulations initiated from equilibrated configurations obtained from NVT simulations. The spectra obtained with the two EVB parameter sets at 50, 100, 200 and 300 K are shown in Fig. 9 in comparison with the experimental IRMPD spectrum.

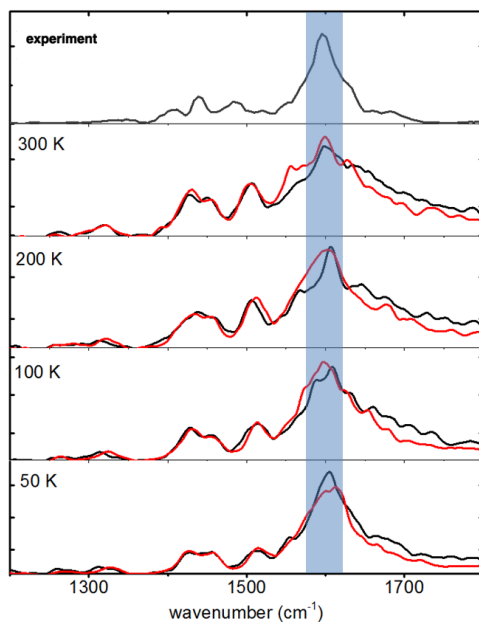


Figure 9: IR absorption spectra obtained with EVB-AMOEBA models for adipic acid as a function of temperature in the range 50–300 K, for the two vdW parameter sets  $\text{EVB}_R$  (black lines) and  $\text{EVB}_{\text{Eps}}$  (red lines). The experimental IRMPD spectrum is shown in the upper panel. The vertical blue box is a visual guide to compare the experimental intensity maximum with the computed spectra.

Consistently with the small geometry fluctuations, the finite temperature spectra do not vary significantly between 50 and 300 K. One broad band centered around  $1600\text{ cm}^{-1}$  is present in all cases, whose blue part becomes wider with increasing temperature. According to Fig. 8 this is correlated with a lesser importance of proton sharing with more frequent occurrence of localized proton with a C=O stretching band on the blue side. The frequency maximum is well reproduced without showing any significant shift with temperature in sharp contrast to the results obtained from simulation of a single diabatic state (Fig. 5). Our computations agree very well with experiments with three bands in the  $1400\text{--}1500\text{ cm}^{-1}$  range and one low-intensity band between  $1300\text{--}1350\text{ cm}^{-1}$ .

### 4.3 Extension to other dicarboxylic acids $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$ , ( $n = 2, 3$ and $6$ )

Having shown the performance of the EVB-AMOEBAs approach on the specific case of adipic acid on which the model was fitted, we now challenge its transferability on the other dicarboxylic acids of the same family. The EVB-AMOEBAs simulations were thus performed without altering or reparametrizing the EVB coupling term  $V_{12}$ . The specific parameters for the diabatic AMOEBAs part required the extraction of new multipoles and vdW parameters to reproduce the MP2 geometries.

#### 4.3.1 Glutaric acid ( $n = 3$ )

The IRMPD spectrum measured for glutaric acid  $\text{HO}_2\text{C}-(\text{CH}_2)_3-\text{CO}_2^-$  shows an intense band at  $1600\text{ cm}^{-1}$ , very similar to the band of adipic acid (Fig. 1). This band can be attributed to proton transfer in the absence of typical bands associated with hydrogen-bound carboxylic acid and carboxylate. Another band can be identified at ca.  $1500\text{ cm}^{-1}$ . Because of the small signal-to-noise ratio in the  $1350\text{--}1500\text{ cm}^{-1}$  range, it is difficult to discuss possible differences between the spectra for  $n = 3$  and  $n = 4$  in this region, however the slightly more strained conformation for  $n = 3$  is expected to lead to some frequency changes.

For this molecule, it turned out to be impossible to fit the geometric parameters  $R_{\text{OO}}$ ,  $r_{\text{OH}}$  and  $r_{\text{O}\dots\text{H}}$  simultaneously with a good accuracy for the bent MP2 reference geometry shown in Fig. 2 that we refer to as  $\text{C}_5$  since glutaric acid has 5 carbon atoms. These distances are probably out of the transferability range of the model with too large of a role then given to the coupling potential.

A slightly different strategy was thus used, using a reference geometry similar to that obtained for adipic acid, thus named  $\text{C}_6$ . This was achieved by performing a constrained optimization of glutaric acid at the MP2/cc-pVTZ level by fixing some geometry parameters around the proton, namely  $R = 2.565\text{ \AA}$ ,  $r_{\text{OH}} = 1.021\text{ \AA}$  and  $\widehat{\text{O}_2\text{O}_1\text{H}} = 2^\circ$ . Multipoles and vdW parameters involving the proton were then extracted in the same conditions



Table 3: Internal coordinates (in Å) of glutaric acid specific to proton transfer at the MP2 and AMOEBA levels for different vdW parameter modifications for O and H atoms involved in the hydrogen bond.

	MP2	EVB <sub>R</sub>	EVB <sub>Eps</sub>
$R_{OO}$	2.565	2.568	2.564
$R_{O\dots H}$	1.550	1.581	1.574
$r_{OH}$	1.021	0.992	0.994

as for the C<sub>5</sub> geometry. As was achieved for adipic acid, vdW parameters were further corrected for the hydrogen bond with a decrease of 20.5% for both  $R_{HH}^0$  and  $R_{OO}^0$ , or a decrease of 76.5% for both  $\epsilon_{HH}$  and  $\epsilon_{OO}$ , yielding two EVB-AMOEBA parameter sets again referred to as EVB<sub>R</sub> and EVB<sub>Eps</sub>. As seen in Table 3, both parameter sets produce new equilibrium geometries that are in good agreement with the constrained MP2 structure. Using these new parameter sets, EVB-AMOEBA MD simulations were performed at 100 and 300 K thermal equilibrium, and the distributions obtained for  $R$  and  $q$  are depicted in Fig. 10.

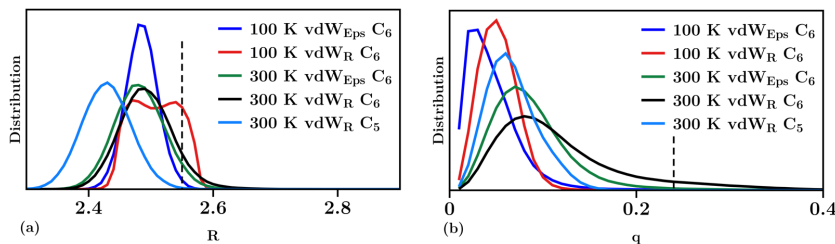


Figure 10: Equilibrium distributions of distances  $R$  (a) and  $q$  (b) obtained from EVB-AMOEBA MD trajectories of glutaric acid at 100 and 300 K, for the two parameter sets EVB<sub>R</sub> and EVB<sub>Eps</sub> fitted on the C<sub>6</sub> geometry, the results with the parameter set EVB<sub>R</sub> fitted on the C<sub>5</sub> geometry being shown as well. Vertical dashed lines highlight the equilibrium values of  $R$  and  $q$  in the MP2 equilibrium geometries for adipic acid. C<sub>5</sub> and C<sub>6</sub> notations correspond to MP2 reference geometries that have been used to adjust the vdW parameters (see text for details).

At 300 K,  $R$  lies between 2.4 and 2.6 Å with a maximum at 2.48 Å. For this system, the two vdW corrections are not as equivalent as they were in the case of adipic acid, and correcting the well-depth parameters favors proton sharing. If the equilibrium distances are altered, the steric hindrance of the CH<sub>2</sub> group allows proton localization on one carboxylate site. This explains the slightly broader  $R$  distribution at 300 K, while being bimodal at 100 K with some proton sharing at short distances ( $R \sim 2.48$  Å) concomitant with a lower steric effect of the CH<sub>2</sub> group and a more localized proton at longer distances ( $R \sim 2.55$  Å) interacting more strongly with CH<sub>2</sub>.

IR absorption spectra obtained with the EVB-AMOEBA models are shown in Fig. 11 at 100, 200, and 300 K.

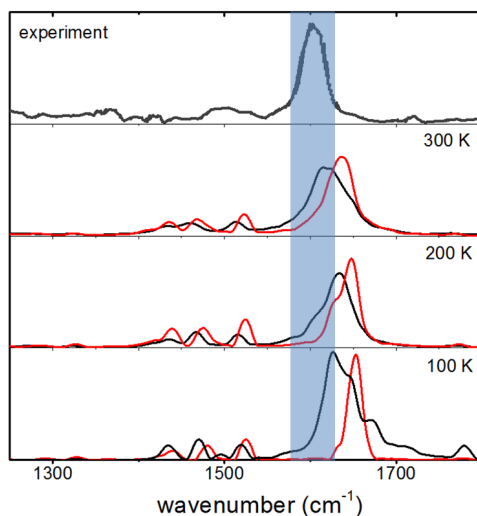


Figure 11: IR absorption spectra obtained with EVB-AMOEBA models for glutaric acid as a function of temperature in the range 100–300 K, for the two vdW parameter sets EVB<sub>R</sub> (black lines) and EVB<sub>Eps</sub> (red lines). The IRMPD experimental spectrum is shown in the upper panel. The vertical blue box is a visual guide to compare the experimental intensity maximum with the computed spectra.

The proton transfer signature is visible on the IR spectra, which also depend more sensitively on the vdW model details. In particular, the secondary hydrogen bonds with the CH<sub>2</sub> groups give rise to a redshift of the

main band and a stronger proton mobility with the  $\text{EVB}_R$  model, which agrees better with the experimental spectrum. The best agreement is obtained at 300 K with a residual blueshift of only  $15 \text{ cm}^{-1}$  relative to the measured band.

### 4.3.2 Succinic acid ( $n = 2$ )

At the price of having to readjust the vdW corrections, the  $\text{EVB-AMOEB}$ A model could be successfully transferred from adipic acid to glutaric acid. These changes notably allowed the proton not being stuck at equilibrium, with large  $R$  distances permitted by the presence of the  $\text{CH}_2$  group. In succinic acid, the equilibrium geometry has a very stable 7-membered ring (see Fig. 2). Performing  $\text{EVB-AMOEB}$ A simulations with the  $\text{EVB}_R$  parameter set fitted for glutaric acid preserves this ring very tightly and a strongly shared proton. The equilibrium distances, located in the range of  $2.40\text{--}2.45 \text{ \AA}$ , lie away from the fitting range of the model. We have therefore chosen not to consider this case further for IR spectroscopy.

### 4.3.3 Suberic acid ( $n = 6$ )

We finally consider suberic acid, for which no experimental spectrum is available yet. The equilibrium structure of isolated  $\text{HO}_2\text{C}-(\text{CH}_2)_6\text{-CO}_2^-$  presents similar constraints as adipic acid, the  $\text{MP2/cc-pVTZ}$  geometry being depicted in Fig. 12(a).

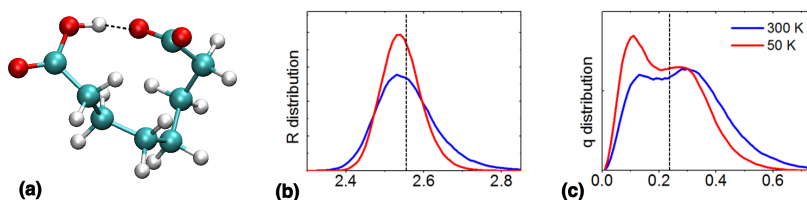


Figure 12: (a)  $\text{MP2/cc-pVTZ}$  optimized geometry of  $\text{HO}_2\text{C}-(\text{CH}_2)_6\text{-CO}_2^-$  and equilibrium distributions predicted by the  $\text{EVB-AMOEB}$ A model for the  $R$  (b) and  $q$  (c) distances at 50 and 300 K. Vertical dashed black lines highlight the values for both quantities at the  $\text{MP2}$  equilibrium geometry.

The same fitting procedure was applied to correct the vdW parameters of the atoms involved in the hydrogen bond, only focusing on the  $R$  correction since this procedure turned out to be the best for glutaric acid and with minor effects for adipic acid.

EVB-AMOEBA simulations at thermal equilibrium predict a distribution of interoxygen distance  $R$  shown in Fig. 12(b) that is similar to that obtained for adipic acid, with values ranging between 2.4 and 2.7 Å at 50 K, and between 2.4 and 2.8 Å at 300 K. The  $q$  distribution is clearly bimodal with both shared and localized protons, the former situation being favored at low temperature. We interpret this bimodal distribution as originating from the greater flexibility of the molecule caused by its longer CH<sub>2</sub> chain. Comparison of Fig. 8(a) and Fig. 12(c) shows that  $q$  is sampled over larger values for suberic acid than for adipic acid, up to 0.6 Å or more, indicating a higher propensity for localized proton.

IR absorption spectra obtained with the EVB-AMOEBA model at various temperatures in the 50–300 K range are shown in Fig. 13.

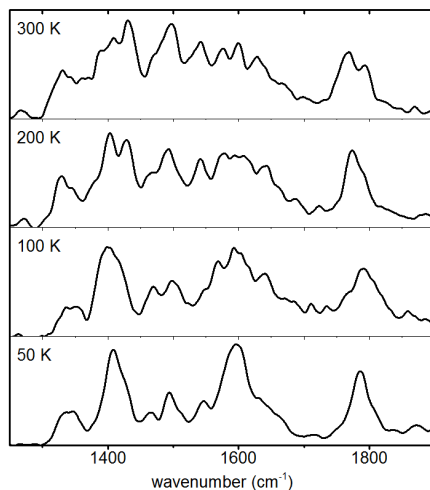


Figure 13: IR absorption spectra obtained with EVB-AMOEBA for suberic acid  $\text{HO}_2\text{C}-(\text{CH}_2)_6-\text{CO}_2^-$  at 50, 100, 200 and 300 K.

At 50 K, there are three main bands with maxima near 1400, 1600 and 1800  $\text{cm}^{-1}$ . The shape of the band around 1600  $\text{cm}^{-1}$  is rather similar to

that in the adipic acid spectrum (see Fig. 9), suggesting dynamical proton exchange. However the band near  $1800\text{ cm}^{-1}$  is suggestive of a carboxylic C=O stretch [60] while that near  $1400\text{ cm}^{-1}$  may be attributed to a symmetric stretching motion of a hydrogen bonded carboxylate [61, 62]. This superposition of carboxylic and carboxylate bands may be completed with a carboxylate antisymmetric stretching band near  $1600\text{ cm}^{-1}$ , although expected to be narrower than that in Fig. 13. Thus the 50 K spectrum appears to be best described as a mixture of a non classical spectrum as discussed before, and a more conventional spectrum exhibiting features that are typical of carboxylic and carboxylate groups. All these bands exist at higher temperatures up to 300 K (Fig. 13), however nearly continuous absorption is predicted to occur between  $1350$  and  $1700\text{ cm}^{-1}$ , probably indicative of conformational rearrangements associated with large amplitude motion.

These results suggest that proton sharing and exchange can be very sensitive to structural details in these series of dicarboxylic acids, as already suggested by IRPD [13, 14] and photoelectron spectra [12].

## 5 Conclusions

The AMOEBA polarizable force field, already known to provide an accurate computational framework for vibrational spectroscopy in the gas phase [48–50], was extended here to account for possible intramolecular proton transfer using a two-state empirical valence bond approach [10], with the aim of describing deprotonated dicarboxylic acids  $\text{HO}_2\text{C}-(\text{CH}_2)_n-\text{CO}_2^-$ . IRMPD spectra measured for  $n = 2$  to 4 reveal signatures of proton sharing that are accounted for neither by static quantum chemistry calculations in the harmonic approximation, nor by finite temperature anharmonic spectra based on the Fourier transform of the dipole moment time autocorrelation function using the nonreactive AMOEBA potential.

An EVB-AMOEBA reactive potential was then designed to reproduce *ab initio* reference geometries and energies, without introducing any experimental input. Parameterization of the EVB-AMOEBA model requires adjusting the coupling term, for which a simple functional form was proposed

in this work, as well as specific components for the diabatic AMOEBA contribution. This fitting part deals with the van der Waals terms involved in the description of the migrating proton. Fitting these parameters on the medium size adipic acid ( $n = 4$ ) molecule indicated that the choice of the reference geometry requires special care in order for the proton to display some effective mobility and not be stuck in a shared position between two oxygen atoms of the carboxylate groups. In practice, we found that adjusting the vdW parameters was sufficient to reproduce the MP2 data, the barriers for proton transfer being satisfactorily described by adjusting the EVB coupling term alone.

The finite temperature equilibrium dynamics obtained with the resulting EVB-AMOEBA model satisfactorily reproduces the measured IRMPD spectrum for adipic acid, especially the broad band at ca.  $1600\text{ cm}^{-1}$  which is characteristic of proton sharing. Trajectory analysis allows to distinguish between proton sharing and frequent proton transfer. We find that proton transfer dominates and occurs about every 500 fs at 100 K.

In applying the model to glutaric and suberic acids, the coupling term was found to be transferable, only the vdW parameters of the atoms involved in proton transfer requesting adjustment. Again, the experimentally measured IRMPD spectrum for glutaric acid ( $n = 3$ ) could be satisfactorily reproduced by our simulations, and our predicted spectrum for suberic acid ( $n = 6$ ) shows interesting temperature effects related to proton transfer. However, for succinic acid ( $n = 2$ ) the EVB-AMOEBA approach is probably less appropriate, since this shorter and highly constrained molecule has its proton in a very strongly shared position with reduced mobility.

Tunneling effects were not included in our modeling, as they are probably not essential at the relatively high temperatures considered here. However, they might further enhance proton mobility and lead to calculated spectra in better agreement with experimental data. It would thus be worth considering them in future extensions of the present EVB-AMOEBA model, e.g. through path-integral based approaches such as ring-polymer MD. Future extensions of this approach should first consider less symmetric cases, in which the two diabatic states are not just mirror images of one another.

A more ambitious extension should consider multiple protonation sites in a multistate EVB description. Work along these lines is in progress.

## **Acknowledgments**

CNRS and Ecole Polytechnique are acknowledged for funding. F.T. thanks Ecole Polytechnique for a PhD fellowship. Implementation of the EVB-AMOEBA model was supported by travel grants from GDR 3533 EMIE funded by CNRS. Additional financial support from the National FT-ICR network (FR 3624 CNRS) for conducting the research is gratefully acknowledged.

## References

- [1] P. Phatak, N. Ghosh, H. Yu, Q. Cui, M. Elstner, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 19672–19677.
- [2] P. Goyal, N. Ghosh, P. Phatak, M. Clemens, M. Gaus, M. Elstner, Q. Cui, *J. Am. Chem. Soc.* **2011**, *133*, 14981–14997.
- [3] A. Rahman, M. Mupa, C. Mahamadi, *Catal. Lett.* **2016**, *146*, 788–799.
- [4] J. Oomens, B. G. Sartakov, G. Meijer, G. von Helden, *Int. J. Mass Spectrom.* **2006**, *254*, 1–19.
- [5] L. MacAleese, P. Maitre, *Mass Spectrom. Rev.* **2007**, *26*, 583–605.
- [6] N. C. Polfer, J. Oomens, *Mass Spectrom. Rev.* **2009**, *28*, 468–494.
- [7] N. C. Polfer, P. Dugourd (Eds.), *Laser Photodissociation and Spectroscopy of Mass-separated Biomolecular Ions*, Springer, **2013**.
- [8] D. A. Thomas, M. Marianski, E. Mucha, G. Meijer, M. A. Johnson, G. von Helden, *Angew. Chem. Int. Ed.* **2018**, *57*, 10615–10619.
- [9] J. Oomens, J. D. Steill, B. Redlich, *J. Am. Chem. Soc.* **2009**, *131*, 4310–4319.
- [10] F. Thauay, F. Calvo, G. Ohanessian, C. Clavaguéra, *Theory and Applications of the Empirical Valence Bond Approach: From Physical Chemistry to Chemical Biology*, Vol. ISBN: 978-1-119-24539-1, F. Duarte, S. C. L. Kamerlin (Eds.), John Wiley & Sons, Ltd., **2017**, chapter Empirical Valence-Bond Models Based on Polarizable Force Fields for Infrared Spectroscopy.
- [11] E. Nicol, C. Clavaguéra, G. Ohanessian, *Int. J. Mass Spectrom.* **2018**, in press, <https://doi.org/10.1016/j.ijms.2018.10.006>.
- [12] H.-K. Woo, X.-B. Wang, K.-C. Lau, L.-S. Wang, *J. Phys. Chem. A* **2006**, *110*, 7801–7805.



- [13] C. T. Wolke, A. F. DeBlase, C. M. Leavitt, A. B. McCoy, M. A. Johnson, *J. Phys. Chem. A* **2015**, *119*, 13018–13024.
- [14] M. Z. Kamrath, R. A. Relph, T. L. Guasco, C. M. Leavitt, M. A. Johnson, *Int. J. Mass Spectrom.* **2011**, *300*, 91–98.
- [15] J. Lobaugh, G. A. Voth, *J. Chem. Phys.* **1996**, *104*, 2056–2069.
- [16] R. Vuilleumier, D. Borgis, *J. Chem. Phys.* **1999**, *111*, 4251–4266.
- [17] U. W. Schmitt, G. A. Voth, *J. Chem. Phys.* **1999**, *111*, 9361–9381.
- [18] D. C. Marinica, G. Grégoire, C. Desfrancois, J. P. Schermann, D. Borgis, M. P. Gaigeot, *J. Phys. Chem. A* **2006**, *110*, 8802–8810.
- [19] P. D. Dopieralski, Z. Latajka, I. Olovsson, *Acta Crystal. Section B* **2010**, *66*, 222–228.
- [20] Y. Yang, M. Meuwly, *J. Chem. Phys.* **2010**, *133*, 064503.
- [21] O. Vendrell, F. Gatti, H.-D. Meyer, *Angew. Chem. Int. Ed.* **2007**, *46*, 6918–6921.
- [22] Y. Wang, B. J. Braams, J. M. Bowman, S. Carter, D. P. Tew, *J. Chem. Phys.* **2008**, *128*, 224314.
- [23] T. Hammer, U. Manthe, *J. Chem. Phys.* **2011**, *134*, 224305.
- [24] M. Schröder, F. Gatti, H.-D. Meyer, *J. Chem. Phys.* **2011**, *134*, 234307.
- [25] M. Marciante, F. Calvo, *Mol. Simul.* **2014**, *40*, 176–184.
- [26] D. L. Howard, H. G. Kjaergaard, J. Huang, M. Meuwly, *J. Phys. Chem. A* **2015**, *119*, 7980–7990.
- [27] R. Biswas, Y.-L. S. Tse, A. Tokmakoff, G. A. Voth, *J. Phys. Chem. B* **2016**, *120*, 1793–1804.
- [28] Z.-H. Xu, M. Meuwly, *J. Phys. Chem. A* **2017**, *121*, 5389–5398.

- [29] S. Wang, L. MacKay, G. Lamoureux, *J. Chem. Theory Comput.* **2014**, *10*, 2881–2890.
- [30] X. Wu, W. Thiel, S. Pezeshki, H. Lin, *J. Chem. Theory Comput.* **2013**, *9*, 2672–2686.
- [31] S. Pezeshki, H. Lin, *J. Chem. Theory Comput.* **2015**, *11*, 2398–2411.
- [32] D. Riccardi, S. Yang, Q. Cui, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2010**, *1804*, 342 – 351.
- [33] A. Warshel, R. M. Weiss, *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.
- [34] A. Warshel, *Acc. Chem. Res.* **2002**, *35*, 385–395.
- [35] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- [36] P. Ren, J. W. Ponder, *J. Comput. Chem.* **2002**, *23*, 1497–1506.
- [37] P. Ren, J. W. Ponder, *J. Phys. Chem. B* **2003**, *107*, 5933–5947.
- [38] P. Maître, S. L. Caër, A. Simon, W. Jones, J. Lemaire, H. Mestdagh, M. Heninger, G. Mauclair, P. Boissel, R. Prazeres, F. Glotin, J.-M. Ortega, *Nucl. Instrum. Methods Phys. Res., Sect. A* **2003**, *507*, 541 – 546.
- [39] J. Oomens, J. D. Steill, *J. Phys. Chem. A* **2008**, *112*, 3281–3283.
- [40] B. M. Marsh, E. M. Duffy, M. T. Soukup, J. Zhou, E. Garand, *J. Phys. Chem. A* **2014**, *118*, 3906–3912.
- [41] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara,

- K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc. Wallingford CT 2009.
- [42] T. D. Rasmussen, P. Ren, J. W. Ponder, F. Jensen, *Int. J. Quantum Chem.* **2007**, *107*, 1390–1395.
- [43] D. Semrouni, G. Ohanessian, C. Clavaguera, *Phys. Chem. Chem. Phys.* **2010**, *12*, 3450–3462.
- [44] A. Grossfield, P. Ren, J. W. Ponder, *J. Am. Chem. Soc.* **2003**, *125*, 15671–15682.
- [45] J.-P. Piquemal, L. Perera, G. A. Cisneros, P. Ren, L. G. Pedersen, T. A. Darden, *J. Chem. Phys.* **2006**, *125*, 054511.
- [46] A. Marjolin, C. Gourlaouen, C. Clavaguéra, P. Ren, J. Wu, N. Gresh, J.-P. Dognon, J.-P. Piquemal, *Theo. Chem. Acc.* **2012**, *131*, 1–14.
- [47] D. Semrouni, W. C. I. Isley, C. Clavaguéra, J.-P. Dognon, C. J. Cramer, L. Gagliardi, *J. Chem. Theory Comput.* **2013**, *9*, 3062–3071.
- [48] D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian, C. Clavaguéra, *J. Chem. Theory Comput.* **2014**, *10*, 3190–3199.
- [49] F. Thaunay, J.-P. Dognon, G. Ohanessian, C. Clavaguera, *Phys. Chem. Chem. Phys.* **2015**, *17*, 25968–25977.
- [50] F. Thaunay, C. Jana, C. Clavaguéra, G. Ohanessian, *J. Phys. Chem. A* **2018**, *122*, 832–842.

- [51] A. J. Stone, *Chem. Phys. Lett.* **1981**, *83*, 233–239.
- [52] A. J. Stone, M. Alderton, *Mol. Phys.* **1985**, *56*, 1047–1064.
- [53] D. R. Nutt, M. Meuwly, *Biophys. J.* **2003**, *85*, 3612–3623.
- [54] T. A. Halgren, *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.
- [55] N. L. Allinger, Y. H. Yuh, J. H. Lii, *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- [56] J. W. Ponder, *TINKER - Software Tools for Molecular Design (version 7)*, <http://dasher.wustl.edu/tinker>, **2015**.
- [57] J. C. Wu, Ph.D. thesis, The University of Texas at Austin, **2011**.
- [58] R. Vuilleumier, D. Borgis, *J. Mol. Struct.* **1997**, *436–437*, 555 – 565.
- [59] C. M. Maupin, K. F. Wong, A. V. Soudackov, S. Kim, G. A. Voth, *J. Phys. Chem. A* **2006**, *110*, 631–639.
- [60] *NIST Chemistry WebBook* (<https://webbook.nist.gov/chemistry>).
- [61] T. Wende, M. Wanko, L. Jiang, G. Meijer, K. R. Asmis, A. Rubio, *Angew. Chem. Int. Ed.* **2011**, *50*, 3807–3810.
- [62] M. T. Rodgers, P. B. Armentrout, J. Oomens, J. D. Steill, *J. Phys. Chem. A* **2008**, *112*, 2258–2267.