

Article

# Solvation and Aggregation of Meta-Aminobenzoic Acid in Water: Density Functional Theory and Molecular Dynamics Study

Etienne Gaines and Devis Di Tommaso \* 

School of Biological and Chemical Sciences, Materials Research Institute, Queen Mary University of London, Mile End Road, London E1 4NS, UK; e.gaines@qmul.ac.uk

\* Correspondence: d.ditommaso@qmul.ac.uk; Tel.: +44-207-882-6226

Received: 4 January 2018; Accepted: 19 January 2018; Published: 23 January 2018

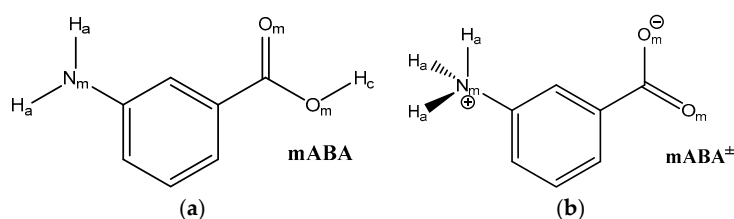
**Abstract:** Meta-aminobenzoic acid, an important model system in the study of polymorphism and crystallization of active pharmaceutical ingredients, exist in water in both the nonionic (mABA) and zwitterionic ( $\text{mABA}^\pm$ ) forms. However, the constituent molecules of the polymorph that crystallizes from aqueous solutions are zwitterionic. This study reports atomistic simulations of the events surrounding the early stage of crystal nucleation of meta-aminobenzoic acid from aqueous solutions. Ab initio molecular dynamics was used to simulate the hydration of  $\text{mABA}^\pm$  and mABA and to quantify the interaction of these molecules with the surrounding water molecules. Density functional theory calculations were conducted to determine the low-lying energy conformers of meta-aminobenzoic acid dimers and to compute the Gibbs free energies in water of nonionic,  $(\text{mABA})_2$ , zwitterionic,  $(\text{mABA}^\pm)_2$ , and nonionic-zwitterionic,  $(\text{mABA})(\text{mABA}^\pm)$ , species. Classical molecular dynamics simulations of mixed  $\text{mABA}$ – $\text{mABA}^\pm$  aqueous solutions were carried out to examine the aggregation of meta-aminobenzoic acid. According to these simulations, the selective crystallization of the polymorphs whose constituent molecules are zwitterionic is driven by the formation of zwitterionic dimers in solution, which are thermodynamically more stable than  $(\text{mABA})_2$  and  $(\text{mABA})(\text{mABA}^\pm)$  pairs. This work represents a paradigm of the role of molecular processes during the early stages of crystal nucleation in affecting polymorph selection during crystallization from solution.

**Keywords:** meta-aminobenzoic acid; solvation; aggregation; polymorphism; atomistic simulations

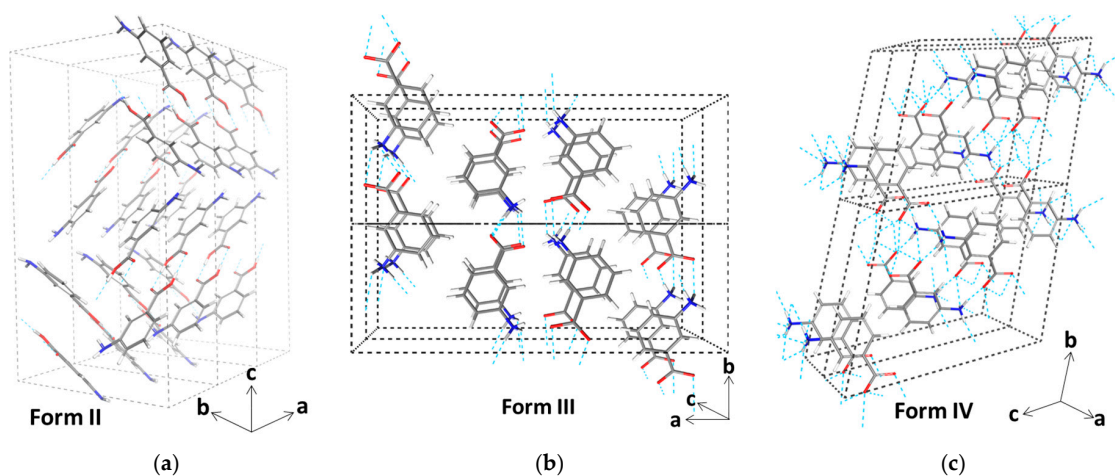
## 1. Introduction

The substance meta-aminobenzoic acid is of considerable importance in the pharmaceutical industry, widely used in the synthesis of analgesics, antihypertensives, vasodilators, and other drugs [1]. This molecule also represents a fascinating model system for polymorphic research because it can crystallize in five different crystal structures (I–V) [2]. The very strong polymorphic character of meta-aminobenzoic acid can be related to the manifold of inter-molecular interactions between meta-aminobenzoic acid molecules (hydrogen (H) bonding,  $\pi$ – $\pi$  interactions, and H– $\pi$  interactions) but also to the ability of this molecule to exist in either of both the nonionic (mABA) and zwitterionic ( $\text{mABA}^\pm$ ) forms (Figure 1) [3]. In fact, in the polymorphs denoted I, III, and V the molecules of meta-aminobenzoic acid are zwitterionic, and in the polymorphs II and V they are nonionic [2,4]. In Form II, two mABA molecules interact through the O–H $\cdots$ O acid dimer of an  $R_2^2(8)$  ring motif (Figure 2a). In Form III, the  $\text{mABA}^\pm$  molecules form ionic  $\text{N}^+–\text{H}\cdots\text{O}^-$  interactions in an  $R_4^4(8)$  ring motif (Figure 2b). In Form IV, two independent molecules form a linear C(7) chain through ionic  $\text{N}^+–\text{H}\cdots\text{O}^-$  interactions (Figure 2c). So far, the crystal structure of Form I has not been determined and the crystal structure of Form V shows disorder [2,4,5]. The nature of the

solvent can significantly influence the thermodynamics and kinetics of crystal growth [6–8] and, consequently, control the formation of one specific polymorph over another [9–11]. In the case of meta-aminobenzoic acid, Form II preferentially crystallizes from dimethyl sulfoxide (DMSO) [4], where meta-aminobenzoic acid only exist in the nonionic form. Hughes and co-workers [12] monitored the crystallization of meta-aminobenzoic acid from organosulfur solutions using a combined liquid- and solid-state in-situ NMR apparatus and proposed the existence of nonionic mABA aggregates linked by H bonds; the authors could not, however, uniquely determine the identity of these species. A recent theoretical study conducted in our group showed that mABA molecules in DMSO aggregate to form thermodynamically stable dimers and tetramers, whose structure is consistent with the classic carboxylic dimer  $\pi$ – $\pi$  stacking synthon found in this polymorph [8].



**Figure 1.** Schematic picture of the two tautomeric forms of meta-aminobenzoic acid: (a) nonionic mABA; (b) zwitterion  $\text{mABA}^\pm$ . The oxygen and nitrogen atoms of mABA or  $\text{mABA}^\pm$  are denoted by  $\text{O}_m$  and  $\text{N}_m$ , the hydrogen atoms of the amino group are denoted by  $\text{H}_a$ , and the hydrogen atoms of the carboxylic group are denoted by  $\text{H}_c$ .



**Figure 2.** Crystal structure of the polymorphs of meta-aminobenzoic acid denoted II, III, and IV: (a)  $(1 \times 3 \times 1)$  unit cell of Form II (neutral); (b)  $(1 \times 2 \times 2)$  unit cell of Form III (zwitterionic); (c)  $(2 \times 2 \times 1)$  unit cell of Form IV (zwitterionic) [2].

On the other hand, Form I preferentially crystallizes from aqueous environments [4], even though it has been reported that the values of the equilibrium constant  $K_Z = [\text{mABA}^\pm]/[\text{mABA}]$  for aminobenzoic acids are of the order of unity in water [13–15], implying a comparable distribution of  $\text{mABA}^\pm$  and mABA molecules. The fundamental details of factors controlling the selection between zwitterionic and nonionic forms of meta-aminobenzoic acid during crystal nucleation from aqueous solution are not yet known [12,16]. This work aims therefore to solve this conundrum by applying a combination of atomistic methods to follow the events surrounding the crystal nucleation of meta-aminobenzoic acid from aqueous solutions: *ab initio* molecular dynamics simulations of the hydration of  $\text{mABA}^\pm$  and mABA in water; density functional theory calculations of the structure and energetics of formation in water of  $(\text{mABA})_2$ ,  $(\text{mABA})(\text{mABA}^\pm)$ , and  $(\text{mABA}^\pm)_2$  dimers; classical

molecular simulations of mixed mABA–mABA<sup>±</sup> aqueous solutions to quantify the aggregation of meta-aminobenzoic acid.

## 2. Computational Methods

### 2.1. Density Functional Theory Calculations

Density functional theory (DFT) calculations were carried out with the NWChem (version 6.3, Valiev et al., Richland, WA, USA) [16] and Gaussian09 [17] codes (Frisch et al., Wallingford, CT, USA). The Grimme's density functional (B97-D) [18] and the Minnesota 06 global hybrid functional with 54% Hartree-Fock (HF) exchange (M06-2X) [19] were used together with the Gaussian 6-31+G(d,p) basis set. Free energies of solvation were calculated using the SMD solvation model [20].

The free energies of formation of nonionic, (mABA)<sub>2</sub>, nonionic-zwitterionic, (mABA)(mABA<sup>±</sup>), and zwitterionic, (mABA<sup>±</sup>)<sub>2</sub>, dimers were computed according to the following equation:

$$\Delta G_{ass}^* = G_{AB}^* - G_A^* - G_B^*. \quad (1)$$

In Equation (1),  $G_X^*$  is the total Gibbs free energy of the species  $X$  ( $X = AB, A$  or  $B$ ) in the liquid. This quantity was evaluated using two different approaches. The first one follows the recommendation by Ho et al. that free energies of molecules in solution should be obtained from separate gas- and solution-phase calculations [21]; the following expression was used to evaluate the Gibbs free energy of the species  $X$ :

$$G_X^* = E_{e,gas} + \delta G_{VRT,gas}^\circ + \Delta G_{solv}^* + RT \ln[\tilde{R}T]. \quad (2)$$

In Equation (2),  $E_{e,gas}$  is the gas-phase total electronic energy of the gas-phase optimized geometry of the species  $X$ ,  $\delta G_{VRT,gas}^\circ$  is the vibrational-rotational-translational contribution to the gas-phase Gibbs free energy at  $T = 298$  K under a standard-state partial pressure of 1 atm,  $\Delta G_{solv}^*$  is the solvation free energy of the solute corresponding to transfer from an ideal gas at a concentration of 1 mol·L<sup>-1</sup> to an ideal solution at a liquid-phase concentration of 1 mol·L<sup>-1</sup>, and the last term is the free energy change of 1 mol of an ideal gas from 1 atm to 1 mol·L<sup>-1</sup> ( $RT \ln[\tilde{R}T] = 1.89$  kcal·mol<sup>-1</sup> at 298 K,  $\tilde{R} = 0.082$  K<sup>-1</sup>) [22]. However, the gas-phase optimization of zwitterionic, (mABA<sup>±</sup>)<sub>2</sub>, and nonionic-zwitterionic, (mABA)(mABA<sup>±</sup>), dimers caused the H-transfer between molecular units (e.g., (mABA<sup>±</sup>)<sub>2</sub> → (mABA)<sub>2</sub>). In these instances, stationary points in the solution do not correspond to stationary points in the gas-phase, making it impossible to compute relevant gas-phase vibrational, translational, and rotational contributions ( $\delta G_{VRT,gas}^\circ$ ). The other approach adopted was to optimize the structures of (mABA<sup>±</sup>)<sub>2</sub>, (mABA)(mABA<sup>±</sup>), and of the monomers mABA and mABA<sup>±</sup>, in the aqueous phase; the following expression was then used to evaluate the free energy of the species:

$$G_X^* = E_{solv}^{Tot} + \delta G_{VRT,solv}^* \quad (3)$$

where  $\delta G_{VRT,solv}^*$  is the vibrational-rotational-translational contribution to the liquid-phase, and  $E_{solv}^{Tot}$  is given by the sum of the liquid-phase expectation value of the gas-phase Hamiltonian ( $E_{e,solv}$ ), the electronic polarization contribution to the solvation free energy based on bulk electrostatic ( $\Delta G_{EP}$ ), and the contribution from cavitation, dispersion, and solvent structural effects ( $G_{CDS}$ ):

$$E_{solv}^{Tot} = E_{e,solv} + \Delta G_{EP} + G_{CDS}. \quad (4)$$

The potential energy surface of a molecular cluster is characterized multiple low-lying energy isomers [23]. The free energy of the dimers (mABA)<sub>2</sub>, (mABA)(mABA<sup>±</sup>), and (mABA<sup>±</sup>)<sub>2</sub> was therefore determined from the Boltzmann ensemble average:

$$\langle G(X) \rangle = \sum_{i=1}^N f_i G(X_i) \quad (5)$$

where  $f_i$  is the Boltzmann factor corresponding to the  $i$ -th configuration,  $G(X_i)$  is the corresponding free energy, and  $N$  is the number of low-lying energy isomers. The Boltzmann factor was determined according to

$$f_i = \frac{e^{-G(X_i)/RT}}{\sum_j e^{-G(X_j)/RT}} \quad (6)$$

where  $R$  is the universal gas constant,  $T$  is the absolute temperature ( $T = 298$  K), and the index  $j$  runs over all isomers. The low-lying energy structures of the meta-aminobenzoic acid dimers were located using the following computational protocol: (1) For each type of dimer [(mABA)<sub>2</sub>, (mABA)(mABA<sup>±</sup>), and (mABA<sup>±</sup>)<sub>2</sub>], hundreds of thousands of candidate structures were generated using Granada (Montero et al., Havana, Cuba) [24,25], a code designed to distribute randomly one or more molecules around a central unit (a monomer, dimer, trimer, etc.) placed at the center of a cube of defined side length. (2) Configurations satisfying the condition that at least one atom of each mobile molecule was within 4 Å from at least one atom of the central unit were selected as potential low-lying energy structures. (3) The energies of these structures were evaluated at the B97-D/6-31+G(d,p) level of theory and the Boltzmann factor  $f_i$  corresponding to the  $i$ -th configuration was determined as

$$f_i = \frac{e^{-(E_i-E_0)/RT}}{\sum_j e^{-(E_j-E_0)/RT}} \quad (7)$$

where  $E_i$  was the energy of the  $i$ -th candidate structure, and  $E_0$  was the energy of the most stable candidate structure. (4) The candidate structures with a Boltzmann factor  $f_i \geq 0.01$  and 10–15 randomly selected structures such that  $3 \leq E_i - E_0 \leq 15$  kcal mol<sup>-1</sup> were selected. (5) Geometry optimization, thermochemical properties, and solvation energies of the selected configurations were computed at the M06-2X/6-31+G(d,p) level of theory.

## 2.2. Molecular Dynamics Simulations

Ab initio (Born-Oppenheimer) molecular dynamics (AIMD) simulations were conducted with the electronic structure code CP2K/Quickstep code, version 4.1 (Hutter et al., Zurich, Switzerland) [26]. CP2K implements density functional theory (DFT) based on a hybrid Gaussian plane wave. We used the PBE [27] generalized gradient approximation for the exchange and correlation terms together with the general dispersion correction termed DFT-D3. Goedecker–Teter–Hutter pseudopotentials [28] were used to describe the core–valence interactions. All atomic species were represented using a double-zeta valence polarized basis set. The plane wave kinetic energy cut off was set to 1000 Ry.  $k$ -sampling was restricted to the  $\Gamma$  point of the Brillouin zone. Simulations were carried out with a wave function optimization tolerance of 10<sup>-6</sup> au that allows for 1.0 fs time steps with reasonable energy conservation. Periodic boundary conditions were applied throughout. Simulations were carried out in the NVT ensemble using a Nosé–Hoover chain thermostat to maintain the average temperature at  $T = 300$  K.

Classical MD simulations were performed using version 5.0.4 of the GROMACS molecular dynamics package (van der Spoel et al., Uppsala, Sweden) [29]. The leapfrog algorithm with a time step of 2 fs was used to integrate the equations of motion. The isothermal–isobaric (constant NPT) ensemble was used to maintain a temperature of 300 K and a pressure of 1 bar. The velocity rescale thermostat and the isotropic Parrinello–Rahman barostat were used with 0.4 ps and 2.0 ps as the thermostat and barostat relaxation times, respectively. The electrostatic forces were calculated by means of the particle-mesh Ewald approach with a cutoff of 1.2 nm. The same cutoff was used for the van der Waals forces. The LINCS algorithm was applied at each step to preserve the bond lengths. The general AMBER forcefield (GAFF) [30] was used to model the nonionic and zwitterionic (mABA<sup>±</sup>) forms of meta-aminobenzoic acid; this family of forcefields has been previously used to compute the aggregation and crystal growth of organic molecules [8,31–33]. Water molecules were modeled using the SPC/E potential [34]. The interactions between mABA and mABA<sup>±</sup> molecules and between these molecules and water were described using the GAFF potential. To generate the GAFF parameters

for mABA and mABA<sup>±</sup>, the structure and molecular electrostatic potential of these molecules were computed using the HF method and the 6–31G\* basis set, and the *Antechamber* package was then used to compute partial charges according to the restrained electrostatic potential formalism. The GAFF forcefields and partial charges of mABA and mABA<sup>±</sup> are given in Tables S1 and S2.

Aqueous solutions of a single nonionic and a single zwitterionic meta-aminobenzoic acid molecule were carried out by embedding one mABA in a cubic box of 210 water molecules (0.26 mol·L<sup>-1</sup>) and one mABA<sup>±</sup> in a cubic box of 215 water molecules (0.25 mol·L<sup>-1</sup>). Classical MD simulations were first conducted for approximately 5 ns and the last snapshot was used to conduct 20 ps of ab initio MD simulations.

The insert-molecules utility of GROMACS was used to generate aqueous meta-aminobenzoic acid solutions of different concentrations by inserting equal amounts of mABA and mABA<sup>±</sup> molecules in an empty cubic box of size 5 nm. The solvate utility was then used to solvate the cubic boxes with SPC/E water. Each solution was at first minimized using the conjugate-gradient algorithm with a tolerance on the maximum force of 200 kJ·mol<sup>-1</sup>, and the temperature and volume of each system were equilibrated by running 100 ps of constant volume, constant temperature (NVT) simulation followed by 200 ns of NPT simulations. Analysis was conducted on the last 40 ns of simulation. Details of the simulation times, the number of solute and solvent molecules, equilibrated values of the average cell length (Table S3), and the convergence of the box cell volume during the period of equilibration are shown in Figures S1 and S2).

### 3. Results

#### 3.1. Intermolecular Properties and Hydration Structure

This section is concerned with the stability of the nonionic (mABA) and zwitterionic (mABA<sup>±</sup>) forms of meta-aminobenzoic acid in aqueous solution, and with the interaction of these molecules with the surrounding water molecules. Hereafter, the oxygen and nitrogen atoms of mABA or mABA<sup>±</sup> are denoted by O<sub>m</sub> and N<sub>m</sub>, the hydrogen of amino group are denoted by H<sub>a</sub>, the hydrogen atoms of carboxylic group are denoted by H<sub>c</sub>, and oxygen and hydrogen of water are denoted by O<sub>w</sub> and H<sub>w</sub>, respectively.

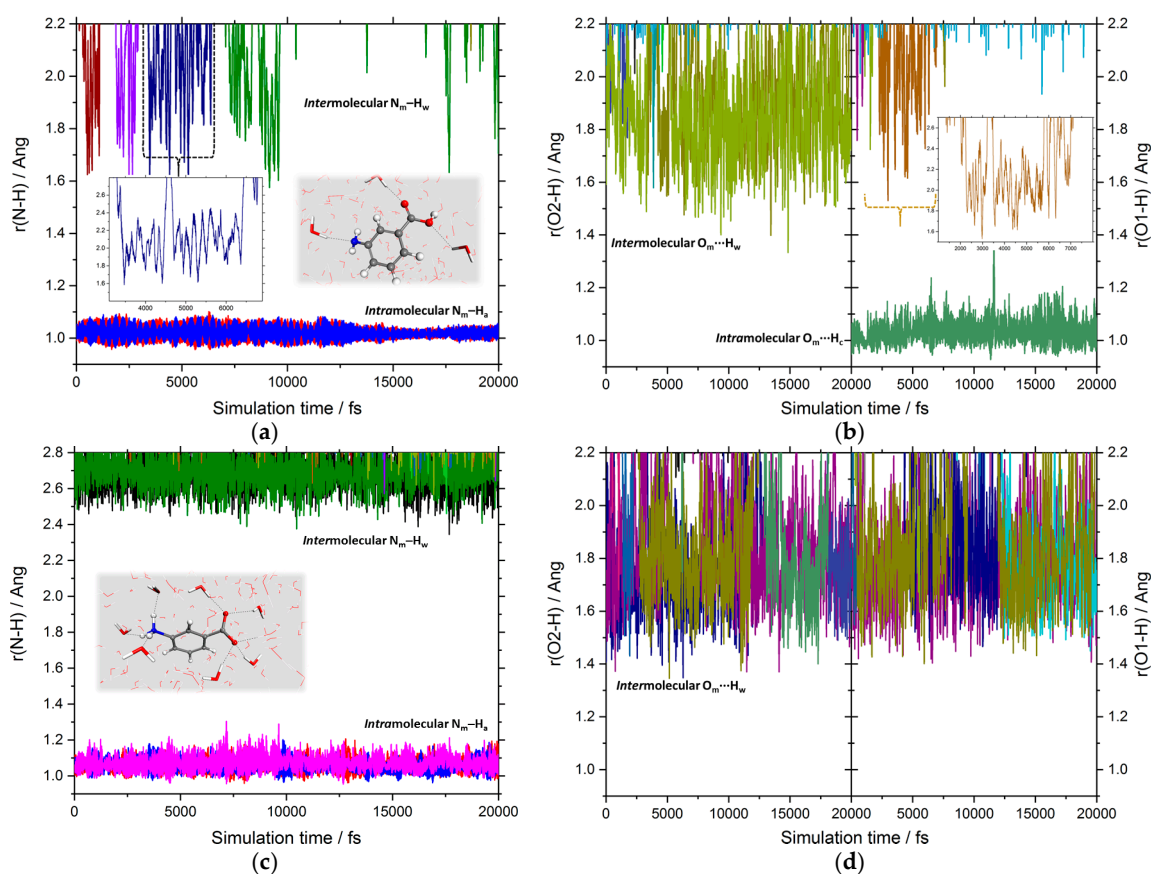
Figure 3 reports the time evolution of the intra- (O<sub>m</sub>–H<sub>c</sub> and N<sub>m</sub>–H<sub>a</sub>) and inter-molecular (O<sub>m</sub>···H<sub>w</sub> and N<sub>m</sub>···H<sub>w</sub>) distances during the AIMD simulations of the mABA and mABA<sup>±</sup> species in water. If 1 Å is taken as the average intramolecular X<sub>m</sub>–H (X = N, O) bond distance, then mABA and mABA<sup>±</sup> are not involved in any proton transfer reactions with the surrounding water molecules. Both mABA and mABA<sup>±</sup> molecules are therefore stable in water and should be considered when modeling the aggregation of meta-aminobenzoic acid in aqueous solution. If we use 2.5 Å to define the existence of intermolecular O<sub>m</sub>···H<sub>w</sub> and N<sub>m</sub>···H<sub>w</sub> interactions, then, as the insets of Figure 3a,b show, the interaction of mABA with the surrounding water molecules occurs during a very short time range (<5 ps).

A detailed characterization of intermolecular O<sub>m</sub>···H<sub>w</sub> and N<sub>m</sub>···H<sub>w</sub> interactions can be obtained from the analysis of the radial distribution function (RDF),  $g_{\alpha\beta}(r)$ , which represents the probability relative to a random distribution of finding an atom of type  $\beta$  at a distance  $r$  from an atom of type  $\alpha$ . Figure 4 reports the O<sub>m</sub>–H<sub>w</sub> and N<sub>m</sub>–H<sub>w</sub> RDFs together with the running coordination number,  $n(r) = (4\pi N/V) \int_0^r g(r') dr'$ , where  $N$  is the number of hydrogen or oxygen atoms and  $V$  is the volume of the simulation cell. In the X<sub>m</sub>–H<sub>w</sub> (X = N or O) RDFs, a maximum in the [1.5–2.0] Å region and a minimum at around 2.5 Å indicates the presence of an H-bond with the surrounding water molecules [35]. On average, less than one water molecule is coordinated to each oxygen atom of the –COOH group and to the nitrogen atom of the –NH<sub>2</sub> group. On the other hand, approximately four water molecules are coordinated to the –COO<sup>-</sup> group of mABA<sup>±</sup> and no water molecule is H-bonded to the nitrogen atom of –NH<sub>3</sub><sup>+</sup>. Table 1 summarizes the positions ( $r_{\max}$  and  $r_{\min}$ ) and amplitudes ( $g_{\max}$  and  $g_{\min}$ ) of the maxima and minima of the X<sub>m</sub>–H<sub>w</sub> (X = N, O) RDFs together with

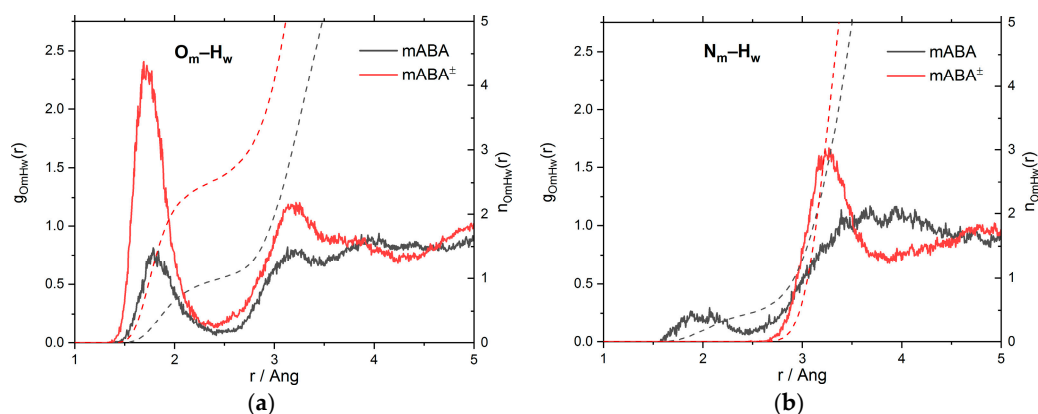


the ratios  $g_{max}^{X_m-H_w} / g_{min}^{X_m-H_w}$ , and these values can be used as a proxy for the strength of the H-bonding interactions between the  $X_m H_w$  pairs ( $X = O, N$ ) [35,36]. For mABA, the  $g_{max}^{O_m-H_w} / g_{min}^{O_m-H_w}$  ratio of the carboxyl oxygen atoms (9.0) is higher than that of nitrogen (4.5) but lower than the value of  $g_{max}^{O_w-H_w} / g_{min}^{O_w-H_w} = 19.6$  obtained from AIMD simulations of pure water. Similar behavior is observed for mABA $^\pm$ , but the interaction of the COO $^-$  group ( $g_{max}^{O_w-H_w} / g_{min}^{O_w-H_w} = 14.0$ ) is significantly stronger than mABA.

The RDFs and structural data of the H $_c$ -O $_w$  and H $_a$ -O $_w$  intermolecular interactions are reported in Figure 5 and Table 2. For the carboxylic group of mABA, the H $_c$ -O $_w$  RDF has a very well-defined maximum at 1.51 Å, and the running coordination number ( $n_w^{H_c}$ ) is characterized by a clear plateau at the first RDF minimum (Figure 5a). The value of  $g_{max}^{H_c-O_w} / g_{min}^{H_c-O_w}$  is significantly larger than  $g_{max}^{O_w-H_w} / g_{min}^{O_w-H_w}$  of pure water (19.6), so the H $_c$ -O $_w$  interaction is stronger than the intermolecular H-bonding in bulk water. The hydrogen of -COOH is therefore stably coordinated to a single water molecule. For the amino group of mABA, the hydrogen atoms of the -NH $_2$  group do not interact significantly with the surrounding water molecules because, in the [1.5–2.0] Å region, the H $_a$ -O $_w$  RDF is not characterized by a well-defined peak (Figure 5b). On the other hand, the H $_a$ -O $_w$  RDF of the -NH $_3^+$  in mABA $^\pm$  is characterized by a distinct peak at 1.77 Å.



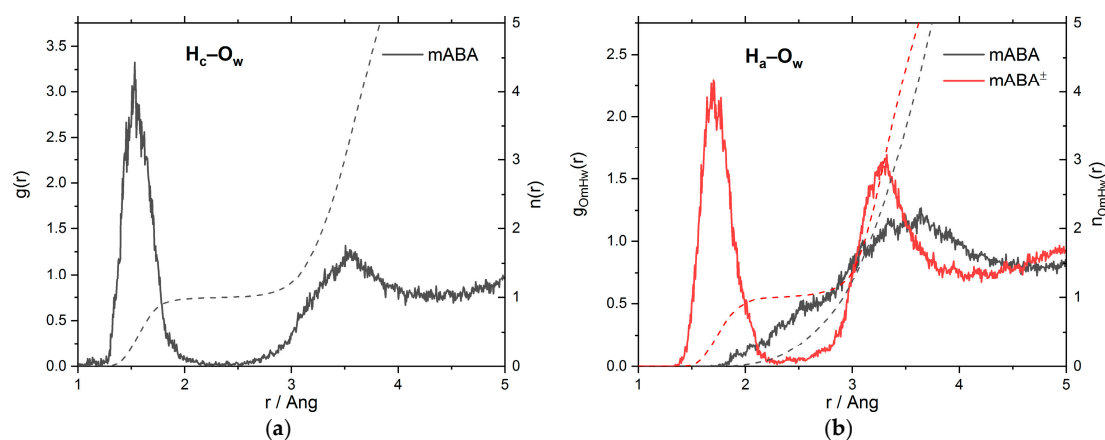
**Figure 3.** Time evolution of the  $X_m$ -H ( $X = N, O$ ) distances during the AIMD simulation of the nonionic (mABA) and the zwitterionic (mABA $^\pm$ ) forms of meta-aminobenzoic acid in water: (a) intramolecular ( $N_m$ -H) and intermolecular ( $N_m \cdots H_w$ ) distances of the mABA molecule; (b) intramolecular ( $O_m$ -H) and intermolecular ( $O_m \cdots H_w$ ) distances of the mABA molecule; (c) intramolecular ( $N_m$ -H) and intermolecular ( $N_m \cdots H_w$ ) distances of the mABA $^\pm$  molecule; (d) intramolecular ( $O_m$ -H) and intermolecular ( $O_m \cdots H_w$ ) distances of the mABA $^\pm$  molecule.



**Figure 4.** The radial distribution functions,  $g(r)$ , and running coordination numbers,  $n(r)$ , of mABA and mABA $^{\pm}$  with water obtained from AIMD simulations: (a)  $O_m-H_w$  RDFs ( $O_m$  = oxygen atoms of meta-aminobenzoic acid;  $H_w$  = hydrogen atoms of water); (b)  $N_m-H_w$  RDFs ( $N_m$  = nitrogen atoms of meta-aminobenzoic acid;  $H_w$  = oxygen atoms of water).

**Table 1.** Positions ( $r_{max}^{X-H}$  and  $r_{min}^{X-H}$  in Å) and amplitudes ( $g_{max}^{X-H}$  and  $g_{min}^{X-H}$ ) of the maxima and minima of the first peak of the  $X_m-H_w$  ( $X = O_m, N_m$ ) RDFs, and first hydration shell numbers ( $n_w$ ) obtained from the AIMD simulations of mABA and mABA $^{\pm}$  in water.

	mABA	mABA $^{\pm}$
$r_{max}^{O_m-H_w}$	1.79	1.72
$g_{max}^{O_m-H_w}$	0.81	2.38
$r_{min}^{O_m-H_w}$	2.50	2.52
$g_{min}^{O_m-H_w}$	0.09	0.17
$g_{max}^{O_m-H_w} / g_{min}^{O_m-H_w}$	9.00	14.00
$n_w^{O_m}$	1.0	2.6
$r_{max}^{N_m-H_w}$	1.88	-
$g_{max}^{N_m-H_w}$	0.27	-
$r_{min}^{N_m-H_w}$	2.46	-
$g_{min}^{N_m-H_w}$	0.06	-
$g_{max}^{N_m-H_w} / g_{min}^{N_m-H_w}$	4.50	-
$n_w^{N_m}$	0.5	0



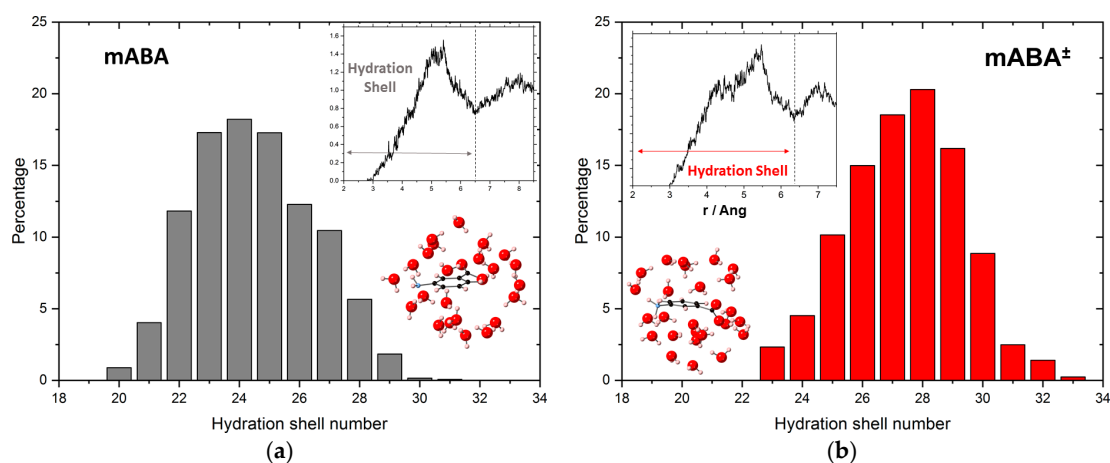
**Figure 5.** The radial distribution functions,  $g(r)$ , and running coordination numbers,  $n(r)$ , of mABA and mABA $^{\pm}$  with water obtained from AIMD simulations: (a)  $H_c-O_w$  RDFs ( $O_c$  = oxygen atoms of the carboxylic group of mABA;  $H_w$  = hydrogen atoms of water); (b)  $H_a-O_w$  RDFs ( $N_m$  = nitrogen atoms of the amino group of mABA and mABA $^{\pm}$ ;  $O_w$  = oxygen atoms of water).

To summarize, the analysis of the  $X_m-H_w$  ( $X = N, O$ ),  $H_c-O_w$  and  $H_a-O_w$  RDFs indicates that in aqueous solution the  $mABA^\pm$ -water interaction is stronger than  $mABA$ -water. Moreover, the interaction of both species with the surrounding water molecules is stronger around the carboxylic acid than around the amino group.

**Table 2.** Positions ( $r_{max}^{H-O}$  and  $r_{min}^{H-O}$  in Å) and amplitudes ( $g_{max}^{H-O}$  and  $g_{min}^{H-O}$ ) of the maxima and minima of the first peak of the  $H_a-O_w$  and  $H_c-O_w$  RDFs, and first hydration shell numbers ( $n_w$ ) obtained from the AIMD simulations of  $mABA$  and  $mABA^\pm$  in water.

	mABA	mABA <sup>±</sup>
$r_{max}^{H_c-O_w}$	1.51	-
$g_{max}^{H_c-O_w}$	3.06	-
$r_{min}^{H_c-O_w}$	2.31	-
$g_{min}^{H_c-O_w}$	0.01	-
$g_{max}^{H_c-O_w} / g_{min}^{H_c-O_w}$	306.00	-
$n_w^{H_c}$	1.0	-
$r_{max}^{H_a-O_w}$	-	1.77
$g_{max}^{H_a-O_w}$	-	2.15
$r_{min}^{H_a-O_w}$	-	2.23
$g_{min}^{H_a-O_w}$	-	0.03
$g_{max}^{H_a-O_w} / g_{min}^{H_a-O_w}$	-	71.7
$n_w^{H_a}$	0	1.0

The probability distribution of the number of water molecules in the first hydration shell (HS) of  $mABA$  and  $mABA^\pm$  was determined from the pair correlation functions between the center-of-mass (COM) of meta-aminobenzoic acid and the COM of the water molecules (Figure 6). The position of the first HS was approximated by the first minimum in the COM( $mABA$ )-COM( $H_2O$ ) RDFs (insets of Figure 6), and although a hydration shell can be located for both molecules, the probability distributions of the number of water molecules surrounding  $mABA$  and  $mABA^\pm$  show the flexibility of their HS as there are, on average, 24 water molecules in the HS of  $mABA$  with a mean absolute deviation (MAD) of 1.4, and 27 water molecules in the HS of  $mABA^\pm$  with an MAD of 1.0.



**Figure 6.** (a) Probability distribution of the coordination number in the hydration shell of  $mABA$ , the  $mABA$ - $H_2O$  radial distribution function of the center-of-masses of  $mABA$  and water (inset), and the optimized structure of  $mABA$  with its hydration shell. (b) Probability distribution of the coordination number in the hydration shell of  $mABA^\pm$ , the  $mABA^\pm$ - $H_2O$  radial distribution function of the center-of-masses of  $mABA^\pm$  and water, and the optimized structure of  $mABA^\pm$  with its hydration shell.



### 3.2. Dimerization of Meta-Aminobenzoic Acid

Stable dimers in solution have often been linked to the structural synthon found in the crystal polymorph that crystallizes from solution [37,38]. This section reports therefore results from extensive DFT calculations to determine the structure and the thermodynamic stability in water of dimers of meta-aminobenzoic acid. The Boltzmann averaged energetics of formation of the nonionic, (mABA)<sub>2</sub>, zwitterionic, (mABA<sup>±</sup>)<sub>2</sub>, and nonionic-zwitterionic, [(mABA)(mABA<sup>±</sup>)], dimers are reported in Table 3. The free energy of formation of (mABA)<sub>2</sub> ranges from −0.1 to 2.4 kcal·mol<sup>−1</sup>, depending on the method used to compute the total free energies of the dimers and monomer in water. The formation of (mABA)(mABA<sup>±</sup>) (2.4 kcal·mol<sup>−1</sup>) is also endergonic. On the other hand, the dimerization free energy of the zwitterionic dimer (mABA<sup>±</sup>)<sub>2</sub> is large and negative (−5.8 kcal·mol<sup>−1</sup>).

**Table 3.** Energetics of dimerization of meta-aminobenzoic acid:  $\Delta E_{e,gas}$  is the gas phase interaction energy;  $\Delta G_{ass}^{\circ}$  is the standard state (1 atm) gas-phase association free energy at 298 K;  $\Delta G_{ass}^*$  is the standard state (1 mol·L<sup>−1</sup>) free energy of reactions in the liquid-phase. Calculations conducted at the M06-2X/6-31+G(d,p) level of theory using the SMD solvation model. Values obtained from the Boltzmann average of the energies, or free energies, of the isomers of nonionic (mABA)<sub>2</sub>, zwitterionic (mABA<sup>±</sup>)<sub>2</sub>, and mixed (mABA)(mABA<sup>±</sup>) dimers. Values in kcal·mol<sup>−1</sup>.

Reaction	$\Delta E_{e,gas}$	$\Delta G_{ass}^{\circ}$	$\Delta G_{ass}^*$
2 mABA → (mABA) <sub>2</sub>	−18.3	−6.6	−0.1 <sup>1</sup> 2.4 <sup>2</sup>
mABA + mABA <sup>±</sup> → (mABA)(mABA <sup>±</sup> )			1.3 <sup>2</sup>
2 mABA <sup>±</sup> → (mABA <sup>±</sup> ) <sub>2</sub>	−	−	−5.8 <sup>2</sup>

<sup>1</sup> Gas-phase optimized geometries and free energies in water obtained using Equation (2); <sup>2</sup> Solution-phase optimized geometries and free energies in water obtained using Equation (3).

Figure 7 reports the structures of the thermodynamically most stable (mABA)<sub>2</sub> and (mABA)(mABA<sup>±</sup>) species in water. The (mABA)<sub>2</sub> dimer corresponds to the structural synthon found in Form II [2], where the two nonionic meta-aminobenzoic acid molecules interact through a double H-bond to form a classic carboxylic dimer (Figure 7a). In the (mABA)(mABA<sup>±</sup>) dimer, the two monomers are arranged to maximize the concomitant H-bonding and  $\pi$ – $\pi$  interactions (Figure 7b). All other (mABA)<sub>2</sub> and (mABA)(mABA<sup>±</sup>) dimeric structure have significantly higher free energies of formation in water ( $2.5 \text{ kcal}\cdot\text{mol}^{-1} < \Delta G_{ass}^* < 10 \text{ kcal}\cdot\text{mol}^{-1}$ ), so they are very unstable in aqueous solution. On the other hand, several stable zwitterionic dimers, (mABA<sup>±</sup>)<sub>2</sub>, were found in solution (Figure 7c). Therefore, even though the distribution between zwitterions and nonionic molecules in water is close to unity [13–15], the selective crystallization of the polymorphs that only contain zwitterionic molecules (Forms I, III, and V) could be driven by the higher stability in water of zwitterionic (mABA<sup>±</sup>)<sub>2</sub> dimers.

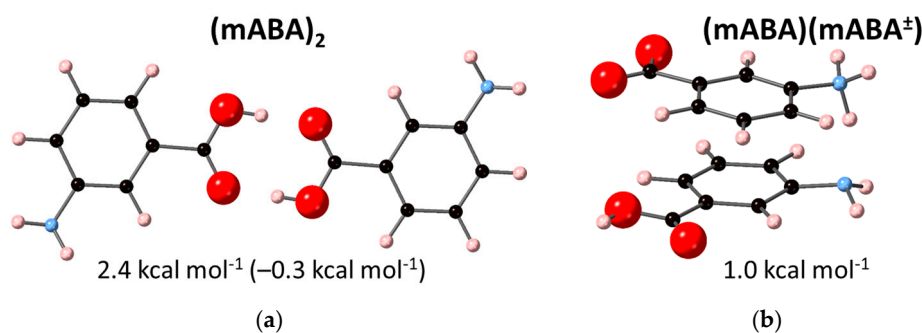
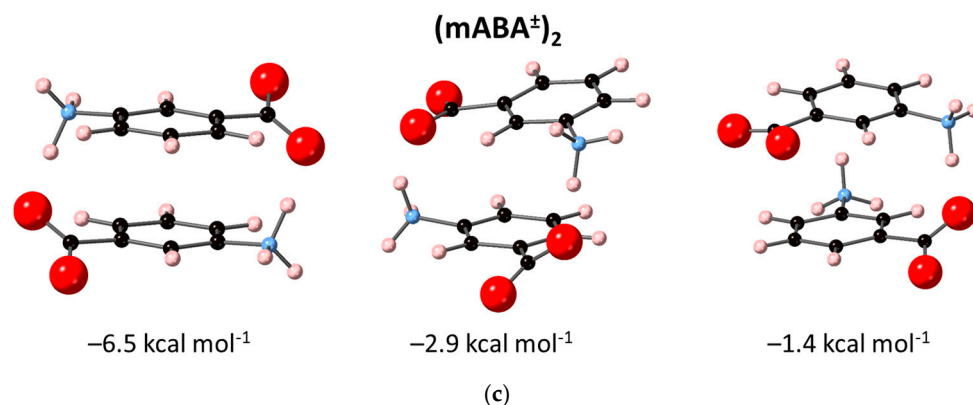


Figure 7. Cont.



**Figure 7.** Optimized structures of most stable meta-aminobenzoic acid dimers in water: (a) nonionic (mABA)<sub>2</sub> dimer (in parenthesis value obtained using the gas-phase optimized geometries of (mABA)<sub>2</sub> and mABA); (b) nonionic–zwitterionic (mABA)(mABA<sup>±</sup>) dimer; (c) zwitterionic (mABA<sup>±</sup>)<sub>2</sub> dimer. Beneath the structure is reported free energy of dimer formation in water.

### 3.3. Molecular Aggregation in Mixed mABA–mABA<sup>±</sup> Aqueous Solutions

Classical MD simulations ( $\geq 200$  ns) of mixed mABA–mABA<sup>±</sup> aqueous solutions were conducted to examine the aggregation behavior of meta-aminobenzoic acid as a function of concentration. Four concentrations were considered: 0.04 mol·L<sup>-1</sup>, 0.08 mol·L<sup>-1</sup>, 0.16 mol·L<sup>-1</sup>, and 0.31 mol·L<sup>-1</sup>. Svård et al. (2010) reported crystallization experiments of meta-aminobenzoic acid at saturated solution. At 300 K, the solubility in water of polymorph I is 5.4 g·L<sup>-1</sup> [1], whereas for the other polymorphs higher solubility values have been reported: 7.8 g·L<sup>-1</sup> for Form II, 6.07 g·L<sup>-1</sup> for Form III, and 6.25 g·L<sup>-1</sup> for Form III [5]. The 0.04 mol·L<sup>-1</sup> solution (5.3 g·L<sup>-1</sup>) corresponds therefore to conditions just below the solubility limit of Form I, while the others simulated systems (10.8 g·L<sup>-1</sup>, 21.5 g·L<sup>-1</sup>, and 42.6 g·L<sup>-1</sup>) correspond to increasingly supersaturated solutions with respect to all polymorphs of meta-aminobenzoic acid. Representative configurations of these solutions are reported in Figure 8, where the number of molecular aggregates that form in solution increases as a function of solute concentration. This aggregation process has been quantified in terms of the number of (mABA⋯mABA), (mABA<sup>±</sup>⋯mABA<sup>±</sup>), and (mABA⋯mABA<sup>±</sup>) pairs within 4.0 Å (Figure 9 and Figure S3). The number of molecular pairs increases with the concentration but the number of nonionic clusters is significantly higher than mixed and zwitterionic species. As the dehydration of the molecules of solute is a crucial step during crystal nucleation from solution [39], the stronger interaction of mABA<sup>±</sup> with the surrounding water molecules discussed in Section 3.1 could explain the observed different level of aggregation of nonionic and zwitterionic species in water. Moreover, a close view of the clusters formed during the MD simulations reveals that meta-aminobenzoic acid interact via a manifold of inter-molecular interactions: H-bonding X–H⋯X (X = O or N) between amino (NH<sub>2</sub> and NH<sub>3</sub><sup>+</sup>) and carboxylic (COOH and COO<sup>-</sup>) groups,  $\pi$ – $\pi$  interactions between benzene (C<sub>6</sub>H<sub>4</sub>) groups, and X–H⋯ $\pi$  interactions.

To characterize these interactions, a three-body simplified representation of the nonionic mABA (A–B–C) and zwitterionic mABA<sup>±</sup> (A\*–B\*–C\*) molecules has been adopted (Figure 10), where A and A\* represent the center-of-masses of –NH<sub>2</sub> and –NH<sub>3</sub><sup>+</sup>, B and B\* represent the center-of-masses of the benzene (C<sub>6</sub>H<sub>4</sub>) groups, and C and C\* represent the center-of-masses of –COOH and –COO<sup>-</sup>.

A symmetric pairwise interaction matrix (PIM) can therefore be used to quantify the interactions between (A–B–C) and (A\*–B\*–C\*):

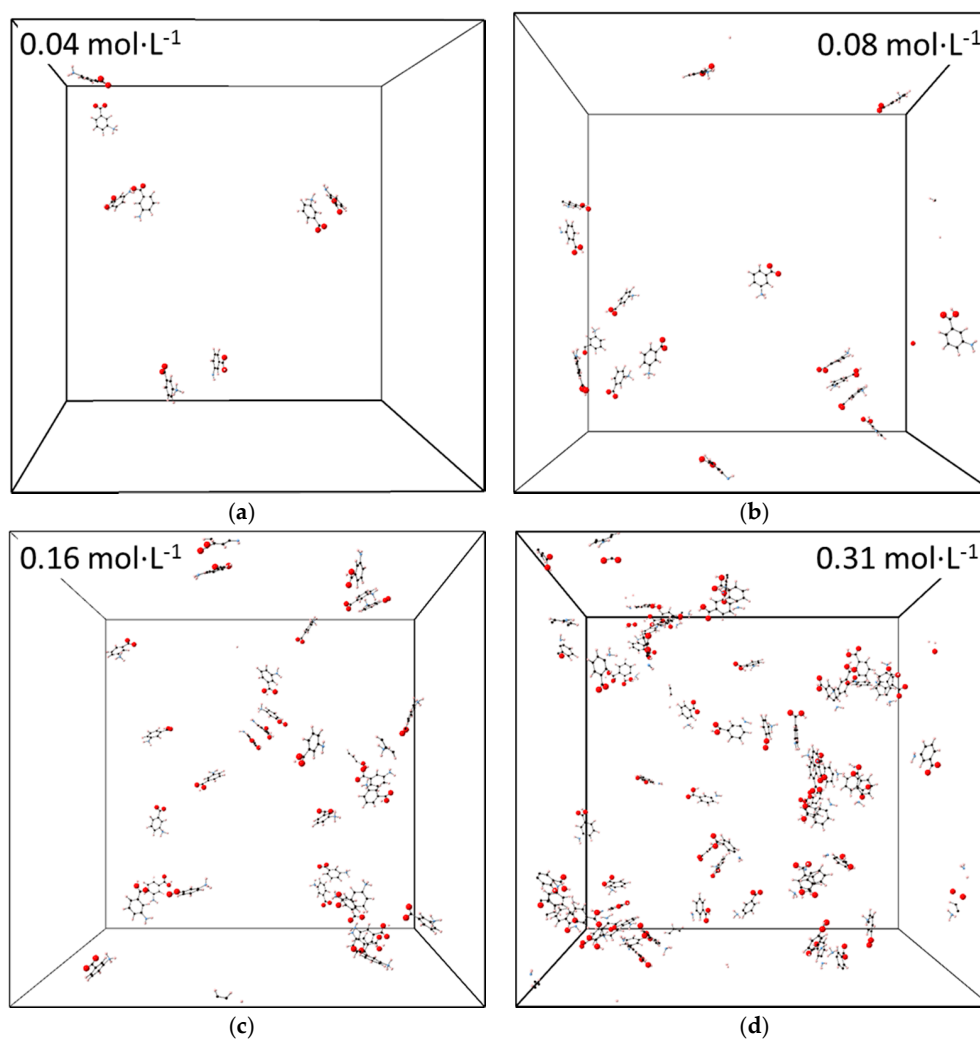
$$PIM = \begin{cases} p_{A^*A^*} & p_{A^*B^*} & p_{A^*C^*} & p_{A^*A} & p_{A^*B} & p_{A^*C} \\ & p_{A^*B^*} & p_{B^*C^*} & p_{B^*A} & p_{B^*B} & p_{B^*C} \\ & & p_{C^*C^*} & p_{C^*A} & p_{C^*B} & p_{C^*C} \\ & & & p_{AA} & p_{AB} & p_{AC} \\ & & & & p_{BB} & p_{BC} \\ & & & & & p_{CC} \end{cases} \quad (8)$$

In Equation (6), the elements of the PIM matrix are defined as

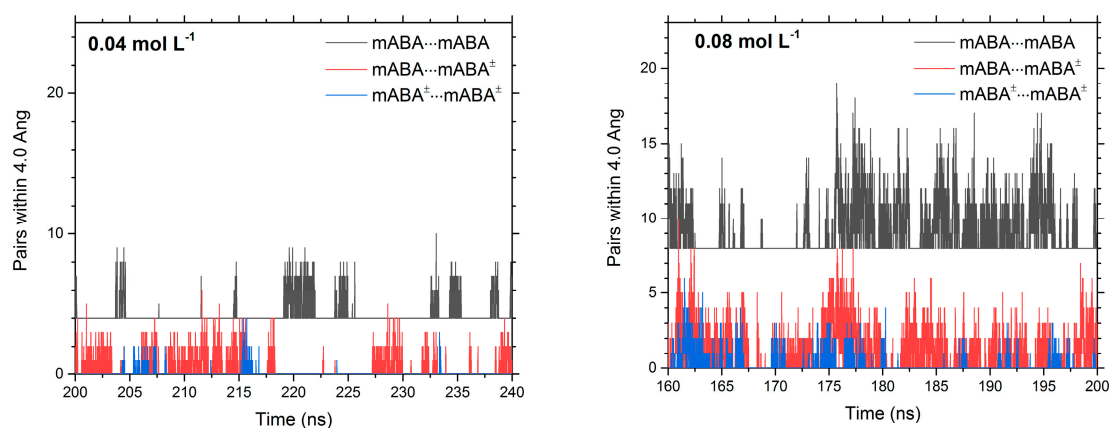
$$p_{ij} = \left\langle \sum_i \sum_{i>j} f(r_{ij}) \right\rangle \quad (9)$$

where the pairwise interaction function  $f(r_{ij})$  quantifies the existence of a  $(i, j)$  pair within a cutoff distance of 4.0 Å:

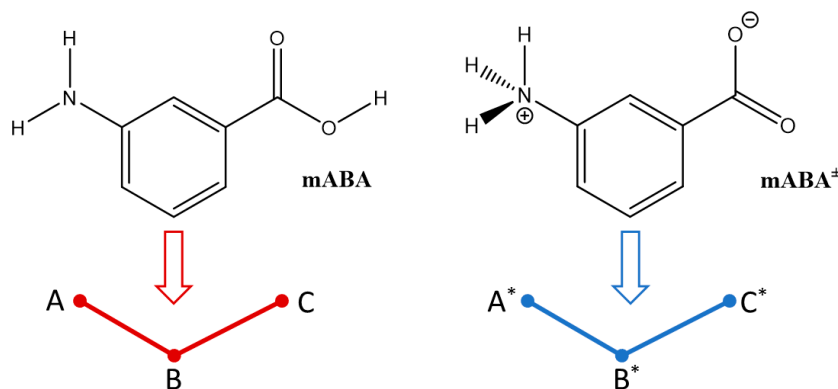
$$f(r_{ij}) = \begin{cases} 0, & r_{ij} > 4.0 \text{ \AA} \\ 1, & r_{ij} < 4.0 \text{ \AA} \end{cases} \quad (10)$$



**Figure 8.** Configuration at 200 ns of mixed mABA–mABA<sup>±</sup> aqueous solutions: (a) 0.04 mol·L<sup>−1</sup> aqueous solution; (b) 0.09 mol·L<sup>−1</sup> aqueous solution; (c) 0.16 mol·L<sup>−1</sup> aqueous solution; (d) 0.31 mol·L<sup>−1</sup>. Water molecules have been removed. The grey outlines represent the cubic simulation box.



**Figure 9.** Time evolution of the number of pairs between meta-benzoic acid molecules in mixed  $\text{mABA–mABA}^{\pm}$  aqueous solutions computed during the last 40 ns of the MD simulations.



**Figure 10.** Three-body representations ( $A\text{--}B\text{--}C$ ) and ( $A^*\text{--}B^*\text{--}C^*$ ) of the nonionic,  $\text{mABA}$ , and zwitterionic,  $\text{mABA}^{\pm}$  forms of meta-aminobenzoic acid:  $A$  and  $A^*$  are the center-of-masses (COMs) of the  $\text{–NH}_2$  and  $\text{–NH}_3^+$  groups,  $B$  and  $B^*$  are the COMs of the benzene ( $\text{C}_6\text{H}_4$ ) group, and  $C$  and  $C^*$  are the COMs of the  $\text{–COOH}$  and  $\text{–COO}^-$  groups.

The cutoff value of  $4.0 \text{ \AA}$  was based on the analysis of the intermolecular distances between amino ( $\text{NH}_2$  and  $\text{NH}_3^+$ ), carboxylic ( $\text{COOH}$  and  $\text{COO}^-$ ), and benzene ( $\text{C}_6\text{H}_4$ ) groups in the most thermodynamically stable  $(\text{mABA})_2$ ,  $(\text{mABA})(\text{mABA}^{\pm})$ , and  $(\text{mABA}^{\pm})_2$  dimers in water (Figure S4). For example, the element  $p_{AA}$  corresponds to  $\text{COOH}\cdots\text{COOH}$  interactions found in the classic carboxylic dimer  $(\text{mABA})_2$  (Figure 7a), the elements  $p_{A^*A}$  and  $p_{C^*C}$  correspond to the  $\text{COO}^-\cdots\text{COOH}$  and  $\text{NH}_3^+\cdots\text{NH}_2$  interactions in the nonionic-zwitterionic dimer  $(\text{mABA})(\text{mABA}^{\pm})$  (Figure 7b), and the elements  $p_{B^*B}$  and  $p_{A^*C^*}$  correspond to  $\pi\cdots\pi$  and  $\text{NH}_3^+\cdots\text{COO}^-$  interacting pairs in the structures of the most stable zwitterionic dimers  $(\text{mABA}^{\pm})_2$  (Figure 7c). For the mixed  $0.08 \text{ mol}\cdot\text{L}^{-1}$   $\text{mABA–mABA}^{\pm}$  aqueous solutions, the pairwise interaction matrix in Table 4 reveals a higher proportion of  $\text{NH}_3^+\cdots\text{COO}^-$  ( $A^*\cdots C^* = 8.7\%$ ) and  $\pi\cdots\pi$  ( $B^*\cdots B^* = 9.1\%$ ) pairs than  $\text{COOH}\cdots\text{COOH}$  ( $C\cdots C = 6.5\%$ ),  $\text{COO}^-\cdots\text{COOH}$  ( $C^*\cdots C = 6.5\%$ ) and  $\text{NH}_3^+\cdots\text{NH}_2$  ( $A^*\cdots A = 5.3\%$ ). Very similar PIM matrices were obtained from the calculation of the three-body pairwise interactions of the other systems (Tables S4–S6). This analysis implies that aqueous solutions of meta-aminobenzoic acid contain a higher proportion of stable zwitterionic  $(\text{mABA}^{\pm})_2$  pairs, in agreement with the DFT calculations of dimerization free energies.

**Table 4.** Matrix elements  $p_{ij}$  of the pairwise interaction matrix for the mixed 0.08 mol·L<sup>-1</sup> mABA–mABA<sup>±</sup> aqueous solutions. Values of  $p_{ij}$  expressed as percentage.

	A*	B*	C*	A	B	C
A*	0.2	0.6	8.7	5.3	3.4	3.5
B*		9.1	2.7	2.6	10.0	7.6
C*			0.1	3.6	2.3	3.6
A				4.3	4.2	5.2
B					6.1	10.6
C						6.5

#### 4. Conclusions

The solvation and aggregation of the nonionic (mABA) and zwitterionic (mABA<sup>±</sup>) forms of meta-aminobenzoic acid in water were investigated by means of atomistic simulations.

Ab initio molecular dynamics of two aqueous solutions containing one mABA and one mABA<sup>±</sup> molecules in approximately 200 water molecules were conducted to determine the stability, intermolecular and hydration properties of these two species. A detailed analysis of the number and strength of hydrogen bonds of mABA and mABA<sup>±</sup> with the surrounding water molecules shows that the mABA<sup>±</sup>–water interaction is stronger than mABA–water, and that the interaction with the surrounding water molecules is stronger around the carboxylic acid than around the –NH<sub>2</sub> (mABA) and –NH<sub>3</sub><sup>+</sup> (mABA<sup>±</sup>) groups. Although a coordination shell can be located for both molecules, the probability distributions of the number of water molecules surrounding mABA and mABA<sup>±</sup> show a great degree of flexibility of the hydration environment.

Density functional theory calculations with a polarizable continuum model to describe the aqueous environment were used to locate the low-lying energy structures and thermodynamic stability in water of nonionic, (mABA)<sub>2</sub>, zwitterionic, (mABA<sup>±</sup>)<sub>2</sub> and nonionic-zwitterionic, (mABA)(mABA<sup>±</sup>), dimers. Results show that the only thermodynamically dimers in solution are (mABA<sup>±</sup>)<sub>2</sub>, whereas the formation of the nonionic classic carboxylic dimer (mABA)<sub>2</sub> and the π–π stacked (mABA)(mABA<sup>±</sup>) dimer is endoergonic.

Classical molecular dynamics simulations of meta-aminobenzoic acid aqueous solutions containing an equal amount of nonionic and zwitterionic species were conducted to examine the aggregation behavior as a function of concentration of solute. Analysis of the aggregates formed during the simulation shows a higher proportion of π···π and NH<sub>3</sub><sup>+</sup>···COO<sup>-</sup> pairs, whose interactions occur in the most stable zwitterionic dimers (mABA<sup>±</sup>)<sub>2</sub> located using DFT calculations.

According to the atomistic simulations reported in this work the selective crystallization of the polymorphs of meta-aminobenzoic acid whose constituent molecules are zwitterionic is driven by the higher stability of zwitterionic dimers in solution.

This work represents a paradigm of the role of molecular processes during the early stages of crystal nucleation in affecting polymorph selection during crystallization from solution.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4923/10/1/12/s1>; Table S1: General AMBER forcefield parameters used to model mABA in GROMACS; Table S2: General AMBER forcefield parameters used to model mABA<sup>±</sup> in GROMACS; Table S3: Details of molecular dynamics simulation; Table S4: Matrix elements  $p_{ij}$  of the pairwise interaction matrix for the mixed 0.04 mol·L<sup>-1</sup> mABA–mABA<sup>±</sup> aqueous solutions; Table S5: Matrix elements  $p_{ij}$  of the pairwise interaction matrix for the mixed 0.16 mol·L<sup>-1</sup> mABA–mABA<sup>±</sup> aqueous solutions; Table S6: Matrix elements  $p_{ij}$  of the pairwise interaction matrix for the mixed 0.31 mol·L<sup>-1</sup> mABA–mABA<sup>±</sup> aqueous solutions; Figure S1: Convergence of the volume of the two simulation boxes containing one mABA and one mABA<sup>±</sup> in water; Figure S2: Convergence of the volume of the simulation boxes containing mixed mABA–mABA<sup>±</sup> solutions; Figure S3: Time evolution of the number of pairs between meta-benzoic acid molecules in mixed mABA–mABA<sup>±</sup> aqueous solutions computed during the last 40 ns of the MD simulations: (a) 0.16 mol·L<sup>-1</sup>; (b) 0.31 mol·L<sup>-1</sup>; Figure S4: Intermolecular distances between the amino (NH<sub>2</sub> and NH<sub>3</sub><sup>+</sup>), carboxylic (COOH and COO<sup>-</sup>), and benzene (C<sub>6</sub>H<sub>4</sub>) groups in the most thermodynamically stable (mABA)<sub>2</sub>, (mABA)(mABA<sup>±</sup>) and (mABA<sup>±</sup>)<sub>2</sub> dimers in water.



**Acknowledgments:** Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202), the researchers made use of the ARCHER UK National Supercomputing Service (<http://www.archer.ac.uk>). Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT, was also utilized. <http://doi.org/10.5281/zenodo.438045>.

**Author Contributions:** Etienne Gaines and Devis Di Tommaso conceived and designed the experiments; Etienne Gaines and Devis Di Tommaso performed the experiments; Etienne Gaines and Devis Di Tommaso analyzed the data; Devis Di Tommaso and Etienne Gaines wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Palafox, M.A.; Gill, M.; Núñez, J.L. Meta-aminobenzoic acid: Structures and spectral characteristics. *Spectrosc. Lett.* **1996**, *29*, 609–629. [[CrossRef](#)]
2. Williams, P.A.; Hughes, C.E.; Lim, G.K.; Kariuki, B.M.; Harris, K.D. Discovery of a new system exhibiting abundant polymorphism: M-aminobenzoic acid. *Cryst. Growth Des.* **2012**, *12*, 3104–3113. [[CrossRef](#)]
3. Théorêt, A. Structure moléculaire et zwitterion des acides aminés—I. Spectres infrarouges des acides o, m et p-aminobenzoïques dans différentes formes cristallines. *Spectrochim. Acta Part A Mol. Spectrosc.* **1971**, *27*, 11–18. [[CrossRef](#)]
4. Svärd, M.; Nordström, F.L.; Jasnobulka, T.; Rasmuson, Å.C. Thermodynamics and nucleation kinetics of m-aminobenzoic acid polymorphs. *Cryst. Growth Des.* **2010**, *10*, 195–204. [[CrossRef](#)]
5. Kumar, S.S.; Nangia, A. A solubility comparison of neutral and zwitterionic polymorphs. *Cryst. Growth Des.* **2014**, *14*, 1865–1881. [[CrossRef](#)]
6. Lahav, M.; Leiserowitz, L. The effect of solvent on crystal growth and morphology. *Chem. Eng. Sci.* **2001**, *56*, 2245–2253. [[CrossRef](#)]
7. Ter Horst, J.H.; Geertman, R.M.; van Rosmalen, G.M. The effect of solvent on crystal morphology. *J. Cryst. Growth* **2001**, *230*, 277–284. [[CrossRef](#)]
8. Gaines, E.; Maisuria, K.; Di Tommaso, D. The role of solvent in the self-assembly of m-aminobenzoic acid: A density functional theory and molecular dynamics study. *CrystEngComm* **2016**, *18*, 2937–2948. [[CrossRef](#)]
9. Blagden, N.; Davey, R.J. Polymorph selection: Challenges for the future? *Cryst. Growth Des.* **2003**, *3*, 873–885. [[CrossRef](#)]
10. Musumeci, D.; Hunter, C.A.; McCabe, J.F. Solvent effects on acridine polymorphism. *Cryst. Growth Des.* **2010**, *10*, 1661–1664. [[CrossRef](#)]
11. Kitamura, M.; Umeda, E.; Miki, K. Mechanism of solvent effect in polymorphic crystallization of BPT. *Ind. Eng. Chem. Res.* **2012**, *51*, 12814–12820. [[CrossRef](#)]
12. Hughes, C.E.; Williams, P.A.; Harris, K.D.M. "CLASSIC NMR": An In-Situ NMR Strategy for Mapping the Time-Evolution of Crystallization Processes by Combined Liquid-State and Solid-State Measurements. *Angew. Chem. Int. Ed.* **2014**, *53*, 8939–8943. [[CrossRef](#)] [[PubMed](#)]
13. Kumler, W.D. Acidic and basic dissociation constants and structure. *J. Org. Chem.* **1955**, *20*, 700–706. [[CrossRef](#)]
14. Cohn, E.J.; Edsall, J.T. *Proteins, Amino Acids and Peptides*; Reinhold Publishing Corporation: New York, NY, USA, 1943.
15. Bjerrum, N.Z. Die Konstitution der Ampholyte, besonders der Aminosäuren, und ihre Dissoziationskonstanten. *Phys. Chem.* **1923**, *104*, 147–173.
16. Valiev, M.; Bylaska, E.J.; Govind, N.; Kowalski, K.; Straatsma, T.P.; Van Dam, H.J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.L.; et al. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489. [[CrossRef](#)]
17. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian Software*; Gaussian, Inc.: Wallingford, CT, USA, 2009.
18. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799. [[CrossRef](#)] [[PubMed](#)]
19. Zhao, Y.; Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acta* **2008**, *120*, 215–241.

20. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [[CrossRef](#)] [[PubMed](#)]
21. Ho, J.; Klamt, A.; Coote, M.L. Comment on the correct use of continuum solvent models. *J. Phys. Chem. A* **2010**, *114*, 13442–13444. [[CrossRef](#)] [[PubMed](#)]
22. Ribeiro, R.F.; Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. The solvation, partitioning, hydrogen bonding, and dimerization of nucleotide bases: A multifaceted challenge for quantum chemistry. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10908–10922. [[CrossRef](#)] [[PubMed](#)]
23. Do, H.; Besley, N.A. Structural optimization of molecular clusters with density functional theory combined with basin hopping. *J. Chem. Phys.* **2012**, *137*, 134106. [[CrossRef](#)] [[PubMed](#)]
24. Montero, L.A.; Esteva, A.M.; Molina, J.; Zapardiel, A.; Herna, L.; Màrquez, H.; Acosta, A. A theoretical approach to analytical properties of 2, 4-diamino-5-phenylthiazole in water solution. Tautomerism and dependence on pH. *J. Am. Chem. Soc.* **1998**, *120*, 12023–12033. [[CrossRef](#)]
25. Sanchez-Garcia, E.; Studentkowski, M.; Montero Luis, A.; Sander, W. Noncovalent Complexes between Dimethyl Ether and Formic Acid—An Ab Initio and Matrix Isolation Study. *ChemPhysChem* **2005**, *6*, 618–624. [[CrossRef](#)] [[PubMed](#)]
26. Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25. [[CrossRef](#)]
27. Perdew, J.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [[CrossRef](#)] [[PubMed](#)]
28. Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710. [[CrossRef](#)]
29. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [[CrossRef](#)] [[PubMed](#)]
30. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)] [[PubMed](#)]
31. Salvalaglio, M.; Perego, C.; Giberti, F.; Mazzotti, M.; Parrinello, M. Molecular-dynamics simulations of urea nucleation from aqueous solution. *Proc. Natl. Acad. Sci. USA* **2014**, *112*, E6–E14. [[CrossRef](#)] [[PubMed](#)]
32. Salvalaglio, M.; Giberti, F.; Parrinello, M. 1,3,5-Tris (4-bromophenyl) benzene prenucleation clusters from metadynamics. *Acta Crystallogr. Sect. C Struct. Chem.* **2014**, *70*, 132–136. [[CrossRef](#)] [[PubMed](#)]
33. Toroz, D.; Hammond, R.B.; Roberts, K.J.; Harris, S.; Ridley, T. Molecular dynamics simulations of organic crystal dissolution: The lifetime and stability of the polymorphic forms of para-amino benzoic acid in aqueous environment. *J. Cryst. Growth* **2014**, *401*, 38–43. [[CrossRef](#)]
34. Berendsen, H.J.C.; Grigera, J.R.; Straatsma, T.P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271. [[CrossRef](#)]
35. Gageot, M.-P.; Sprik, M. Ab initio molecular dynamics study of uracil in aqueous solution. *J. Phys. Chem. B* **2004**, *108*, 7458–7467. [[CrossRef](#)]
36. Tang, E.; Di Tommaso, D.; de Leeuw, N.H. Hydrogen transfer and hydration properties of  $H_nPO_4^{3-n}$  ( $n = 0-3$ ) in water studied by first principles molecular dynamics simulations. *J. Chem. Phys.* **2009**, *130*, 234502. [[CrossRef](#)] [[PubMed](#)]
37. Davey, R.J.; Dent, G.; Mughal, R.K.; Parveen, S. Concerning the relationship between structural and growth synthons in crystal nucleation: Solution and crystal chemistry of carboxylic acids as revealed through IR spectroscopy. *Cryst. Growth Des.* **2006**, *6*, 1788–1796. [[CrossRef](#)]
38. Du, W.; Cruz-Cabeza, A.J.; Woutersen, S.; Davey, R.J.; Yin, Q. Can the study of self-assembly in solution lead to a good model for the nucleation pathway? The case of tolfenamic acid. *Chem. Sci.* **2015**, *6*, 3515–3524. [[CrossRef](#)]
39. Khamar, D.; Zeglinski, J.; Mealey, D.; Rasmuson, Å.C. Investigating the Role of Solvent–Solute Interaction in Crystal Nucleation of Salicylic Acid from Organic Solvents. *J. Am. Chem. Soc.* **2014**, *136*, 11664–11673. [[CrossRef](#)] [[PubMed](#)]

