

Formic Acid Dimerization: Evidence for Species Diversity from First Principles Simulations

Pawel Rodziewicz[†] and Nikos L. Doltsinis^{*‡}

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

Received: January 26, 2009; Revised Manuscript Received: April 19, 2009

Dimerization of formic acid has been simulated using ab initio molecular dynamics at conditions mimicking rare gas matrix isolation experiments. Aggregation product distributions and the corresponding reaction pathways have been studied as a function of temperature. At higher temperatures, the cyclic, C_{2h} symmetric, global minimum structure **A** with two $O-H\cdots O=C$ hydrogen bonds predominates over the second most stable, acyclic, local minimum isomer **B** with one $O-H\cdots O=C$ and one $C-H\cdots O=C$ hydrogen bond, whereas the latter is the main species at low temperature. Significant concentrations of two additional, less stable, local minimum species, **C** exhibiting an $O-H\cdots O=C$ and an $O-H\cdots O(H)-C$ hydrogen bond, and **D** with an $O-H\cdots O(H)-C$ and a $C-H\cdots O=C$ hydrogen bond, which should be detectable in experiment, are predicted at low temperature. Theoretical vibrational spectra are provided to guide the experimental search for these species. Furthermore, free-energy barriers for thermal interconversion of different dimer species have been calculated using targeted molecular dynamics in conjunction with thermodynamic integration. The small barrier for the $C \rightarrow A$ reaction of 7.0 kJ/mol indicates that the **C** species can only be stabilized at ultracold conditions. The data presented here thus hold important clues for the execution and interpretation of low-temperature vibrational spectroscopy at matrix isolation or ultracold helium droplet conditions.

I. Introduction

Noncovalent interactions play a key role in biology, chemistry, physics, and materials science. In supramolecular chemistry, for example, complex, weakly interacting, molecular aggregates are created with tailor-made properties with applications in catalysis and new materials.¹ Herein, important concepts such as molecular recognition and self-assembly are based on noncovalent interactions. Typically, the supramolecular structures formed are in thermal equilibrium and have the strongest possible intermolecular interactions. Very complex phenomena such as the formation of molecular crystals or the folding of proteins, where the number of noncovalent contacts is much larger, can no longer be explained by thermodynamic arguments alone. Instead they are often kinetically controlled.² These aggregation processes are still poorly understood in general. In particular, the prediction of the structure of molecular crystals is an unsolved problem.

Insight into the formation and structures of noncovalently bonded aggregates has been gained experimentally using a variety of techniques, including spectroscopy in rare gas matrix isolation, superfluid helium droplets, and supersonic molecular beams. It has been demonstrated for dimers of small organic molecules that at sufficiently low temperatures, T , local minimum structures can be predominantly formed, while the global minimum structure is only reached at high T thermal equilibrium.³ Going beyond dimers, larger aggregates are frequently dominated by cooperative effects, leading to nonadditive stabilization by long-range order of electrostatic moments and

electronic delocalization. Such cooperativeness can add to the fact that many larger aggregates are composed of dimer building blocks which do not correspond to the global minimum dimer structure.

Formic acid (FA) is the simplest carboxylic acid and serves as a prototypical molecule that can form two hydrogen bonds, either two strong $O-H\cdots O$ bonds or one strong $O-H\cdots O$ and a weaker $C-H\cdots O$ contact. In the liquid phase, recent first-principles and classical molecular dynamics simulations^{4,5} suggest the presence of large branched structures which consist of small clusters characterized by strong $O-H\cdots O$ H-bonds, which are held together by a weak $C-H\cdots O$ type of interaction. The population of the cyclic dimer configuration, which dominates in the gas phase, has been estimated to be 23%.⁵ The FA crystal structure, on the other hand, does not consist of dimeric species and reveals an infinite polymeric structure.⁶ Gas-phase FTIR molecular beam spectroscopy suggests the formation of the global minimum cyclic (C_{2h}) structure with two $O-H\cdots O=C$ hydrogen bonds,^{7,8} which is the only populated FA dimer (FAD) at room temperature. In the following, we shall refer to this structure as dimer **A**. Matrix isolation spectroscopy has revealed that the second most stable, local minimum, structure **B** with one $O-H\cdots O=C$ and one $C-H\cdots O=C$ hydrogen bond is formed at low temperatures below 40 K.⁹ Upon heating the matrix above 40 K, the dimers are converted to the thermodynamically most stable cyclic form. At ultracold conditions in superfluid helium droplets ($T = 0.37$ K), the local minimum structure was exclusively found.¹⁰ In total, six different local minimum structures have been described in the literature.^{11–13} However, most theoretical studies are static quantum chemical calculations merely characterizing the different dimer forms without attempting to understand their formation, that is, the dynamical aggregation process. A recent ab initio study¹⁴ describes the FA dimerization as a stepwise mechanism where the thermodynamically less stable acyclic dimer converts into the global minimum. According to the results utilizing the MP2-

* To whom correspondence should be addressed. E-mail: nikos.doltsinis@kcl.ac.uk.

[†] Present address: Interdisziplinäres Zentrum für Molekulare Materialien (ICMM) and Computer-Chemie-Centrum (CCC), Department Chemie und Pharmazie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nögelsbachstr. 25, 91052 Erlangen, Germany.

[‡] Present address: Department of Physics, King's College London, London WC2R 2LS, United Kingdom.

R12/K2 level of theory, the global minimum is 6.3 kcal/mol more stable than the acyclic structure. The barrier for rotation about the C=O bond, necessary to convert **A** to **B**, has been estimated to be 2.9 kcal/mol. DFT calculations (B3LYP/TZ2P(f,d)+diff) predict the values of 6.7 and 2.8 kcal/mol for the energy difference and rotational barrier, respectively, in a good agreement with the ab initio data.¹⁴

In the present work, we have carried out DFT-based first-principles molecular dynamics simulations of the FA dimerization process mimicking the conditions of rare gas matrix isolation experiments. The simulations provide detailed insight into the aggregation pathways and predict product distributions of different dimer structures as a function of temperature. Similar calculations have recently been employed successfully to interpret matrix isolation spectroscopic data on formamide dimers.¹⁵

In particular, we classify the possible first contact structures in terms of the H-bond formed, that is, either a strong O–H···O=C or weaker C–H···O=C or O–H···O(H)–C hydrogen bond, and quantify their abundance. For each initial contact structure, we follow the dynamical paths through to the final products, which often proceed via a number of local potential minima, and calculate a conditional product distribution.

In addition, we have investigated the dynamical mechanism of thermal interconversion of different local minimum structures using the targeted MD (TMD) method.¹⁶ This method has been applied previously to study double proton transfer in FA dimers¹⁷ and yields free-energy profiles, including activation barriers, along the transition pathways. Our theoretical results provide guidance for future IR spectroscopic studies, which, in turn, will be able to probe the theoretical predictions. As a concrete help for the experimental search of the new dimer species proposed here, we have calculated the IR vibrational spectra of all relevant species. In addition to identifying the experimentally observable dimer products, our simulations contribute to a more complete understanding of the aggregation process, including precursor and intermediate structures that are not easily detectable or sometimes not accessible at all in spectral analysis.

II. Computational Methods

All calculations were performed using the plane wave density functional code CPMD, version 3.10.¹⁸ The PBE exchange–correlation functional¹⁹ and a plane wave basis truncated at 25 Ry were used. Core electrons were taken into account through the use of Vanderbilt ultrasoft pseudopotentials.²⁰ The runs were carried out screening the periodic boundary conditions using Martyna and Tuckerman's method²¹ to solve the Poisson equation. Geometry optimizations were performed in a smaller unit cell of size $13.2 \times 10.6 \times 10.6 \text{ \AA}^3$. Vibrational frequencies were calculated for the optimized structures to verify that they indeed correspond to potential minima. Most of the starting structures for geometry optimization were taken from the literature;^{11–13} others were constructed taking into account all possible hydrogen bonding arrangements. A recent study of a similar system has shown that no important additional structures could be found using extensive random sampling of starting geometries.¹⁵ For more complex systems, a random sampling procedure using molecular dynamics could be employed.²² The set of optimized FAD structures found here has proven to ensure a complete description of all stages of the dimerization process.

To enhance the predictive power of our theoretical vibrational frequencies, these calculations were carried out using a higher, 30 Ry, cutoff. Further increase of the cutoff to 35 Ry has been found to change the monomer frequencies by less than 1% and

dimer frequencies by less than 3%. Furthermore, we performed DFT calculations using the PBE exchange–correlation functional with a localized basis set, that is, 6-31+G(d,p), using the Gaussian program²³ to estimate the intensities of the vibrational lines. All aggregation simulations were carried out using Car–Parrinello molecular dynamics (CP–MD) in the same orthorhombic unit cell of size $19.1 \times 11.1 \times 11.1 \text{ \AA}^3$ with a time step of 4 au and a fictitious electron mass of 400 au at ionic temperatures of 50, 100, and 150 K. We verified that the shortest distance between the molecules and the edge of the box did not fall below 3.0 Å at any time. For each temperature, a set of 25 initial configurations was picked at random from a CP–MD simulation at 300 K with fixed positions of the C atoms of both monomers at a C–C distance of 8.15 Å. During these constrained initial runs, a Nosé–Hoover^{24,25} chain thermostat was coupled to each degree of freedom, resulting in heavy rotation of the two monomers. Subsequently, the distance constraint was released, and a single Nosé–Hoover^{24,25} chain of length three was applied to the entire system. During these unconstrained runs of about 10 ps length, 11, 15, and 11 successful dimerization events took place at 50, 100, and 150 K, respectively. Unsuccessful events are those in which the two monomers do not approach each other at all or even move away from each other during the 10 ps simulation time.

In addition, free-energy calculations for the transition between relevant local and global minimum structures were performed using the targeted CPMD technique.¹⁷ The temperature was set to 100 K, and a cubic unit cell of size 13.2 \AA was used. For each fixed value of the target distance, average constraint forces were calculated from runs exceeding a length of 1 ps each. On the reagent's side of the barrier, the runs were 2 ps long, while on the product side, which was not used for quantitative analysis, the runs were 1 ps long.

III. Results and Discussion

A. Local Minimum Dimer Structures. We first discuss the most stable optimized dimer structures as they will form the basis of our later analysis of aggregation dynamics. Figure 1 shows the seven lowest-energy FAD structures as well as their relative energies as obtained from our DFT geometry optimizations. Within the set of seven structures, one can distinguish the following types of H-bond interactions (ordered by increasing strength): (i) O–H···O=C, (ii) O–H···O(H)–C, (iii) C–H···O=C, and (iv) C–H···O(H)–C. The lowest-energy structure is the cyclic dimer **A**, which has two strong O–H···O=C hydrogen bonds and C_{2h} symmetry. Dimer **B** (C_s symmetry), being the second most stable isomer with one O–H···O=C and one C–H···O=C hydrogen bond, is already significantly higher in energy by 33.5 kJ/mol. The least stable dimer **G**, having two C–H···O(H)–C hydrogen bonds and C_{2h} symmetry, is 65.1 kJ/mol higher in energy than dimer **A** and is therefore not expected to play a significant role.

Our DFT/PBE dimerization energies for the seven isomers are collected in Table 1 together with other data from the literature. It is particularly noteworthy that the MP2 results by Qian and Krimm¹² are in good agreement with the DFT/PBE values. Although it is unclear which of the methods presented in Table 1 yields the most accurate binding energies, it is important to note that they all produce the same energetic ordering. In addition, in contrast to atom-centered basis sets, the present plane wave calculations do not suffer from the basis set superposition error. Thus, we are confident that the current approach is sufficiently reliable for the purpose of studying FAD aggregation and isomerization dynamics.

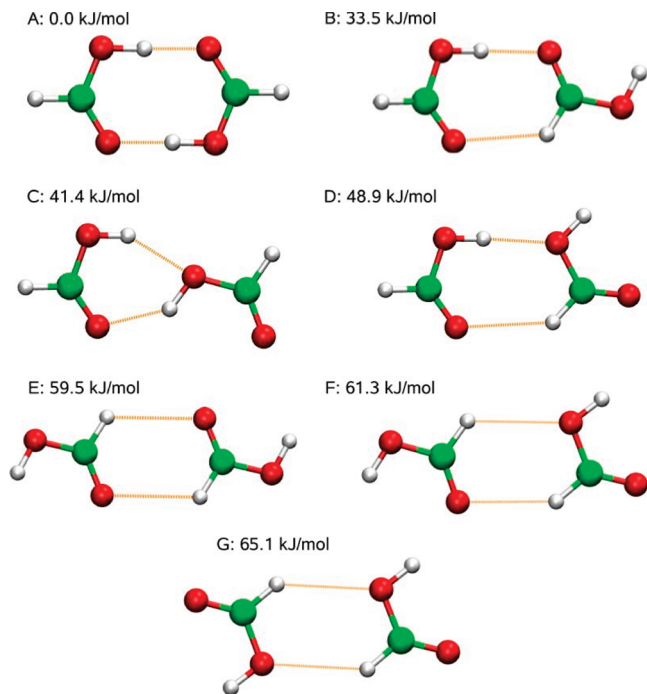


Figure 1. Optimized FA dimer structures and their relative energies.

TABLE 1: Comparison of FAD Absolute Binding Energies (in kJ/mol) Obtained in This Work (tw) Using the Plane Wave DFT/PBE with Literature Data

method/ isomer	PBE [tw] pw 25 Ry	BLYP ⁵ pw 70 Ry	BLYP ⁵ aug-cc-pVTZ	MP2 ¹² 6-311++G**
A	69.1	59.0	58.2	56.9
B	35.6	33.1	31.0	32.7
C	28.1	26.8	24.7	26.4
D	20.1	19.3	18.4	22.6
E	9.6	7.5	9.2	13.4
F	7.5			10.9
G	3.8			8.0

In order to provide guidance to IR spectroscopy experiments, with the aid of which our theoretical predictions for dimer structures derived from the aggregation dynamics simulations described below may be tested, we have calculated harmonic vibrational frequencies for all dimers in Figure 1. Table 2 contains calculated frequencies of the most important bond vibrations for the monomer and the five lowest-energy dimers. A graphical representation of the dimer vibrational modes of Table 2 is displayed in Figure 2.

Dimer **A** is easily distinguishable from the other isomers by its very intense OH stretching vibrational line, which is massively red-shifted with respect to the monomer by 871 cm^{-1} . The OH signatures of the other four dimer structures are also all very distinct and should be straightforwardly identifiable spectroscopically. Compared to the monomer, the dimer OH lines exhibit a strong increase in intensity—by 2 orders of magnitude in the case of dimer **A**! The extent of this enhancement depends on the strength of the H-bonds in the dimer. Interestingly, this phenomenon is also observed in dimer **E**, although this structure does not possess any O—H \cdots O H-bonds. This can be explained as an indirect effect caused by the structural changes of the other parts of the molecules upon dimerization.

The differences are more subtle in the case of the CH vibrations. Dimers **A** and **E** both have a single line blue-shifted

by 20 cm^{-1} , while for dimers **B**, **C**, and **D**, one line is blue-shifted by $21\text{--}43\text{ cm}^{-1}$, and the other is red-shifted by $12\text{--}33\text{ cm}^{-1}$. The blue-shifted feature, the so-called improper H-bond, which has been predicted for weak C—H \cdots O interactions previously,²⁶ is typically accompanied by the unusual decrease of the intensity of this vibration, as is the case here. However, CH intensities are rather weak in comparison with OH lines, so that this spectral region may be less suitable for the identification of dimer species.

The CO lines appear to be more suitable again for distinguishing between the five different dimers. Dimers **A** and **E** exhibit a single line, red-shifted by 63 and 14 cm^{-1} , respectively, while dimer **B** has two red-shifted lines, the stronger one shifted by 29 , the weaker one by 64 cm^{-1} . For dimers **C** and **D**, one CO line is blue-shifted by 3 and 10 cm^{-1} , respectively, while the other line is red-shifted by 50 and 25 cm^{-1} , respectively.

B. Aggregation Kinetics. In order to shed light on the mechanisms governing aggregation in rare gas matrix isolation experiments and in ultracold helium droplets, three series of CP—MD simulations were carried out at the different temperatures, 50 , 100 , and 150 K , each run starting from two well-separated, randomly oriented monomers (see Computational Details). Thus, our theoretical experiment captures the entire aggregation mechanism, from the first contact structure, that is, the initial metastable structure that we shall refer to as precursor, via a first (local) energy minimum to the final product structure.

At this point, it is important to realize that our CP—MD simulations treat the atomic nuclei as classical particles. The simulation temperatures should therefore be compared to somewhat lower experimental temperatures.²⁷ Classical simulations close to 0 K are unrealistic since they do not account for zero-point vibrational motion. Despite the fact that there is no rigorous mapping of classical onto real temperature, our simulations at three different temperatures allow us to predict general trends, for example, for dimer product abundances and aggregation mechanisms.

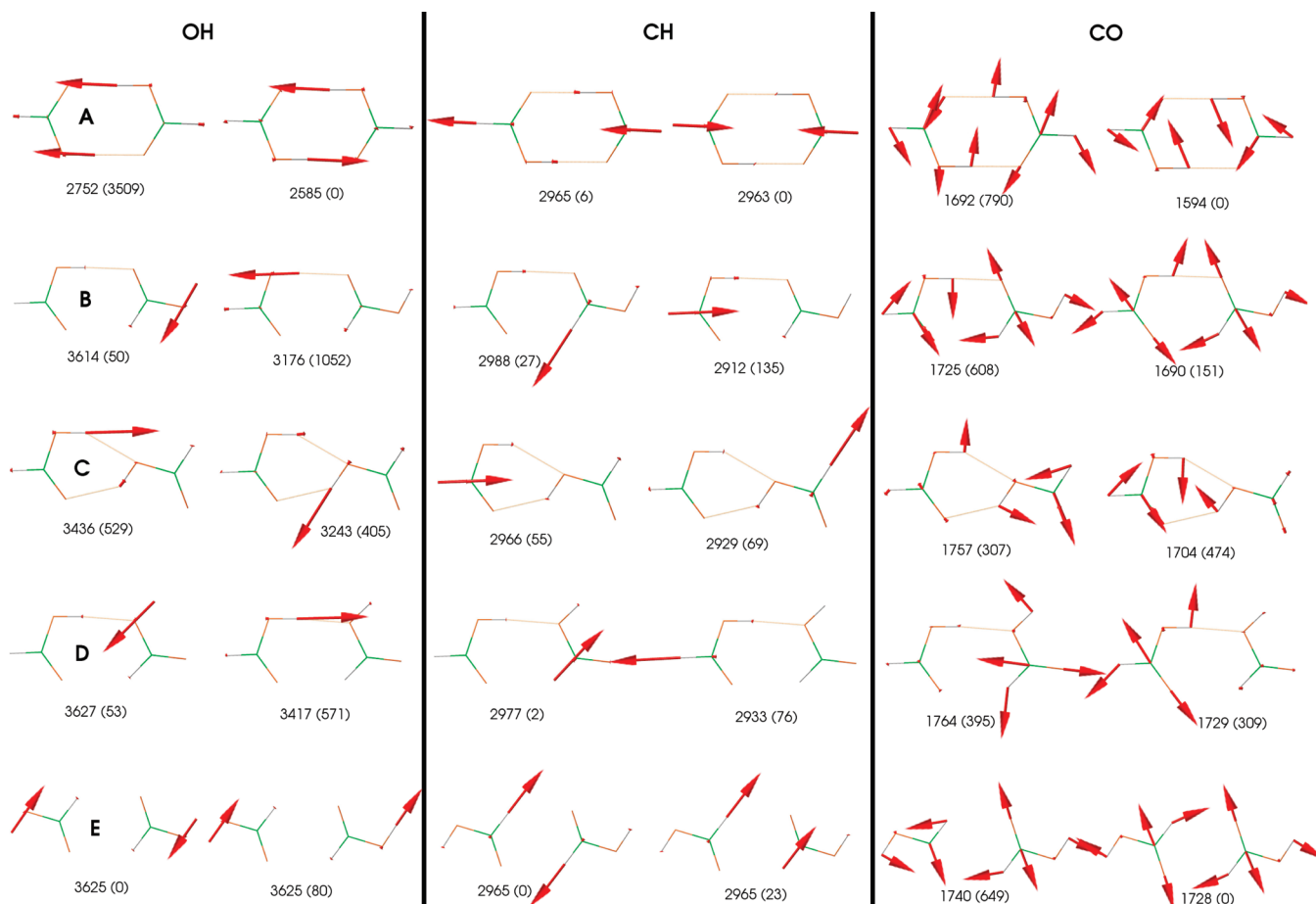
Let us begin with an overview of the aggregation product species obtained from the simulations as a function of temperature and compare it to the experimental findings available in the literature. The relative abundances of final dimer products at the different temperatures are plotted in Figure 3. At the lowest temperature considered here, that is, 50 K , the dominant product is dimer **B**, occurring in roughly half (46%) of all aggregation events. This is in accord with spectroscopic measurements in superfluid helium droplets,¹⁰ where this species was exclusively observed. The local minimum structure **B** has also been observed in argon matrix isolation experiments at low (15 K) temperatures before annealing to 40 K .⁹ No other isomers have been reported so far experimentally at very low temperature. Our simulations suggest, however, that, initially, also FAD isomers **A**, **C**, and **D** are formed in low concentrations (18% each). Bearing in mind that the aggregation simulations only cover a time scale of 10 ps , we cannot exclude at this point that thermal rearrangement between different isomers takes place at a later stage. This may be expected, in particular, for the less stable local minimum structures **C** and **D**. We shall return to this issue in section III.C, where we characterize free-energy landscapes and reaction paths for thermal rearrangement.

Upon increasing the simulation temperature to 100 K , the abundance of dimer **B** decreases to 40% , while that of dimer **A** increases to 33% (see Figure 3). The population of dimer **C** changes only a little to 20% ; dimer **D** falls to 7% . Raising the temperature further to 150 K confirms the main trend, namely,

TABLE 2: Selected Stretching Vibrational Frequencies in cm^{-1} for the Formic Acid Monomer and Its Global and Local Minima Dimers^a

structure	OH	CH	C=O
monomer	3623 [42]	2945 [43]	1754 [353]
dimer A	2752 (−871), 2585 (−1038) [3517, 0]	2965 (20), 2963 (18) [6, 0]	1692 (−63), 1594 (−161) [790, 0]
dimer B	3614 (−9), 3176 (−447) [50, 1052]	2988 (43), 2912 (−33) [27, 135]	1725 (−29), 1690 (−64) [608, 151]
dimer C	3436 (−187), 3243 (−380) [529, 405]	2966 (21), 2929 (−16) [55, 69]	1757 (3), 1704 (−50) [307, 474]
dimer D	3627 (4), 3417 (−206) [53, 571]	2977 (32), 2933 (−12) [2, 76]	1764 (10), 1729 (−25) [395, 309]
dimer E	3625 (2), 3625 (2) [0, 80]	2965 (20), 2965 (20) [0, 23]	1740 (−14), 1728 (−26) [649, 0]

^a Shifts of the dimer lines relative to the monomer are given in parentheses, and IR intensities are in units of km/mol in square brackets.

**Figure 2.** Vibrational normal modes of Table 2 with frequencies in cm^{-1} (parentheses contain the corresponding intensities in km/mol).

the decrease of the dimer **B** population (now 27%) at the expense of a growing dimer **A** population (now 55%; see Figure 3). Interestingly, dimer **C** completely vanishes at 150 K, while dimer **D** retains a population of 18%.

We must bear in mind, of course, the significant statistical uncertainty in the theoretical abundances associated with the limited number of successful dimerization events (11, 15, and 11 at 50, 100, and 150 K, respectively). However, our observation that dimer **B** is favored at low temperature whereas dimer **A** predominates at higher temperature is in line with experimental evidence from matrix isolation and ultracold helium droplet spectroscopy.^{9,10} Beyond that, our data clearly suggest that it should be possible to detect low concentrations of dimers **C** and **D** by spectroscopic measurements in cold rare gas matrixes.

Why is the local minimum structure **B** the most likely aggregation product at low temperature? In the following, we will try to shed light on this question based on the information obtained from the simulations. Let us first analyze in which way the two monomers approach each other from large distance and which kind of H-bond contacts are initially formed. It is obvious that dipole–dipole interactions are mainly responsible for the intermolecular attraction at very large distance. However, unlike in the case of formamide,¹⁵ which has a strong dipole moment of 4.5 D, the FA dipole moment is too small (1.4 D) to remain the dominant driving force at shorter intermolecular distance. Instead, it is predominantly the interactions between the most polar groups, that is, the OH and CH groups on one monomer with the C=O or OH oxygen lone pairs of the other monomer, that are responsible for the first contact structure,

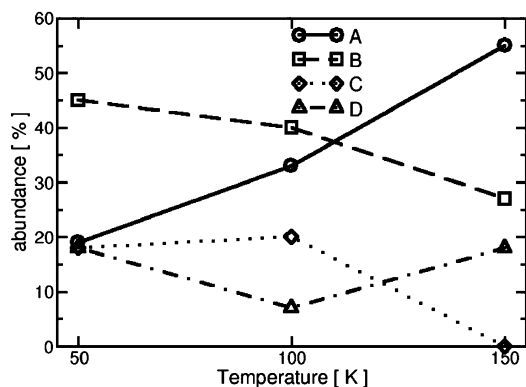


Figure 3. Relative abundances of dimerization products at different temperatures. Dimer **A** (○): 18% at 50 K, 33% at 100 K, 55% at 150 K. Dimer **B** (□): 46% at 50 K, 40% at 100 K, 27% at 150 K. Dimer **C** (◇): 18% at 50 K, 20% at 100 K, 0% at 150 K. Dimer **D** (△): 18% at 50 K, 7% at 100 K, 18% at 150 K.

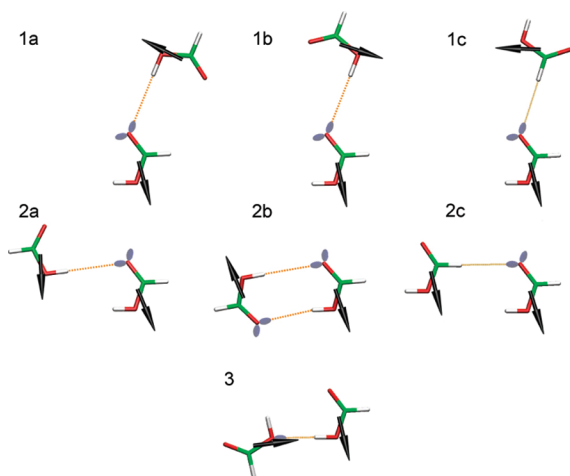


Figure 4. Idealized model structures of first dimerization contacts. The structures labeled **1a**, **1b**, and **1c** all involve lone pair **1** of the C=O oxygen atom; **2a**, **2b**, and **2c**, involve lone pair **2** of the C=O oxygen atom; **3** involves the lone pair of the OH group.

the precursor. We have compared the dipole–dipole interaction energy with the total DFT interaction energy for various configurations at an intermolecular center-of-mass distance of 5 Å and found that the magnitude of the former is typically smaller by 1 order of magnitude than the latter.

In an attempt to develop a simplified model, we distinguish between the seven different limiting first contact scenarios depicted in Figure 4. These precursor structures have been constructed on the basis of the actual observations made during the MD runs; they are however idealized and have not been optimized.

The oxygen atom of the carbonyl group, C=O, of monomer **1** can be approached by the hydrogen atom of the OH group of monomer **2** in four different ways to form a H-bond. We label the two contact structures involving the first O lone pair as **1a** and **1b**; they are both planar and related to each other by rotation about the H-bond by 180°. Similarly, the two structures involving the second O lone pair are denoted as **2a** and **2b**.

In addition, H-bonds of the type C–H···O have been frequently observed in our simulations. Thus, we propose a model where the oxygen atom of the C=O group of monomer **1** is approached by the hydrogen atom of the CH group of monomer **2**. Precursors **1c** and **2c** in Figure 4 show the formation of a C–H···O H-bond involving either the first or the second

TABLE 3: Abundances (in Percent) of the Different Precursors Defined in Figure 4, the Distribution of the First (Local) Minimum Structures Reached from Each Precursor, and the Corresponding Final Distributions at 50, 100, and 150 K

precursor	first minimum					final minimum			
	A	B	C	D	E	A	B	C	D
<i>T</i> = 50 K									
1a	9	0	100	0	0	0	100	0	0
1b	27	0	100	0	0	33	66	0	0
1c	0	0	0	0	0	0	0	0	0
2a	46	0	0	100	0	20	40	40	0
2b	0	0	0	0	0	0	0	0	0
2c	9	0	0	0	100	0	0	0	100
3	9	0	0	0	100	0	0	0	100
total	100	0	36	46	18	18	46	18	18
<i>T</i> = 100 K									
1a	7	0	100	0	0	0	100	0	0
1b	13	50	50	0	0	50	50	0	0
1c	7	0	0	0	100	0	100	0	0
2a	20	0	33	66	0	33	33	33	0
2b	7	100	0	0	0	100	0	0	0
2c	26	0	0	0	100	0	50	25	25
3	20	0	0	0	100	0	33	33	33
total	100	13	20	13	47	33	40	20	7
<i>T</i> = 150 K									
1a	27	0	100	0	0	33	66	0	0
1b	27	66	33	0	0	100	0	0	0
1c	9	0	0	0	100	0	100	0	0
2a	9	0	0	100	0	100	0	0	0
2b	0	0	0	0	0	0	0	0	0
2c	9	0	0	0	100	0	0	0	100
3	18	0	0	0	100	0	50	0	50
total	100	18	37	9	27	55	27	0	18

O lone pair of monomer **1**. Furthermore, we also detected a precursor with an O–H···O H-bond between the hydrogen atom of the OH group of monomer **1** and the oxygen atom of the OH group of monomer **2** (see Figure 4, **3**).

Taking into account all our successful aggregation simulations at 50 K, we count 9% first contacts of type **1a**, 27% **1b**, 46% **2a**, 9% **2c**, and 9% **3** (see the second column in Table 3). To fully understand the pathways leading to a particular product, it is now important to investigate which local minimum is first reached from each type of precursor. This information should provide vital clues about the dimers formed at extremely low temperature, as in superfluid He nanodroplets, where the aggregation path is not likely to go beyond the first potential barrier encountered. Such statistics are presented in Table 3.

First, we would like to emphasize that **2b**-type contacts were never observed at 50 K. This is significant because **2b** would be the most direct precursor of the global minimum structure **A**, indicating that dimer **A** is unlikely to be kinetically formed but only via thermodynamic conversion from a local minimum structure.

The most likely precursor was found to be **2a** (46% abundance), which, in all cases, leads to the formation of dimer **C** initially. By the end of the simulation, 40% of these initial **C** structures had rearranged to form dimer **B**, and another 20% formed dimer **A** (see Table 3). Note that **2a** is the only precursor that leads to **C**. Snapshots from a typical **2a** → **C** trajectory are depicted in the top panel of Figure 6.

The most abundant product at 50 K overall, that is, taking into account all of the different precursors, is dimer **B** (46%). As Table 3 demonstrates, another large portion of the final **B** structures is formed via precursor **1b**, which initially always

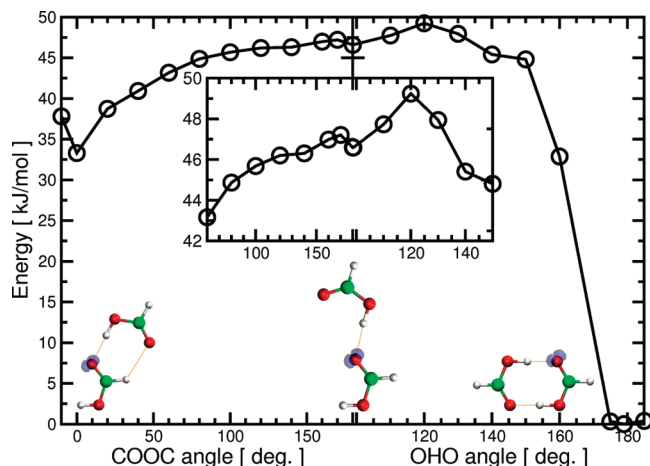


Figure 5. Minimum-energy profiles along the $\text{C}=\text{O}\cdots\text{O}-\text{C}$ dihedral and $\text{O}-\text{H}\cdots\text{O}$ angles starting from the **1b** initial contact structure (see Figure 4) shown at the center. The rotation along the $\text{C}=\text{O}\cdots\text{O}-\text{C}$ dihedral angle from its initial value of 180 to 0° leads to isomer **B**, whereas the change of $\text{O}-\text{H}\cdots\text{O}$ from its initial value of 99 to 180° produces isomer **A**.

transforms to **B**, but one-third of the initial **B** structures rearrange to become **A**. Can we understand why dimer **B** is the preferred product for the **1b** precursor scenario? The potential scan shown in Figure 5 illustrates that precursor **1b** transforms into **B** by straightforward, barrierless rotation about the H-bond. In order to reach dimer **A** from the same starting point, a potential barrier of ~ 2.5 kJ/mol needs to be overcome (Figure 5), which is associated with breaking the H-bond with one O lone pair and forming a new H-bond with the other lone pair. In addition, the former rotation has a much smaller moment of inertia and therefore occurs much faster than the latter. This particular **1b** \rightarrow **B** dimerization path is further illustrated in the lower panel of Figure 6, showing snapshots from a representative trajectory.

The remaining portion of the dimer **B** final products is formed via the first contact structure **1a** (Table 3). This precursor is closest to **B** out of all precursors; thus, the **1a** \rightarrow **B** path is the most direct route to the formation of **B**.

The formation of dimer **D**, hitherto undetected in experiment but accounting for 18% of all final products, is seen to occur with equal probabilities via the precursors **2c** and **3**. Interestingly, both precursors lead to direct formation of **D** without passing through other local minima on the way (Table 3).

The main net changes upon increasing the temperature to 100 K are an increase in the final **A** population and a decrease in the **D** population. This is partly due to thermal rearrangement of **D** structures, initially formed via the most likely precursor **2c**, to the global minimum structure **A**. The top panel of Figure 7 shows that this rearrangement typically proceeds via a number of local minima. The initial **D** population arising from precursor **2c** is hereby completely depleted. In contrast, the second most likely precursor **3**, also exclusively resulting in **D** as the first dimer structure, does not lead to **A** but rather **B**, **C**, and **D** with equal probabilities. A typical direct **3** \rightarrow **D** trajectory is illustrated in the lower panel of Figure 7.

As one would expect, a further increase in temperature to 150 K leads to another rise in the population of the thermodynamically most stable structure **A**. The most likely joint precursors are **1a** and **1b** (27% each; see Table 3). While at lower temperature **1a** exclusively leads to **B**, at 150 K, we observe that a third of these structures transform to **A**. The remainder of **1a** precursors still result in **B**. This most direct formation of **B** is illustrated in the top panel of Figure 8.

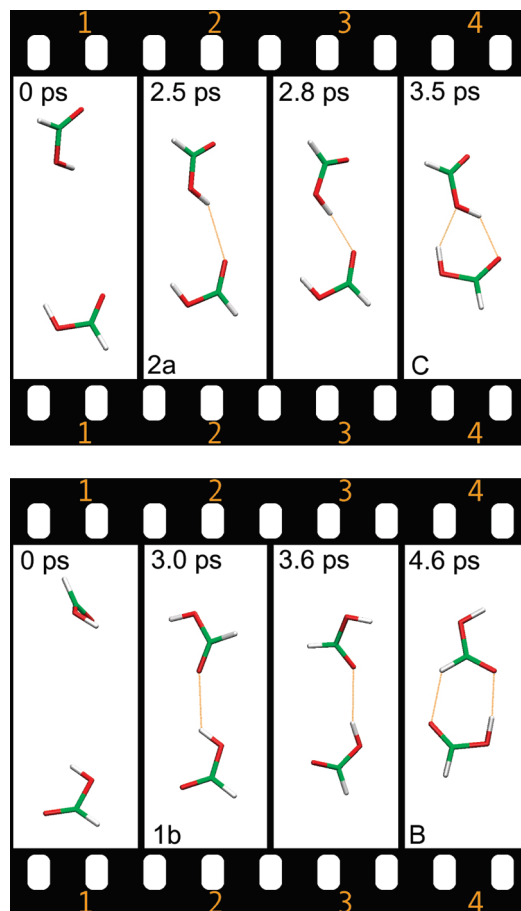


Figure 6. Snapshots from two selected aggregation simulations at 50 K. Top: (1) $t = 0$: The two monomers are released at a C–C distance of 8.15 Å. (2) $t = 2.5$ ps: A first $\text{O}-\text{H}\cdots\text{O}=\text{C}$ contact is established (precursor **2a**). There is additional hydrogen bond interaction of the $\text{O}-\text{H}\cdots\text{O}(\text{H})-\text{C}$ type. (3) $t = 2.8$ ps: The $\text{O}-\text{H}\cdots\text{O}=\text{C}$ hydrogen bond becomes linear and shortens to near-equilibrium distance. (4) $t = 3.5$ ps: The second, $\text{O}-\text{H}\cdots\text{O}(\text{H})-\text{C}$ hydrogen bond is optimized, and dimer **C** formed. Bottom: (1) $t = 0$: The two monomers are released at a C–C distance of 8.15 Å. (2) $t = 3.0$ ps: A first $\text{O}-\text{H}\cdots\text{O}=\text{C}$ contact is established (precursor **1b**). (3) $t = 3.6$ ps: Rotation about the $\text{O}-\text{H}\cdots\text{O}=\text{C}$ hydrogen bond. (4) $t = 4.6$ ps: A second $\text{C}-\text{H}\cdots\text{O}=\text{C}$ hydrogen bond and thus dimer **B** are formed.

Continuing the trend that has emerged from analyzing the product distributions at 50 and 100 K, precursor **1b** now exclusively leads to **A** at 150 K. The same trend prevails also in the distributions of first minimum structures following on from **1b** (Table 3). An example of a trajectory that connects **1b** with the final product **A** via the local minimum **C** is shown in the lower panel of Figure 8.

Perhaps somewhat surprising, there is still a considerable final population of **D** at this temperature, whereas **C** has vanished completely. At this point, the reader should be reminded of (a) the relatively large statistical uncertainty in the percentages listed in Table 3 due to the limited number of successful aggregation events that have been computationally feasible to study and (b) the fact that the final product distributions refer to the simulation time scale of 10 ps and may shift toward the lowest-energy structures for longer times. Having said that, at the ultracold conditions of superfluid helium nanodroplet or rare gas matrix isolation experiments, it should certainly be possible to stabilize even the less stable products **C** and **D**, and we would definitely encourage experimentalists to search for such evidence.

C. Thermodynamic Structural Rearrangements. According to the above aggregation simulations, four different dimer

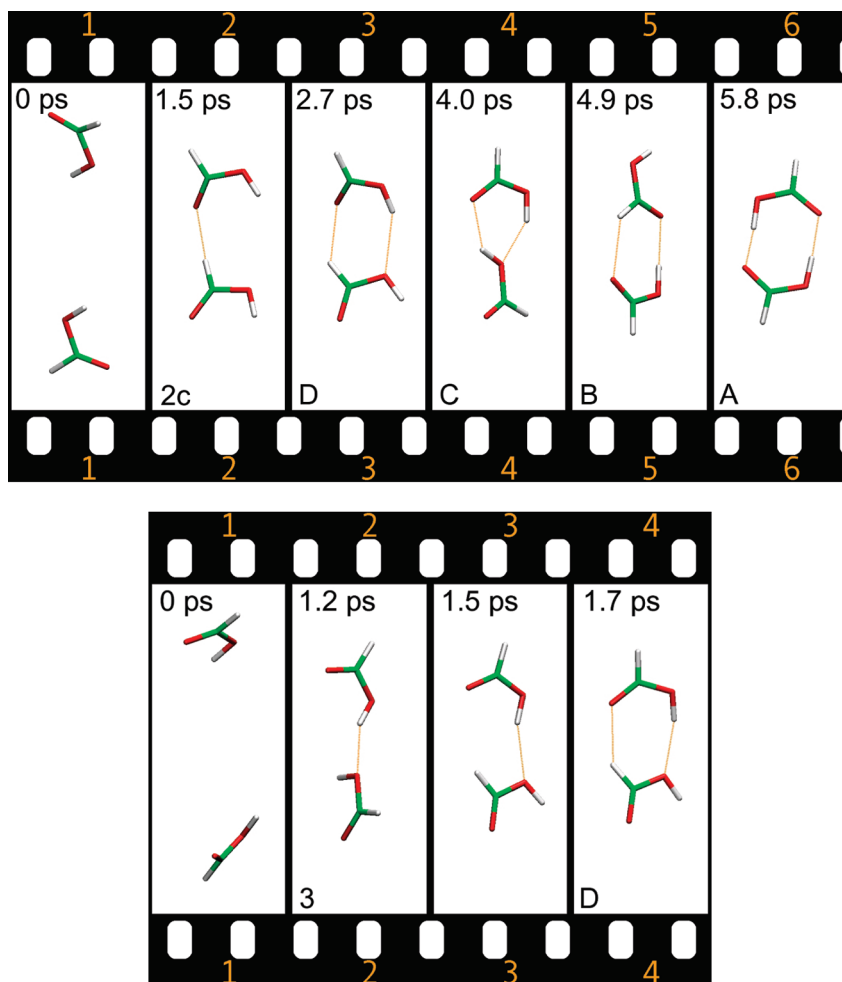


Figure 7. Snapshots from two selected aggregation simulations at 100 K. Top: (1) $t = 0$: The two monomers are released at a C–C distance of 8.15 Å. (2) $t = 1.5$ ps: A first C–H \cdots O=C contact is established (precursor **2c**). (3) $t = 2.7$ ps: The first local minimum, **D**, is traversed. (4) $t = 4.0$ ps: A second local minimum, **C**, is traversed. (5) $t = 4.9$ ps: A third local minimum, **B**, is traversed. (6) $t = 5.8$ ps: The final, global minimum structure **A** is reached. Bottom: (1) $t = 0$: The two monomers are released at a C–C distance of 8.15 Å. (2) $t = 1.2$ ps: A first O–H \cdots O(H)–C contact is established (precursor **3**). (3) $t = 1.5$ ps: Rotation about the first H bond leads to a C–H \cdots O=C second contact. (4) $t = 1.7$ ps: The final structure **D** is reached.

structures could be present in rare gas matrixes at very low temperatures. However, strictly speaking, the temperature-dependent distributions (Figure 3 and Table 3) only reflect the situation a few picoseconds after the first contact between the two monomers. All local minimum FAD structures may be expected to undergo thermal rearrangement to lower-energy isomers beyond the picosecond time scale depending on the temperature. To explore this possibility, we have carried out free-energy calculations using targeted MD (TMD) for the most likely conversions, that is, **C** \rightarrow **A**, **B** \rightarrow **A**, and **B** \rightarrow **C**. These rearrangement scenarios have been suggested by the aggregation pathways described above. Here, we can analyze in detail the reaction mechanisms and the free-energy barriers associated with the particular paths.

Figure 9 shows the free-energy profiles along the TMD reaction path for the three conversions. In each case, a series of simulations was carried out starting from the initial local minimum structure corresponding to the largest target distance shown. The target distance was then incrementally decreased until the target dimer structure was reached. Note, however, that the target distance can never be exactly zero as this would not allow the atoms to move.

This way, the free-energy barrier for the **B** \rightarrow **C** rearrangement process, frequently observed during aggregation (see, for instance, Figure 8), was calculated to be 16.2 kJ/mol (see Figure 9).

The rearrangement process starts with cleavage of the C–H \cdots O bond, followed by out-of-plane rotation about the O–H \cdots O H-bond. In the transition-state structure, the two planes, spanned by the monomers atoms, are roughly perpendicular to each other (see Figure 9). The O–H \cdots O bond remains intact throughout the entire transformation. Further rotation results in the formation of an O–H \cdots O(H)–C H-bond and thus dimer **C**.

We would like to point out that our TMD free-energy curves (Figure 9) should not be used to derive free-energy barriers for the reverse reactions, for example, **C** \rightarrow **B**. This is due to the well-known overestimation of free energies by TMD close to the target structure as a result of the decrease of the accessible configuration space.¹⁷

For the sake of comparison and to obtain an estimate of the entropic contribution to the free energy, we have also calculated the energy difference between the optimized initial structure, **B**, and the optimized transition state. The calculated energy difference of 13.4 kJ/mol is only 2.8 kJ/mol lower than the free-

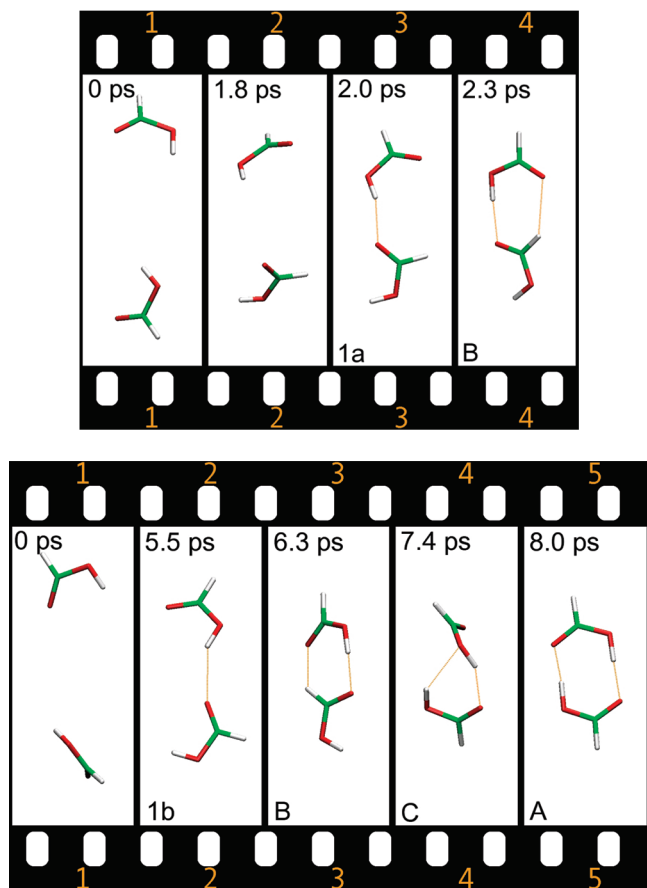


Figure 8. Snapshots from two selected aggregation simulations at 150 K. Top: (1) $t = 0$: The two monomers are at the fixed C–C distance of 8.15 Å. (2) $t = 1.8$ ps: A first O–H...O=C contact is formed, the two monomers not being coplanar. (3) $t = 2.0$ ps: Precursor 1a is formed, the two monomers being quasi-coplanar. (4) $t = 2.3$ ps: Formation of the final structure B. Bottom: (1) $t = 0$: The two monomers are at the fixed C–C distance of 8.15 Å. (2) $t = 5.5$ ps: A first O–H...O=C contact is formed (precursor 1b). (3) $t = 6.3$ ps: The first local minimum, B, is traversed. (4) $t = 7.4$ ps: The second local minimum, C, is traversed. (5) $t = 8.0$ ps: The final, global minimum, A, is reached.

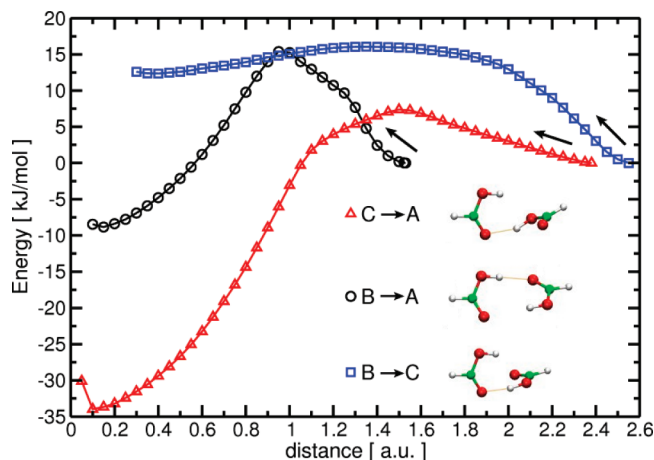


Figure 9. Free-energy curves from TMD simulations for the thermal rearrangements B → A, B → C, and C → A. The arrows above the curves indicate the direction in which the simulations were carried out. The free energies of the initial structures (large target distances) were set to 0. Also shown are the respective transition-state structures.

energy value. However, this corresponds to a relative entropic contribution of ~20%.

Analogously, the free-energy barrier for the B → A reaction pathway was calculated to be 15.5 kJ/mol. This is slightly higher than the energy value at 0 temperature of 9.6 kJ/mol calculated by Gantenberg et al.⁹ at the B3LYP/6-311++G** level. As their result is not a free energy, it should be compared to our energy difference of 12.7 kJ/mol between optimized transition-state and initial structures. As in the B → C case, the entropic contribution to the free energy is thus 2.8 kJ/mol in our simulations.

The mechanism in the B → A reaction pathway is quite similar to the one observed for the B → C process. First, the weaker C–H...O bond breaks, and then, the rotation along the remaining O–H...O enables the formation of the second O–H...O bond. The transition-state structure consists of two quasi-perpendicular planes, that is, the monomers, bridged by the O–H...O H-bond (Figure 9).

Our results mean that an initially formed dimer B can convert with nearly equal probability into dimer A or dimer C. Once C is formed, it can then undergo a transition to the global minimum A. Our TMD simulations predict a rather low free-energy barrier of 7.3 kJ/mol (the energy difference is 5.0 kJ/mol) for the C → A reaction. This is due to the first step of the C → A reaction being the apparent cleavage of a weak O–H...O(H)–C bond, followed by the rotation about the stronger O–H...O bond. The transition-state structure is shown in Figure 9; it is similar to the other two transition states as far as the quasi-perpendicular arrangement of the two molecules is concerned. However, it appears that actually a weak O–H...O(H)–C contact is maintained in the C → A reaction up to transition state. This could indeed contribute to the small barrier height for this process. By further rotation, the system becomes planar, and the second O–H...O bond and thus dimer A are formed.

The low C → A free-energy barrier explains why the C population in our aggregation simulations is totally depleted at 150 K (Figure 3). However, we have also carried out unconstrained simulations of dimer C at thermal equilibrium at a temperature of 100 K. During this run, a spontaneous conversion of C to A was observed, meaning that ultracold conditions are necessary to stabilize dimer C.

IV. Conclusions

We have performed first-principles molecular dynamics simulations of the dimerization process of formic acid mimicking the conditions in matrix isolation or helium droplet experiments. The aggregation product distributions obtained at the different classical simulation temperatures, 50, 100, and 150 K, suggest that the second most stable dimer structure B predominates at low temperature (50 K), while the global minimum structure A is the preferred species at 150 K. This is completely in line with the experimental evidence published to date. However, our theoretical data also predict the formation of significant concentrations of the less stable local minimum dimers C and D at low temperatures. We provide IR vibrational frequencies of all relevant isomers to guide the experimental search for these additional dimers.

Furthermore, our simulations offer detailed insight into the aggregation mechanism, that is, the dynamical pathway connecting the first contact structures to the final dimer products via a number of local potential minima. Our analysis also rationalizes why dimer B is the most favorable kinetic product. Interestingly, unlike in the case of formamide, the dipole–dipole intermolecular interactions do not play the decisive role in the nature of the first contact structure here. This is due to the comparatively small dipole moment of formic acid.

To investigate thermal rearrangement between different local minimum FAD structures, which typically occurs on a time scale beyond the 10 ps of our aggregation simulations, we have performed free-energy calculations using the targeted MD method. For the most frequent conversions, that is, $C \rightarrow A$, $B \rightarrow A$, and $B \rightarrow C$, we have studied the reaction mechanisms and the free-energy profiles associated. The free-energy barriers for the $B \rightarrow A$ and $B \rightarrow C$ reactions are both about 16 kJ/mol, while that for $C \rightarrow A$ is considerably lower (7 kJ/mol), suggesting that **C** can only be stabilized at ultracold conditions.

Acknowledgment. This project was funded by DFG within FOR 618. We are grateful for computer time at RWTH Aachen and the University of Dortmund.

References and Notes

- (1) Schneider, H.-J.; Yatsimirsky, A. K. *Principles and Methods in Supramolecular Chemistry*; Wiley: New York, 2000.
- (2) (a) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press: Oxford, U.K., 1999. (b) Desiraju, G. R. *Nature* **2001**, *412*, 397.
- (3) Madeja, F.; Havenith, M.; Nauta, K.; Miller, R. E.; Chocholousova, J.; Hobza, P. *J. Chem. Phys.* **2004**, *120*, 10554.
- (4) Chelli, R.; Righini, R.; Califano, S. *J. Phys. Chem. B* **2005**, *109*, 17006.
- (5) Bako, I.; Hutter, J.; Palinkas, G. *J. Phys. Chem. A* **2006**, *110*, 2188.
- (6) Nahringerbauer, I. *Acta Crystallogr., Sect. B* **1978**, *34*, 315.
- (7) Georges, R.; Freytes, M.; Hurtmans, D.; Kleiner, I.; Vander Auwera, J.; Herman, M. *Chem. Phys.* **2004**, *305*, 187.
- (8) Matyilitsky, V. V.; Riehn, C.; Gelin, M. F.; Brutschy, B. *J. Chem. Phys.* **2003**, *119*, 10553.
- (9) Gantenberg, M.; Halupka, M.; Sander, W. *Chem.—Eur. J.* **2000**, *6*, 1865.
- (10) Madeja, F.; Havenith, M.; Nauta, K.; Miller, R. E.; Chocholousova, J.; Hobza, P. *J. Chem. Phys.* **2004**, *120*, 10554.
- (11) Gora, R. W.; Grabowski, S. J.; Leszczynski, J. *J. Phys. Chem. A* **2005**, *109*, 6397.
- (12) Qian, W.; Krimm, S. *J. Phys. Chem. A* **2001**, *105*, 5046.
- (13) Senthilkumar, L.; Ghanty, T. K.; Ghosh, S. K.; Kolandaivel, P. *J. Phys. Chem. A* **2007**, *110*, 12623.
- (14) Brinkmann, N. R.; Tschumper, G. S.; Yan, G.; Schaefer, H. F., III. *J. Phys. Chem. A* **2003**, *107*, 10208.
- (15) Mardiyukov, A.; Sanchez-Garcia, E.; Rodziewicz, P.; Doltsinis, N. L.; Sander, W. *J. Phys. Chem. A* **2007**, *111*, 10552.
- (16) Schlitter, J.; Engels, M.; Krüger, P.; Jacobi, E.; Wollmer, A. *Mol. Simul.* **1993**, *10*, 291.
- (17) Markwick, P. R. L.; Doltsinis, N. L.; Marx, D. *J. Chem. Phys.* **2005**, *122*, 054112.
- (18) CPMD 3.10: Hutter J.; Ballone, P.; Bernasconi, M.; Focher, P.; Fois, E.; Goedecker, S.; Marx, D.; Parrinello, M.; Tuckerman M. *MPI für Festkörperforschung*, Stuttgart and IBM Zurich Research Laboratory.
- (19) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (20) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892.
- (21) Tuckerman, M. E.; Martyna, G. J. *J. Chem. Phys.* **1999**, *110*, 2810.
- (22) Gervasio, F. L.; Chelli, R.; Procacci, P.; Schettino, V. *Proteins* **2002**, *48*, 117.
- (23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision E.01; Gaussian, Inc.: Pittsburgh, PA, 2003.
- (24) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.
- (25) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
- (26) Melikova, S. M.; S Rutkowski, K.; Rodziewicz, P.; Koll, A. *Chem. Phys. Lett.* **2002**, *352*, 301.
- (27) Besley, N. A.; Doltsinis, N. L. *J. Chem. Theory Comput.* **2006**, *2*, 1598.

JP9007575