The Astrophysical Journal, 951:32 (17pp), 2023 July 1 https://doi.org/10.3847/1538-4357/accae8

Theoretical Water Binding Energy Distribution and Snowline in Protoplanetary Disks

Lorenzo Tinacci[1,2], Aurèle Germain[2], Stefano Pantaleone[2], Cecilia Ceccarelli[1], Nadia Balucani[3], and Piero Ugliengo[2]
[1] Univ. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France; cecilia.ceccarelli@univ-grenoble-alpes.fr
[2] Dipartimento di Chimica and Nanostructured Interfaces and Surfaces (NIS) Centre, Università degli Studi di Torino, via P. Giuria 7, I-10125 Torino, Italy
[3] Dipartimento di Chimica, Biologia e Biotecnologie, Università di Perugia, I-06123 Perugia, Italy
Received 2022 December 15; revised 2023 March 10; accepted 2023 March 30; published 2023 June 29

Abstract

Water is one of the most important and abundant molecules in star-forming regions. In protoplanetary disks, where planets and comets form, H_2O is in a gas or solid form, depending on the dust temperature, i.e., the distance from the center and its binding energy (BE). Not surprisingly, several experimental and theoretical studies of the H_2O BE have been published. We report new ab initio calculations carried out on a large model of interstellar ice, where we identified 144 different adsorption sites. The BE associated with those sites ranges between 14.2 kJ mol^{-1} (1705 K) and 61.6 kJ mol^{-1} (7390 K). The distribution of the computed BEs as a function of BE follows a Gaussian peaked at 35.4 kJ mol^{-1} (4230 K) with a standard deviation of 9.7 kJ mol^{-1} (1160 K). The computed pre-exponential factor (ν) ranges between 9×10^{12} and 6×10^{14} s^{-1}. We evaluated the impact of the newly calculated BE and ν distributions on the snowline of a generic protoplanetary disk. We found that the region where water is frozen onto the ice is much smaller (a factor of 10 smaller radius) than that computed with the single BE (5600 K) and ν (2×10^{12} s^{-1}) values commonly adopted by astrochemical models. Besides, ∼10% of water remains frozen in relatively warm (∼150 K) regions, where the single BE and ν model would predict a full release of the ice in the gas phase. This last aspect may have an impact on the quantity trapped in the planetesimals eventually forming rocky planets.

Unified Astronomy Thesaurus concepts: Astrochemistry (75); Interstellar molecules (849)

1. Introduction

Water is the second most abundant molecule in the universe after H_2, together with CO. It is "*easily*" formed both in the gas phase (Elitzur & de Jong 1978; Langer & Graedel 1989; Ceccarelli et al. 1996; Kaufman & Neufeld 1996; Hollenbach et al. 2009) and, with a higher efficiency, on the surface of the submicron dust grains that permeate the interstellar medium (ISM; Tielens & Hagen 1982; Oba et al. 2009; Dulieu et al. 2010; Lamberts et al. 2016; Molpeceres et al. 2019; see also the review by van Dishoeck et al. 2013).

Water vapor is observed in large quantities, with abundances up to about 2×10^{-4} with respect to H_2, in a wide variety of warm ($T \geqslant 50$ K) objects: galactic hot cores and hot corinos, innermost regions of protoplanetary disks, protostellar shocks, nearby galaxies (e.g., Cernicharo et al. 1994; Combes & Wiklind 1997; Ceccarelli et al. 1999; Fischer et al. 1999; González-Alfonso et al. 2004; Herpin et al. 2012; Kristensen et al. 2012; Imanishi et al. 2022), and the farthest galaxies at $z \geqslant 3$ (e.g., Omont et al. 2013; Yang et al. 2016, 2020; Pensabene et al. 2022; see also van Dishoeck et al. 2021). Water vapor is also observed, but with a much lower abundance ($\leqslant 10^{-8}$), in cold ($T \leqslant 50$ K) regions such as molecular clouds, prestellar cores, and the outer protoplanetary disks (Melnick & Bergin 2005; Caselli et al. 2010; Hogerheijde et al. 2011; Podio et al. 2013). Finally, large quantities of frozen water, up to about 2×10^{-4} in abundance (as in warm objects), are observed in cold objects. The frozen water is believed to constitute icy mantles that envelop the interstellar dust grains, prevalently in amorphous water surfaces (e.g., Leger et al. 1979; Gibb et al. 2004; Boogert et al. 2015). In summary, water is copiously present in the molecular ISM, and particularly in the regions forming new solar-like planetary systems, with a rather constant abundance, $\sim 2 \times 10^{-4}$, when considering its two different forms, gaseous and solid.

Probably it is not by chance that water is a crucial molecule in terrestrial life, due to its multiple roles in the emergence and sustainability of life (e.g., Westall & Brack 2018) and probably its relatively large abundance in the progenitor of the solar system, whose analogs are the presently forming planetary systems (e.g., Ceccarelli & Du 2022). However, whether water was incorporated, in what quantity, and when in the newly formed Earth are still debated issues (e.g., Morbidelli et al. 2019). One way to answer these questions is to understand the fate of the water synthesized during the formation of solar-like planetary systems (e.g., Morbidelli et al. 2016; Hartmann et al. 2017). Since rocky planets, as well as asteroids and comets, which may later in the evolution bring water to the newly formed planet, are formed by the aggregation of interstellar dust grains, the above questions translate into how much of the interstellar water remains frozen on the dust grain mantles during the aggregation process. In this respect, the crucial parameter that determines whether water is in the gaseous or solid state is, along with the temperature, its binding energy (BE).

In protoplanetary disks, where planets, asteroids, and comets are formed, the distance from the star where water freezes out and becomes a solid is called the "*snowline*" (e.g., Hartmann et al. 2017). Therefore, the snowline is a crucial parameter in all models aiming to explain the origin of water on Earth and clearly depends on the assumed water BE distribution. The larger the BE, the larger is the region of the protoplanetary disk where water remains frozen around the dust grains and, therefore, is passed to the nascent planets, asteroids, and

comets. On the contrary, low-water BEs would allow a smaller fraction of water to be inherited by the planets, asteroids, and comets from the protoplanetary disk matter.

Finally, it is worth mentioning that a crucial parameter in the models focused on the terrestrial water origin is represented by its deuteration, namely the fraction of singly (or doubly) deuterated water with respect to the fully hydrogen-bearing one, HDO/H_2O (or D_2O/H_2O; e.g., Alexander 2017). The Earth's oceans have an HDO/H_2O ratio about 10 times that of the original elemental D/H ratio (Lécuyer et al. 1998). There is ample consensus that this enhanced water deuteration has been inherited from the first phases of the solar system formation (e.g., Ceccarelli et al. 2014). It is, therefore, of interest to understand the difference in the BE distribution of HDO with respect to that of H_2O.

While several experimental studies of the water BE have been carried out for more than 30 yr (see Section 2.1), theoretical studies on (very) small water surfaces have been published only in the last five years (see Section 2.2). In this work, we report new ab initio calculations of the water BE distribution, obtained on a simulated icy grain composed of 200 water molecules, which is, in turn, based on a specifically developed open access computer code, called ACO-FROST[4] (Germain et al. 2022a). To the best of our knowledge, this is the first theoretical study of the water BEs on such large water amorphous ice, which allows us to derive a BE distribution and not only a single value. We have already applied the new grain model and method used here to the adsorption of ammonia and found that its BE distribution presents two peaks. While the first one around 34 kJ mol^{-1} ($\sim$4000 K) was already known from experimental and theoretical studies, the second smaller peak at about 15 kJ mol^{-1} ($\sim$1800 K) had never been detected before (Tinacci et al. 2022). It is, therefore, of major interest to understand whether the water BE distribution has a similar behavior.

The article is organized as follows. After reviewing the literature of the previous water BE estimates (Section 2), we describe in detail the adopted methodology (Section 3) and the results (Section 4) of our simulations. We then discuss the comparison of our new BEs with respect to the previous studies (Section 5) and the astrophysical implications of the newly computed water BE distribution on the snowline of protoplanetary disks (Section 6), and conclude the article (Section 7). The data of our study are publicly available online (Section 8).

2. Review of the Published Water Binding Energy

2.1. Experimental Studies

Generally, in experimental studies, water molecules are condensed on an inert cold finger and then the rate of desorption upon heating the sample is measured. This technique is called temperature-programmed desorption (TPD). The thermal desorption of a species bound to a substrate can be approximately described by the Polanyi–Wigner equation (Kolasinski 2002), where the rate of desorption $k(T)$ is given by

$$k(T) = -\frac{dN}{dT} = N^n \underbrace{\nu \exp\left(-\frac{E_{\rm des}}{T}\right)}_{k_{\rm thdes}}, \quad (1)$$

where N is the number density of molecules adsorbed on the surface, n is the order of the desorption, $k_{\rm thdes}$ is the thermal desorption rate constant, ν is the pre-exponential factor (also called prefactor), T is the temperature of the surface, and $E_{\rm des}$ is the activation energy for desorption, which is approximately BE.

Experiments differ in the substrate (the cold finger), how the iced sample is condensed and treated (namely, the form of the ice), how the desorption carried out in time (the temperature ramp) is monitored, and how the $E_{\rm des}$ is extracted from the desorption curve. It is important to emphasize that, regardless of the method, BE is evaluated always together with the prefactor.

The substrates used are numerous and have a relatively low impact on the derived BE if more than one monolayer (ML) of water molecules (1 ML meaning that the substrate is completely covered by water molecules) is deposited on the substrate, as the water molecules will be desorbed from other water molecules. For a coverage inferior to 1 ML, the BEs could be impacted depending on the wetting of the substrate, i.e., if the water molecules are evenly spread over the substrate or if they coagulate and form multilayered clusters as sometime reported (Collings et al. 2015). On the other hand, when the coverage is larger then 1 ML, the BE correlates with the coverage regime, as pointed out by Rosu-Finsen et al. (2022). In their article, besides the suggested value report in our Table 1, these authors report a BE of 50.3 kJ mol^{-1} and a prefactor of 1×10^{30} cm^{-2} s^{-1} for a water film thickness of 53 nm, and a BE of 69.2 kJ mol^{-1} and a prefactor of 1×10^{35} cm^{-2} s^{-1} for a 101 μm film.

In Table 1, we indicate the different properties of the ice samples only if they were explicitly reported in the relevant articles. For samples where "*crystalline*" or "*amorphous*" is not reported, we could have assumed one or the other, depending on the way water was deposited on the sample. However, one should be aware that amorphous water ice undergoes a phase transition to crystalline ice at around 140–150 K (Jenniskens et al. 1995; Löfgren et al. 2003). This temperature almost coincides with the peak desorption temperature of water, which implies that the ice sample could, at that point, be amorphous, crystalline, or a mixture of the two. It is worth noticing that a multilayer case is the most common one in Table 1. This means that a substantial fraction of the experimental BE is due to lateral water–water interaction, not accounted for by our computational model, which considers the adsorption of an isolated water molecule.

Finally, different authors use either the zero or first order of the Polanyi–Wigner equation.

Table 1 summarizes the different estimates of the water BE values along with the different characteristics of the experiments, as briefly described above. Globally, the measured water BE ranges between 34 and 64 kJ mol^{-1}, where this range includes a variety of ice samples, multilayer regimes, and substrates, not to mention different ways of extracting the BE value (zero or first order) and prefactors. When considering only the experiments on amorphous multilayer samples, which are the closest to our simulations, where the BE and the prefactor are estimated at the first order, the measured BE range is smaller, between 40 and 47 kJ mol^{-1} (or $\sim$4800 to 5640 K), with a prefactor between $\sim 10^{12}$ and $\sim 10^{15}$ s^{-1}, with the outsider value of 64.2 kJ mol^{-1} (7721 K) obtained by Rosu-Finsen et al. (2022).

[4] https://github.com/aurelegermain/ACO-FROST_grain_generator

Table 1
Values of the Water BE and Prefactor as Measured in Different Experiments Using Different Substrates, Ice Samples, and Methods, Described in the Main Text

BE	Prefactor	Substrate (c)	Ice Sample (d)	Method (e)	Reference
40.0 ± 0.1 (4815 ± 15)	2×10^{12} (b)	CsI	Amorphous unannealed multilayer	IR	1
42.2 ± 0.4 (5070 ± 50)	2×10^{12} (b)	CsI	Crystalline annealed multilayer	IR	1
49.8 ± 0.8 (5988 ± 101)	$(2.8 \pm 1.0) \times 10^{30}$ (a)	Sapphire	Crystalline multilayer	OIM	2
43.4 ± 2.9 (5222 ± 348)	1.82×10^{21} (a)	Au	Crystalline multilayer	QCM	3
43.4 ± 2.9 (5222 ± 348)	Not indicated	Graphite (0001)	Crystalline multilayer	HREELS	4
48.3 ± 1.0 (5809 ± 120)	$(3.99 \pm 0.8) \times 10^{15}$ (b)	Ru (001)	Crystalline multilayer	QMS	5
46.9 ± 0.9 (5640 ± 108)	$(3.26 \pm 0.7) \times 10^{15}$ (b)	Ru (001)	Amorphous multilayer	QMS	5
48.0 ± 0.5 (5773 ± 60)	$1 \times 10^{30 \pm 2}$ (a)	Au	Crystalline multilayer	QMS	6
46.6 (5600)	$1 \times 10^{30 \pm 2}$ (a)	Au	Amorphous multilayer	QMS	6
58.2 ± 0.8 (6995 ± 101)	$1 \times 10^{32.6 \pm 0.3}$ (a)	Ru (001)	Crystalline multilayer	OIM	7
39.9 ± 2.9 (4799 ± 349)	4×10^{26} (a)	HOPG	Crystalline multilayer	QMS	8
46.0 ± 3.0 (5533 ± 361)	$9 \times 10^{14 \pm 1}$ (b)	HOPG	Unspecified ML	QMS	9
46.0 ± 3.0 (5533 ± 361)	$9 \times 10^{14 \pm 1}$ (b)	HOPG	Unspecified multilayer	QMS	9
39.9 ± 0.8 (4799 ± 96)	$1 \times 10^{27 \pm 1}$ (a)	HOPG	Crystalline multilayer	QMS	10
39.9 (4800)	1×10^{12} (b)	Amorphous silicate	Unspecified sub-ML	QMS	11
42.9 ± 0.5 (5165 ± 55)	$32.1^{+12.8}_{-8.5} \times 10^{27}$ (a)	KBr	Amorphous multilayer	QMS	12
34-36 (4090–4330)	1×10^{12} (b)	Amorphous silica	Unspecified sub-ML	QMS	13
49.3 ± 2 (5930 ± 240)	1×10^{28} (a)	Amorphous silica	Unspecified multilayer	QMS	13
48.7 ± 0.3 (5857 ± 36)	1×10^{12} (a)	KBr	Unspecified multilayer	QCM	14
45.3 ± 1.9 (5454 ± 232)	Not indicated	Au	Unspecified multilayer	d-ABS	15
64.2 ± 0.3 (7721 ± 36)	1×10^{17} (a)	Cu	Amorphous multilayer	QCM	16

Notes. BEs are reported in kilojoules per mole and kelvin (in parenthesis). (a) This refers to zero-order desorption, in units of mol cm^{-2} s^{-1}. (b) This refers to first-order desorption, in units of s^{-1}. (c) "HOPG" stands for Highly Oriented Pyrolytic Graphite. (d) The ice phase can be amorphous crystalline or unspecified, if it is not explicitly indicated in the cited articles for ML or multilayer regimes. (e) Infrared spectroscopy (IR); optical interference method (OIM); Quartz Crystal Microbalance and its resonance frequency (QCM); high-resolution electron energy loss spectroscopy (HREELS); quadrupole mass spectrometer (QMS); and direct-absorption millimeter/submillimeter (d-ABS).

References. (1) Sandford & Allamandola (1988); (2) Haynes et al. (1992); (3) Sack & Baragiola (1993); (4) Chakarov et al. (1995); (5) Speedy et al. (1998); (6) Fraser et al. (2001); (7) Smith et al. (2003); (8) Bolina et al. (2005); (9) Ulbricht et al. (2006); (10) Brown & Bolina (2007); (11) Dulieu et al. (2013); (12) Martín-Doménech et al. (2014); (13) Collings et al. (2015); (14) Potapov et al. (2018); (15) Yocum et al. (2019); and (16) Rosu-Finsen et al. (2022).

2.2. Theoretical Studies

In theoretical studies, the way the icy grain surface is modeled and the method used to compute the BE is of high importance. In Table 2, we report the different values of the water BE obtained from theoretical studies. The computed BE ranges from 22 to 70 kJ mol^{-1}. Contrary to experimental studies, computational studies of water are scarce. In Wakelam et al. (2017) the icy grain surface is modeled using a cluster composed of one water molecule and the BE value is computed using M06-2X coupled to an aug-cc-pVTZ basis set. They derive a water BE equal to 38 kJ mol^{-1}. Evidently, this model cannot represent accurately the interstellar icy grains, whose sizes are submicrometric and whose state is thought to be mainly amorphous (Boogert et al. 2015). To correct for this approximation, Wakelam et al. (2017) propose applying a coefficient of proportionality derived by comparing TPD experiments of different species against the computed BE for a single water molecule cluster. By applying this coefficient to their computed BE, Wakelam et al. (2017) obtained the aforementioned water BE of 38 kJ mol^{-1} (4600 K), relatively close to the values obtained from experiments on amorphous ice (see Section 2.1).

Das et al. (2018) adopted a similar approach (MP2/aug-cc-pVDZ), but decided to study the impact on the BE of the cluster size by using water models ranging from one to six molecules. They showed that increasing the ice model size can give results closer to the experimental ones, without the help of a correction coefficient. That said, while providing interesting methods for simple BE computations, the two above works do not take into account the importance of the hydrogen bond (H-bond) cooperativity present in large water cluster adsorption, resulting in sites exhibiting strong BE (as was pointed out before by Ferrero et al. 2020).

Ferrero et al. (2020) used two different periodic (crystalline and amorphous) slabs to model the icy grains. The BE value for the crystalline water model is almost 10 kJ mol^{-1} higher than the maximum value found using the amorphous ice model. This difference can be explained by the different length of H-bond chains in the crystalline ice (which are in principle infinite) with respect to the amorphous one, in which they break due to the loss of long range order. The BEs with the amorphous model were between 30 and 51 kJ mol^{-1} (3600 and 6100 K), closer to both the experimental ones (40–47 kJ mol^{-1}) and the previous computational results.

Duflot et al. (2021) studied the adsorption of water using an ONIOM hybrid method similar to the one presented in this work. As in Ferrero et al. (2020), they adopted both crystalline and amorphous ice models. Both types of models are composed of a fixed-geometry low-quantum-mechanic (QM) part and an unfixed high-QM part. The two crystalline models have approximately 160 water molecules in the low-QM zone (ω B97X-D/6-31 + G**:PM6) and 20 water molecules in the high-QM zone (CBS/DLPNO-CCSD(T):PM6). The amorphous slab is generated by classical molecular dynamics (MD) simulations. Several adsorptions are performed on it and, for each adsorption, approximately 20 water molecules around the adsorbate are included in the high-QM zone (CBS/DLPNO-CCSD(T):PM6) and 120–150 molecules around it are treated as a low-QM zone (ω B97X-D/6-31 + G**:PM6). For

Table 2
Values of the Water BE as Computed in Different Theoretical Studies Using Various Methods and Ice Grain Models

BE	Ice Model	Method	Reference
38.2 (4600)	Water monomer	M06-2X/A-VTZ (a), (c)	1
22.2 (2670)	Water tetramer	MP2/A-VDZ (b), (c)	2
34.6 (4166)	Water hexamer	MP2/A-VDZ (b), (c)	2
59.9 (7200)	Crystalline periodic	B3LYP-D3/A-VTZ (a)	3
Min: 30.0 (3605)	Amorphous periodic	B3LYP-D3/A-VTZ (a)	3
Max: 50.8 (6111)			
μ: 41.1 (4941)			
Min: 33.1 (3980)	Crystalline cluster	ONIOM(DLPNO-CCSD(T)/CBS//ωB97X-D/6-31 + G**:PM6)	4
Max: 51.7 (6220)			
μ: 40.7 (4897)			
Min: 34.6 (4166)	Amorphous ice slab	ONIOM(DLPNO-CCSD(T)/CBS//ωB97X-D/6-31 + G**:PM6)	4
Max: 54.4 (6545)			
μ: 45.1 (5419)			
Binding mode 1:	Amorphous clusters	ω—PBE/def2-TZVP//HF-3c/MINIX)	5
μ: 22.7 (2725)			
σ: 3.7 (449)			
Binding mode 2:			
μ: 34.0 (4087)			
σ: 2.8 (338)			

Notes. The units are kilojoules per mole and kelvin for the values in parentheses. The total BE average, excluding the first three values (not accurate), is 50.8 ± 10.7 kJ mol^{-1} (6110 ± 1287 K). (a) "A-VTZ" refers to aug-cc-pVTZ. (b) "A-VDZ" refers to aug-cc-pVDZ. (c) This refers to cases in which BSSE and ZPE corrections were not applied. μ is the distribution mean and σ the standard deviation; Max is the maximum value and Min the minimum value.
References. (1) Wakelam et al. (2017); (2) Das et al. (2018); (3) Ferrero et al. (2020); (4) Duflot et al. (2021); and (5) Bovolenta et al. (2022).

both models, the obtained results are in good agreement with the experimental BEs listed in Table 1.

Bovolenta et al. (2022) adopted 20 amorphous clusters, of 22 water molecules each, coming from high-temperature MD followed by annealing at 10 K. For each cluster, the adsorbate randomly samples the surface until at least 225 BE structures are computed through various selection processes. Geometries are first subjected to a cost-effective optimization (HF-3c/MINIX), then a single-point computation (ω–PBE/def2-TZVP with zero-point energy, or ZPE, and Basis Set Superposition Error, or BSSE, correction) is performed to obtain more accurate BEs. Then, the BEs are separated into two modes, depending on whether the water molecules make one or two H-bonds with the surface (respectively, "Binding mode 1" and "Binding mode 2" in Table 2). Finally, a BE distribution is obtained and fitted using a Gaussian function, whose mean and standard deviation values are reported in Table 2.

3. Methodology

In this work, we have adopted an icy model grain composed of 200 water molecules. It was built using the ACO-FROST method (Germain et al. 2022b), which produces high-density amorphous solid water (hd-ASW). The cluster size is too large to run a full DFT geometry optimization and harmonic frequency evaluation, also considering that up to 100 different water adsorption sites are used to compute the final BE distribution. In order to tackle this problem, then, we took into account the main interactions being the H-bond and dispersion London interactions and both interactions having relatively short ranges. On the other hand, the H-bond acceptor/donor character is influenced by the H-bond cooperativity, which is of long range nature. Therefore, the H-bond short-range features can be captured in a local zone around the adsorption site by using a high-level QM method (in our case, DLPNO-CCSD(T); Guo et al. 2018), while the H-bond cooperativity can be taken into account by adopting the tight-binding semi-empirical GFN2 method (Bannwarth et al. 2019), proved to give excellent interaction energy for noncovalent systems, while treating the electrons explicitly, at variance with molecular mechanics methods. The combination between the two levels of calculus is handled through the ONIOM scheme (Chung et al. 2015), which defines the total energy through a combination between the energy contributions coming from the high-QM method for the local region (*model zone*) and that from the low-level method adopted for the whole cluster (*real system*). Full details of the methods are given in the following sections.

The BE (positive for a bounded system) is defined as the opposite of the interaction energy (ΔE), which is the difference between the energies of the complex (between the grain and the adsorbate, i.e., E_c) and the sum of the energies of the isolated adsorbate ($E_{\rm ads}^{\rm iso}$) and the isolated grain ($E_{\rm grn}^{\rm iso}$), as follows:

$$\mathrm{BE} = -\Delta E = E_{\rm ads}^{\rm iso} + E_{\rm grn}^{\rm iso} - E_c. \tag{2}$$

The BE can be decomposed in two terms: (i) the pure electronic interaction (BE_e), corrected, if needed, for the BSSE; and (ii) the geometry deformation energy ($\delta E_{\rm def}$) contribution. When each electronic energy is corrected by the ZPE term derived from the harmonic frequency calculation, the BE can be corrected for these contributions giving the binding enthalpy at 0 K (BH(0)), as

$$\mathrm{BH}(0) = \underbrace{\mathrm{BE}_e - \delta E_{\rm def}}_{\mathrm{BE}} - \Delta\mathrm{ZPE}, \tag{3}$$

where ΔZPE is the difference between each ZPE term. The mathematical details are reported in Appendix A.

3.1. Strategy to Compute Binding Energy Distribution

The initial BE computation is performed in the same way as in our previous work (Germain et al. 2022a). A high-density amorphous ice cluster of 200 water molecules is used as model for the interstellar icy grains[5] (Germain et al. 2022c). Using the ACO-FROST code (Germain et al. 2022b), we generate a spherical grid of 162 points evenly spaced around the cluster. Each of these points is replaced by three water molecules with a different random orientation. After projecting every water molecule between 2.5 Å and 3.0 Å from the grain, we obtain a total of 486 starting adsorption positions.

The strategy to compute the water BE distribution consists of five consequent steps:

(i) We fix the geometry of the grain model and run a geometry optimization at the semi-empirical QM (SQM) level of the adsorbed water molecule only for each starting position.
(ii) Starting from point (i), all molecules within a 5 Å region around the relaxed adsorbed molecule are allowed to relax at the SQM level.
(iii) On the resulting optimized (ii) structure, a further check is carried out to ensure the 5 Å region has not changed its number of water molecules due to the geometry relaxation. In case this has happened, the (ii)–(iii) cycle is repeated until no compositional changes in the 5 Å region are detected.
(iv) On the final (iii) structure, the two-layer ONIOM(QM:SQM) model chemistry is carried out. The model zone (corresponding to the 5 Å region) is treated with a high-level QM method, while both the model zone and the whole cluster (the real system) are treated at the SQM level. The resulting energies are then combined in a subtractive way as defined by the ONIOM method.
(v) During the ONIOM geometry optimization, all atoms outside the model zone are kept fixed; mechanical embedding and polarization of the model zone are taken into account by activating the corresponding keywords in the computer program. In the frequency calculations (within the harmonic approximation), only the normal modes related to the nuclei inside the model zone are taken into account, keeping fixed all the other nuclei. As for steps (ii) and (iii), we ensure that the model zone remains with the same number of water molecules during the geometry optimization. The isolated grain surface ($E_{\rm grn}^{\rm iso}$) is computed following our previous approach (called "*TPD*"), which resembles closely the physics of the experimental TPD process (Tinacci et al. 2022), relaxing the QM system after withdrawing the adsorbed water molecule.
(vi) The identical BE sites are removed from the final distribution in order to obtain only unique BE sites. The adopted procedure is reported in Appendix E.

In order to deal with the large number of samples, we modified the previously adopted procedure (Tinacci et al. 2022) to save computer resources, while keeping the accuracy of the computed BE as high as possible.

[5] See the 200 water grain model at https://aurelegermain.github.io/JSmol_grain/.

3.2. Computational Methods

After the preliminary geometric optimization via the GFN2 (Bannwarth et al. 2019) SQM method with the xTB program (Grimme et al. 2017), discussed above, we refine each BE sample via the multilevel (electrostatic embedded) ONIOM (Mayhall et al. 2010) QM:QM2 approach implemented in ORCA (v.5.0.2; Neese 2018). The GFN2 method is used in all the ONIOM calculations as the SQM low level for its high accuracy (Germain & Ugliengo 2020; Germain et al. 2021; Tinacci et al. 2022). The B97-3c (Brandenburg et al. 2018) method is used for the high-level ONIOM model zone for the geometries and the related vibrational harmonic frequencies. A benchmark, reported in Appendix F, is carried out to validate the B97-3c method. Default settings for the SCF and the geometry optimizer are adopted. In the few cases in which imaginary frequencies are found, the optimization is restarted with a tighter DFT integral grid (Grid3 in ORCA), ensuring the achievement of a PES minimum. Subsequently, the ONIOM energies are refined using DLPNO-CCSD(T) (Guo et al. 2018) as a method to treat the model zone. This method is coupled with aug-cc-pVTZ (Kendall 1992) as the primary basis set, while aug-cc-pVTZ/C (Weigend et al. 2002) is used as the auxiliary basis set for the resolution of the identity (RI) approximation in electron repulsion integrals. The DLPNO calculations are carried out with a tight pair natural orbital (PNO) setup and the default settings for the SCF, in order to achieve higher accuracy. Due to the possible non-negligible error that stems from the BSSE in the post–Hartree–Fock method, the dimer energy at the DLPNO-CCSD(T) level is corrected with the counterpoise method (Boys & Bernardi 1970). We carried out a benchmark, reported in Appendix G, with an extended explanation of the accuracy of the DLPNO-CCSD(T) method.

The rendering of molecule images is obtained via the VMD software (Humphrey et al. 1996), while the graphics elaboration and plots are performed via the TikZ and PGFPlots LaTeX packages.

3.3. The Desorption Rate Prefactor

As pointed out by previous articles (Ferrero et al. 2022; Minissale et al. 2022; Tinacci et al. 2022), attention should be given to the way in which the prefactor in the desorption rate ($k_{\rm thdes}$ in Equation (1)) is computed. As in the aforementioned articles, we here adopt the Tait et al. (2005) formula:

$$\nu(T) = \frac{k_{\rm B}\,T}{h}\left(\frac{2\pi\,m\;\;k_{\rm B}\,T}{h}\right)A\frac{\sqrt{\pi}}{\sigma h^3}(8\pi^2 k_{\rm B}\,T)^{\frac{3}{2}}\sqrt{I_x\;\;I_y\;\;I_z}, \quad (4)$$

where $k_{\rm B}$ is the Boltzmann constant, m is the mass of the molecule, h is the Planck constant, A is the surface area per adsorbed molecules (usually assumed to be 10^{-19} m^2), I_i is the i-esimal adsorbate principal moment of inertia, and σ is the symmetry adsorbate rotation factor. For H_2O, the principal moments of inertia are 1.83, 1.21, 0.62 a.m.u. Å^2, σ is 2, and m is 18 a.m.u.

In general, the Tait et al. (2005) prefactor does not depend on the adsorption site, due to the immobile particle approximation that sets the vibrational partition function contribution ($q_{\rm vib}^{\rm TST}$) to unity. To take into account the latter, Equation (4) can be multiplied by $q_{\rm vib}^{\rm TST}(T)$, which includes the contribution of the harmonic vibrational partition function of the isolated adsorbate, ${}^{\ddagger}q_{\rm ads}^{\rm iso}$, and that of the isolated surface, ${}^{\ddagger}q_{\rm grn}^{\rm iso}$, divided by the

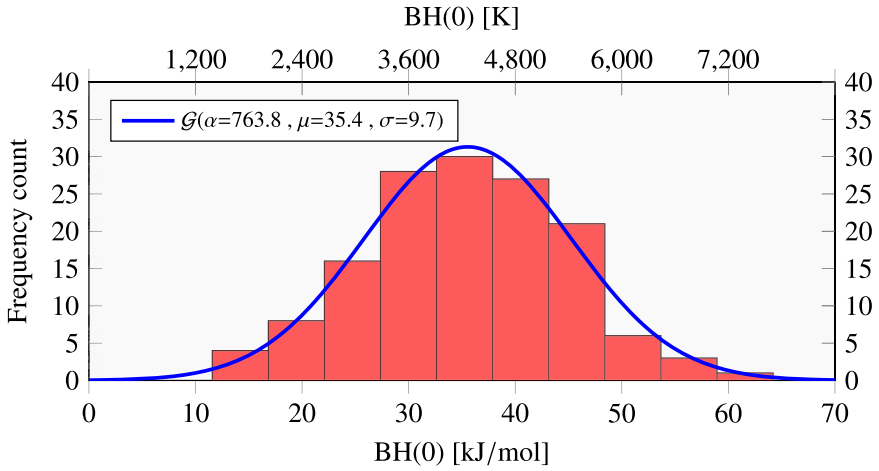

Figure 1. BSSE-corrected BH(0) distribution at the DLPNO-CCSD(T) level. Structures and ZPEs are calculated at the ONIOM(B97-3c:xTB-GFN2) level. The solid blue curve is the $\mathcal{G}(\texttt{hist}(\text{BH}(0)), \alpha, \mu, \sigma)$ unnormalized Gaussian best-fit function for the histogram (see Equation (6)).

harmonic vibrational partition function of complex, q_c, as follows:

$$q_{\text{vib}}^{\text{TST}}(T) = \frac{{}^{\ddagger}q_{\text{ads}}^{\text{iso}}\,{}^{\ddagger}q_{\text{grn}}^{\text{iso}}}{q_c} \approx \exp\left(-\frac{k_{\text{B}}T}{h\nu_f}\right). \quad (5)$$

The last term has been derived by fitting through an exponential function the whole $q_{\text{vib}}^{\text{TST}}(T)$, using ν_f as a free parameter.

4. Results

4.1. Binding Energy Distribution

The final BE distribution is obtained by 144 unique BE values. Note that all the obtained structures are PES minima, i.e., imaginary frequencies are not present. Due to the very low ($\leqslant$100 K) temperature in the molecular ISM, we only considered the ZPE-correct BE (vide supra) and no further thermal correction was included as it is less than about 3 kJ mol^{-1} (see Appendix D). In general, the differences between the two sets of computations are below (1–3 kJ mol^{-1}), which is the commonly accepted chemical accuracy (5 kJ mol^{-1}), and hence they are within the error of our computations.

The distribution of the final BH(0) values (Equation (3)), organized in a bin width following the Freedman Diaconis estimator (Freedman & Diaconis 1981), is shown in Figure 1. With respect to the ammonia distribution, which presents an asymmetric BE distribution (Tinacci et al. 2022), H_2O has a symmetric distribution with a data dispersion ranging from 11.6 to 64.2 kJ mol^{-1}. The distribution is well reproduced by an unnormalized Gaussian function:

$$\mathcal{G}(x, \alpha, \mu, \sigma) = \frac{\alpha}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (6)$$

with the following fitted parameters: mean (μ) of 35.4 kJ mol^{-1}, standard deviation (σ) of 9.73 kJ mol^{-1}, and prefactor (α) of 763.76 kJ mol^{-1}.

4.2. Energetic and Structural Information

As discussed in Section 3, the BH(0) can be decomposed in different contributions, which help to understand the form of the BE distribution. Figure B1, reported in Appendix B, shows the results of the BE decomposition in ΔZPE, BE$_e$, and δE_{def}, which can be summarized as follows:

(i) The ΔZPE contribution is rather sparse, as only a rough correlation exists with the ΔZPE, ranging from 5 to 15 kJ mol^{-1}.
(ii) The BE$_e$ correlates significantly with the BH(0). As expected, the values are much higher than the BH(0), as BE$_e$ does not include the deformation energy.
(iii) The δE_{def} does not correlate with the BH(0), showing a rather large spread of values from 0 to 50 kJ mol^{-1}.

More insight into the natures of the chemical interactions, specifically the network of H-bonds involving the adsorbed water molecule with the surface molecules, can be extracted from the structural features of the adsorbed complex, as shown in Figure 2.

We categorized each case in terms of the number of H-bonds in which the adsorbed water is involved. Figure 2 (left) shows that in the vast majority of cases (90 adsorption sites), the adsorbed water participates in two H-bonds, spanning a BE$_e$ average range of 58 kJ mol^{-1}. When water is involved in one H-bond, only 12 adsorption sites are involved and the BE$_e$ values are on the lower side of the BE$_e$ distribution (BE$_e$ < 60 kJ mol^{-1} and an average range of 37 kJ mol^{-1}). The case of three H-bonds shows a BE$_e$ range peaked at about 69 kJ mol^{-1}, with the lowest BE$_e$ value at 42 kJ mol^{-1}. These data are in agreement with the obvious expectation that the higher the number of H-bonds, the higher the average BE$_e$ values of the corresponding distribution. Another common concept about the H-bond geometrical features is the anticorrelation between the H $\cdots$ O bond length and the O–H$\cdots$O angle. Figure 2 (right) shows, indeed, the expected trend, particularly for the cases involving two H-bonds. In all cases, the O–H$\cdots$O angle ranges from 120° to 180°.

4.3. Binding Energy Distribution of Singly and Doubly Deuterated Water

The degree of deuteration of water, namely the HDO/H_2O abundance ratio, is very important in understanding the origin of terrestrial water and, more generally, to elucidate where and when water molecules are formed in astrophysical objects (see the introduction). For this reason, we studied the BE

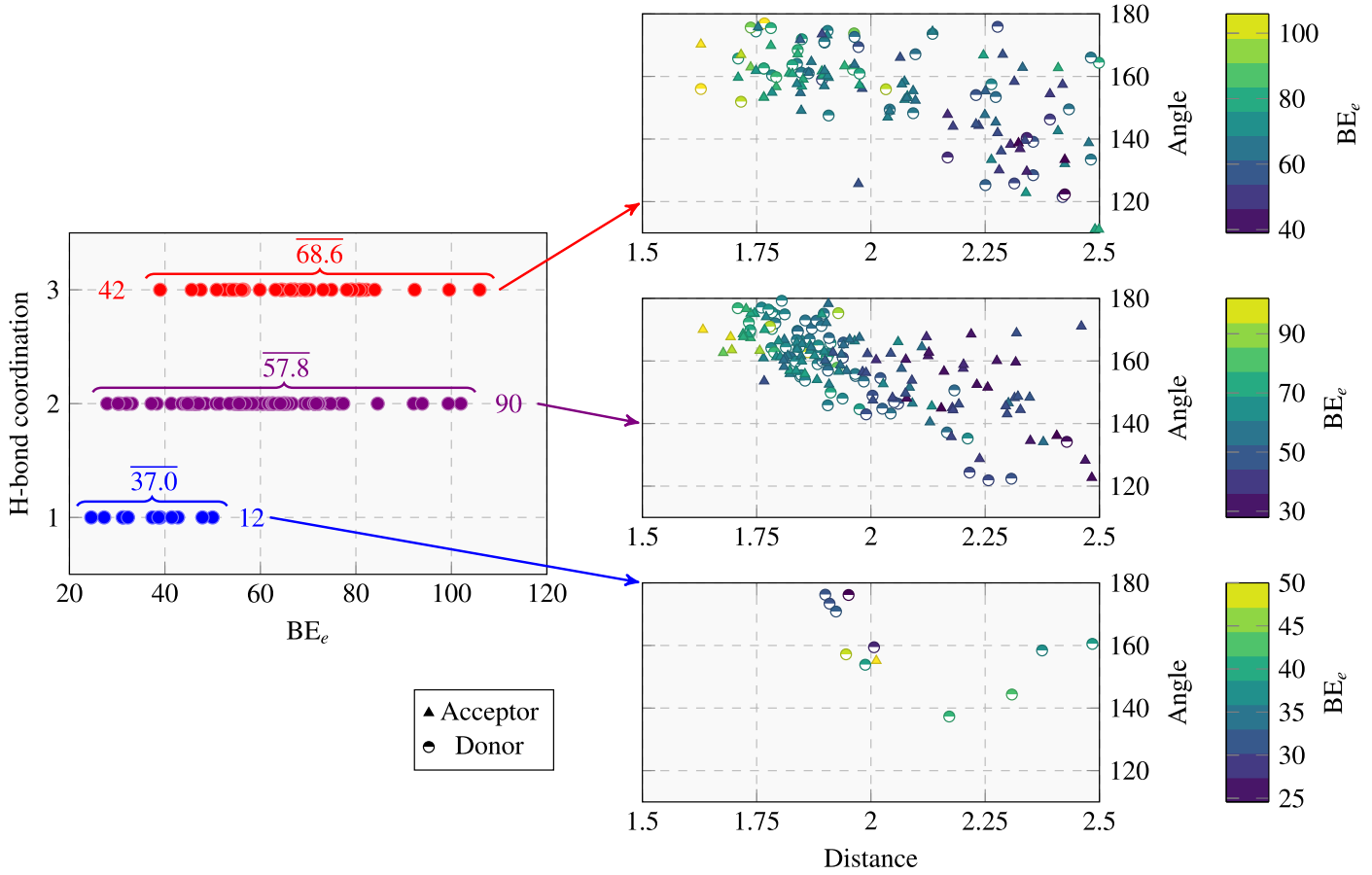

Figure 2. Correlation plots related to the H-bond structural features with respect to the pure electronic BE interaction (BE$_e$). All quantities are in units of angstrom, kilojoules per mole, and degrees. Left: BE$_e$ vs. number of H-bonds; the overbar numbers are the average of the BE$_e$ distribution, while the plain numbers refer to the number of corresponding cases with the specific number of H-bonds. Right: correlation plots of the OH···O distance vs. the O–H···O angle. The different symbols refer to water acting as an H-bond donor/acceptor, respectively.

distributions of singly (HDO) and doubly (D_2O) deuterated water, which are only affected by the difference in the ΔZPE with respect to the undeuterated case.

The correlation plots between the ΔZPE for the nondeuterated water and the singly/doubly deuterated ones are reported in Figure H1 in Appendix H. For the case of HDO, the arithmetic average of the two symmetric deuterated cases is reported. The correlation plot shows a lowering of the ΔZPE by a factor of 0.908 or 0.826 for one or two D substitutions, shifting BH(0) by 1 or 2 kJ mol^{-1}, respectively, a very modest decrease compared to the H_2O case.

4.4. The Desorption Rate Prefactor

Figure 3 shows the desorption rate prefactor (DSP), as a function of the temperature, computed with the Tait et al. (2005) formula (Equation (4)) and the one corrected for the harmonic partition functions, named CDSP (Equation (5); see below), respectively. CDSP varies between $\sim 7 \times 10^{13}$ and $\sim 3 \times 10^{15}$ s^{-1} within the T range 50–250 K, i.e., about a factor of 40. By assuming a desorption temperature of 120 K, as is often the case in astrochemical models, the CDSP is $\sim 8 \times 10^{14}$ s^{-1}, while the DSP is about 3×10^{15} s^{-1}, about a factor of 4 smaller than the CDSP. When considering the Hasegawa et al. (1992) rough estimate, our CDSP is about two orders of magnitude larger at 120 K.

Figure 4 shows the average (over the 144 adsorption cases) and the 95% confidence level of the correction term from the harmonic partition function as a function of temperature. As already pointed out (vide supra), at 120 K the correction factor is around 4, but it can be up to 10 at higher temperatures.

The correction factor (Equation (5)) as a function of the temperature is shown in Figure 4. The figure shows the average of the correction factor obtained for all the 144 computed BE sites and the 95% confidence level band. In general, the correction due to the inclusion of the harmonic vibrational partition functions can be larger than one order of magnitude, and it is larger at larger temperatures, as expected. Finally, by fitting against the average values using the exponential expression of Equation (5), we derive an effective frequency of about 82 cm^{-1}.

5. Comparison with Literature Data

5.1. Experimental Data

As discussed in Section 2.1, the experimentally measured BE on amorphous water ice lie between 40 and 47 kJ mol^{-1} (or ∼4800 to 5640 K), with a prefactor between $\sim 10^{12}$ and $\sim 10^{15}$ s^{-1}, if one excludes the outsider value of 64 kJ mol^{-1} (7720 K) obtained by Rosu-Finsen et al. (2022). The BE values are clearly larger than the BE average value of Figure 1, i.e., 35.4 kJ mol^{-1}.

The comparison between the computed and experimentally derived BE is not at all straightforward. Our approach simulates (as well as the other theoretical estimates of Table 2) the heat

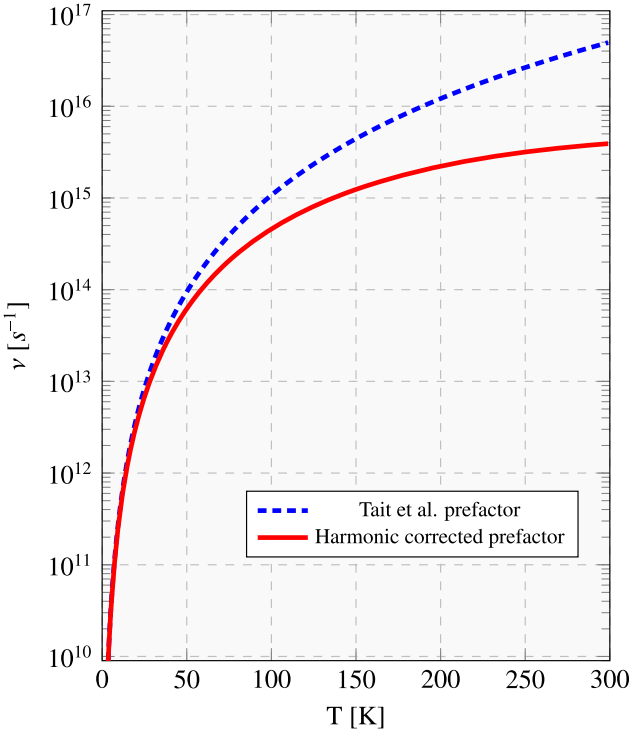

Figure 3. Temperature dependence of the desorption prefactor computed by using the Tait et al. (2005) formula (Equation (4); dashed blue line) and with the inclusion of the correction of the harmonic vibrational partition functions (CSDP; Equation (5); solid red line).

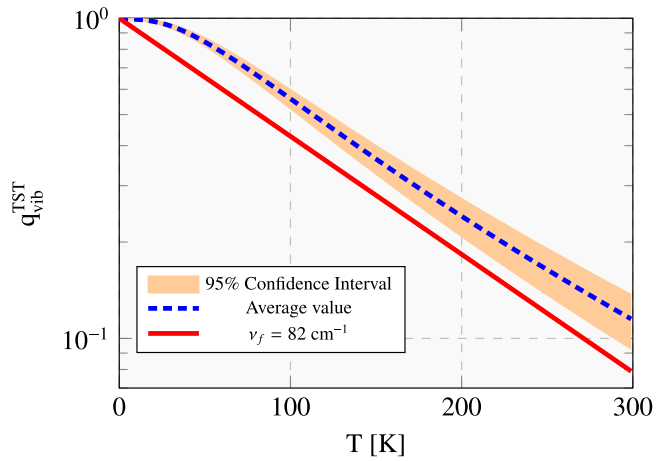

Figure 4. Correction of the desorption prefactor (Equation (5)) as a function of temperature. The dashed blue line shows the q_{vib}^{TST} average value of the water adsorption samples, with its confidence interval pictured in orange. The solid red line reports the fitting function of the q_{vib}^{TST} average value obtained with Equation (5).

released, $\Delta H(0)$, when a single water molecule is adsorbed on an icy model grain. The resulting BE distribution maps the different amounts of heat released by adsorption at different grain sites. These $\Delta H(0)$ data would therefore be directly comparable to the heat measured in a microcalorimetric experiment, as, for instance, when water is adsorbed at an amorphous silica surface or in an acidic zeolite (Bolis et al. 2006). In the astrochemical context, TPD is the method of choice, for many technical and experimental conditions not easily ameneable in a microcalorimeter. In TPD, however, as previously described (vide supra, Section 2.1), the BE is worked out numerically from the Polanyi–Wigner equation, together with the prefactor. A one-by-one comparison with the computed BE is, therefore, only meaningful when the theoretically computed prefactor is also of the same order of magnitude as that derived from the fitted experimental data. Furthermore, in the simulation, the icy grain structure is almost unaltered during the adsorption, but for a few water molecules surrounding the adsorbate. This is not the case during the temperature ramp adopted in TPD, where the heat transferred to the bulk grain can significantly restructure the surface site distribution, altering the BE values along the experiment.

For the specific case of water BE, another subtle difference biases the comparison with the computed data. Indeed, the computed BE are relative to a single specific water molecule that is used to probe different adsorption sites. In TPD, it is almost impossible either to select a specific water molecule to be desorbed or to ensure that the most exterior water molecules are desorbed with respect to more hidden ones. In the latter case, due to the many H-bonds involved, the BE values would be closer to the latent water of liquefaction (44 kJ mol^{-1}). An insight that can be the case comes from the work by Collings et al. (2015), in which they measured a BE of 35 kJ mol^{-1} and a prefactor of 1×10^{12} for water adsorbed on an amorphous silicate, at 0.4 ML of water coverage. And, last but not the least, the heat transferred during the TPD run can induce water surface diffusion to new sites coming from the restructured ice, biasing the BE distribution toward higher values (Minissale et al. 2022). In summary, the soft nature of the icy grains introduces criticalities in the TPD experiments and does not allow for a straightforward comparison between the computed and experimentally derived BE values. A better description and a link between the two approaches could be achieved via a microkinetic model that takes into account the BE distribution and the diffusion across the sites (He & Vidali 2014).

Finally, the observed difference between the measured and computed BEs raises the question of which of the two techniques is more reliable in providing the BE in astrochemical models. Also under this aspect, the use of TPD-derived BEs is not straightforward. Indeed, the timescale of desorption is dramatically different in TPD experiments (hours or days) and in the ISM (thousands or millions years). In the latter, the desorption is hugely slower than in the former, likely allowing just a minor adjustment of the ice surface. In this respect, therefore, the theoretically computed BEs have a better appeal in being chosen as the data inputs in the astrochemical models.

5.2. Theoretical Data

The comparison of the results obtained in this work and the other computational studies is more straightforward than with the experiments. In theoretical studies, the differences may arise by the different adopted ice models, the level of theory, and (if present) the ZPE correction.

Wakelam et al. (2017) used a single water molecule for the ice model. Their computed BE are systematically smaller than those measured in experiments and the larger the BE, the larger the difference (see their Figure 1). These authors decided to apply a proportionality factor to bring the two values closer, which, as explained above, may not be correct. Their water BE, based only on the theoretical computations, is 38 kJ mol^{-1}, indicating that the use of a single molecule for the ice model

leads to the sampling of a binding site whose BE is larger than the mean value of our Gaussian BE distribution.

Das et al. (2018) increased the number of water molecules (from one to six) simulating the ice, but only reported the H_2O BE without any ZPE and BSSE correction. The latter is of fundamental importance for the accuracy of post–Hartree–Fock theories, and even more for the relatively small adopted basis set. For this reason, and against the arguments provided by Das et al. (2018), we think that a comparison with their computations would not be useful.

Ferrero et al. (2020) found water BEs ranging from 30 to 50 kJ mol^{-1} on a periodic amorphous ice model, namely shifted toward the larger BEs with respect to our Figure 1 Gaussian distribution. The difference with the present results can be tracked down to the different icy surface model, the level of theory (they used a low-accuracy HF-3c level for geometry optimization), and, likely, the limited adsorption site sampling procedure.

In the article by Duflot et al. (2021), a procedure similar to the present one (ONIOM(CBS/DLPNO-CCSD(T):PM6)//ONIOM(ωB97X-D/6-31+G**:PM6)) was adopted to compute a ZPE-corrected BE. BE values of $40.1 \pm \sim 10$ kJ mol^{-1} and $40.7 \pm \sim 10$ kJ mol^{-1} have been computed on crystalline clusters and amorphous ice slabs, respectively. These data are in agreement with the present values with a difference of about 5 kJ mol^{-1}, despite the fact that a very different methodology was adopted to build up the ice underneath and the sampling (eight sites for each case) of all the BE sites.

Bovolenta et al. (2022) presented, for the first time, a BE distribution for water, following a new methodology described in Bovolenta et al. (2020). They decomposed the obtained distribution in two Gaussian functions, depending on the number of H-bonds between the adsorbate molecule and the surface. The two Gaussians have means of 22.7 (one H-bond) and 34.0 kJ mol^{-1} (two H-bonds), referring, respectively, to 74.7% and 25.3% of their samples. These results are at odds with what we obtained in our work, where only one Gaussian is present and with a peak at 35.4 kJ mol^{-1}, namely close to the second Gaussian of Bovolenta et al. (2022). Overall, their distribution is shifted toward the lower BE with respect to ours by about $\sim$10 kJ mol^{-1} (equivalent to $\sim$1200 K). In our opinion, this difference can be mostly attributed to (i) a possible huge and nonphysical deformation energy of the cluster, since no constraints are used in the geometry optimization, and (ii) one H-bond BE oversampling, because of the limited size of the cluster. Indeed, the authors themselves notice that, despite their distribution having the largest frequency counts (74.7%) associated with the the lower-energy Gaussian, "*in realistic water surface, evaporation would mostly fall within the higher BE regime*."

Regarding the estimate of the desorption prefactor, it is almost absent in the abovementioned theoretical studies, often assumed to be 10^{12}–10^{13} (the experimental assumption or the Hasegawa et al. 1992 formula). Recently, Molpeceres et al. (2020, 2022) proposed a new method, in which they run MD at the GFN2 level for the adsorption of atoms, H_2, and acetaldehyde on icy clusters. The prefactor, named by those authors the attempt frequency of desorption, was computed with an approach that had never been proposed in the related literature on surface science phenomena. Basically, the position of the adsorbate center of mass radial (COMr) with respect to the grain cluster center is monitored during a short time frame along the MD trajectories. The maximum of the COMr attempts are then worked out from the trajectories and the frequency of desorption is computed as the inverse of the average attempt periods of the simulations. However, if considering the vertical movements of the center of mass can be a smart solution for evaluating the attempt frequency of an adsorbed atom, this could not be the case for complex molecules, which have inner motions and a variety of different bonding interactions with the surface as a function of the change in orientation. Furthermore, a component in the COMr can also be simply due to different rotations of the whole molecule, either lying flat (small COMr) or almost perpendicular to the surface (high COMr), without any attempt to desorb. We notice that the approach of Molpeceres et al. (2020, 2022) has never been proposed by the surface science community, even though many methods have been proposed in the last years in order to estimate the prefactor (Fichthorn & Miron 2002; Sprowl et al. 2016; Rybicki & Sauer 2022). Furthermore, that procedure implies a rather long MD run, which can only be carried out with classical force fields or the GFNx level of theory, reducing the final accuracy.

6. The Water Snowline of Protoplanetary Disks

Several models of the protoplanetary disk chemical structure have been published in the literature, since the pioneer work by Aikawa & Herbst (1999). With the exception of the work by Grassi et al. (2020), all models assume a single value for the species BE, including that of water (e.g., Öberg & Wordsworth 2019; Ruaud & Gorti 2019; Wakelam et al. 2019; Cevallos Soto et al. 2022). Using an ad hoc Gaussian distribution of the water BE, peaked at 4800 K with a dispersion of 600 K, Grassi et al. (2020) approximately computed the midplane snowline. They showed that the latter is moved into smaller radii with respect to the use of the single-peak value. One can anticipate that, in addition to moving the snowline position, the existence of a BE distribution also implies that there are extended regions with gaseous water, although with a low abundance, and warm dusty regions where water remains in iced form, although in small fractions. These two effects have never been exploited before by any existing model, to the best of our knowledge.

In order to estimate the impact of the new computed H_2O BE distribution, we used a simple model to predict the snowline position, namely the gaseous and solid water abundances across a generic protoplanetary disk. The goal of the modeling is to show how the snowline of the disk changes using the newly computed BE and ν distributions with respect to the single values for H_2O (5600 K) and ν (2×10^{12} s^{-1}) adopted by the vast majority of the astrochemical models (e.g., Wakelam et al. 2017; Öberg & Wordsworth 2019, and the model mentioned above). Note that we did not attempt to compare model predictions obtained with our BE distribution with the one recently computed by Bovolenta et al. (2022), because the bias of their BE distribution toward the low BE values makes it unrealistic, as highlighted by the authors themselves (see Section 5.2).

In our model, we adopted a simplified chemistry for the water chemistry, described in Section 6.1, and ran it on a generic disk model, previously used in a similar study (Dominik et al. 2005), as described in Section 6.2. The results of the modeling are then shown and discussed in Section 6.3.

We emphasize that our goal is not to compute in detail the gaseous abundance, but rather the position of the snowline.

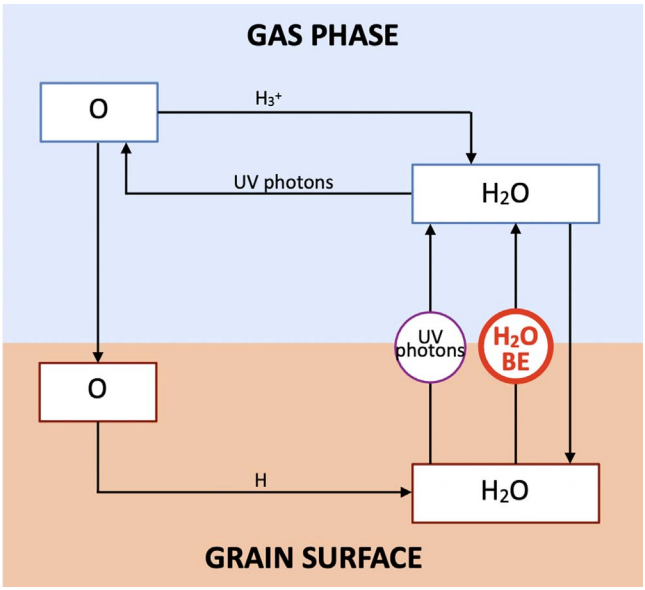

Figure 5. Simplified scheme of the gaseous water chemistry. Gaseous water can be formed by a chain of reactions starting from $O + H_3^+$, and by the photo- or thermal desorption of the water frozen on the grain surface. It is destroyed by the photodissociation or by the freezing onto the grain surface. When landing from the gas onto the grain surface, atomic oxygen is assumed to be instantaneously hydrogenated into frozen water.

6.1. Simplified Water Chemistry

We adopted the simplified scheme of the steady-state water chemistry shown in Figure 5, which accounts for the formation and destruction of gaseous and solid water in the interstellar cold and warm ($\leqslant$200 K) molecular gas. Briefly, the abundance of gaseous water is regulated by the photodesorption of the frozen water, its photodissociation into atomic oxygen, which in part reforms gaseous water,[6] the freezing of gaseous water onto the grain surfaces and its thermal desorption, following the scheme by Dominik et al. (2005). Since the thermal desorption takes the lion's share in this game, the water BE is a crucial parameter in defining the quantity of gaseous and solid water, namely the position of the snowline.

6.2. Model Description

We developed a simplified model for computing the gaseous and solid water abundance across a generic protoplanetary disk, updating and modifying the Dominik et al. (2005) model, i.e., adopting the scheme of Figure 5. Briefly, the model computes the gaseous and frozen abundance of water assuming that (i) when atomic oxygen freezes onto the grain surfaces, it instantaneously forms frozen water; (ii) at each disk position, the H_3^+ density (from which some gaseous water forms; see above) is equal to the electron density; and (iii) a fraction of the ice is sublimated into the gas phase, according to the BE distribution computed in this work.

[6] At steady state, gas-phase atomic oxygen reacts with H_3^+, forming OH^+; the formation of OH^+ is followed by the reactions $OH^+ + H_2 \rightarrow H_2O^+$ and $H_2O^+ + H_2 \rightarrow H_3O^+$ and, finally, the recombination $H_3O^+ + e^- \rightarrow H_2O$. In practice, about one-third of the gaseous O forms gaseous H_2O (Dominik et al. 2005; Hollenbach et al. 2009).

Table 3
Values of the Fraction of the Ice with a Given BE as a Function of BE as Well as the Computed Prefactor ν

BE	ν	Fraction of the Ice
14.2 (1705)	9.64×10^{12}	0.04
19.5 (2336)	2.53×10^{13}	0.08
24.7 (2968)	5.18×10^{13}	0.16
30.0 (3599)	9.10×10^{13}	0.28
35.2 (4230)	1.44×10^{14}	0.30
40.5 (4862)	2.12×10^{14}	0.27
45.8 (5493)	2.95×10^{14}	0.21
51.0 (6125)	3.92×10^{14}	0.06
56.3 (6756)	5.04×10^{14}	0.03
61.6 (7387)	6.29×10^{14}	0.01

Note. The units for BE are kilojoules per mole and kelvin for the values in parenthesis, and seconds for ν.

In practice, at steady state, the water gaseous abundance is computed analytically, solving the following three equations:

$$n_O\, k_{acc}^{O} = k_{phd}\, n_{H_2O} - k_{form}^{H_2O}\, n_O\, n_{H_3^+}$$
$$n_{H_2O}\,(k_{acc}^{H_2O} + k_{phd}) = k_{form}^{H_2O}\, n_O\, n_{H_3^+} + k_{phdes}\, n_{gr}\, k_{thdes}\, n_{ice}$$
$$n_{H_2O} + n_O + n_{ice} = A_{ox}\, f_{ice}\, n,$$

where n_O, n_{H_2O}, $n_{H_3^+}$, n_{gr}, n_{ice}, and n are the number density of gaseous atomic oxygen, water, and H_3^+, of the dust grains, frozen water (ice), and total H-nuclei density, respectively; k_{acc}^{O} and $k_{acc}^{H_2O}$ are the accretion rates onto the grain surfaces of O and H_2O, respectively; k_{phd}, k_{phdes}, and k_{thdes} are the gaseous water photodissociation rate and the frozen water photodesorption and thermal desorption rates; $k_{form}^{H_2O}$ is the formation of gaseous water from atomic oxygen (see footnote 6); A_{ox} is the total volatile oxygen elemental abundance (i.e., of the oxygen not trapped in the refractory dust grains); and f_{ice} is the fraction of sublimated ice. The latter is computed, at each point, by adding up the fraction of ice corresponding to the BE bin where $k_{thdes} \geqslant k_{acc}^{H_2O}$. Note that k_{thdes} is a function of BE, as reported in Equation (2).

The values of the adopted parameters are from Dominik et al. (2005), with the exception of the H_2O photodesorption yield, for which we assumed the value measured by Öberg et al. (2009), 2×10^{-3}. The total volatile oxygen elemental abundance A_{ox} is assumed to be equal to 1×10^{-4} with respect to H_2. The value of the fraction of the ice with a given BE as a function of BE is reported in Table 3, along with the computed prefactor ν.

Finally, we used the physical model (temperature and density) of DM Tau, published in Dominik et al. (2005).

6.3. Results of the Modeling and Implications

Figure 6 shows the results of the modeling using the Wakelam et al. (2017) single value for the BE (5600 K) and ν (2×10^{12} s^{-1}) as well as the BE and ν distributions computed in the present work, respectively.

When the Wakelam et al. (2017) values are used (upper panel in Figure 6), the water ice is all sublimated at about 150 K (light green region in the figure) and, therefore, the water abundance reaches 10^{-4} in the $\leqslant$3 au zone and, more

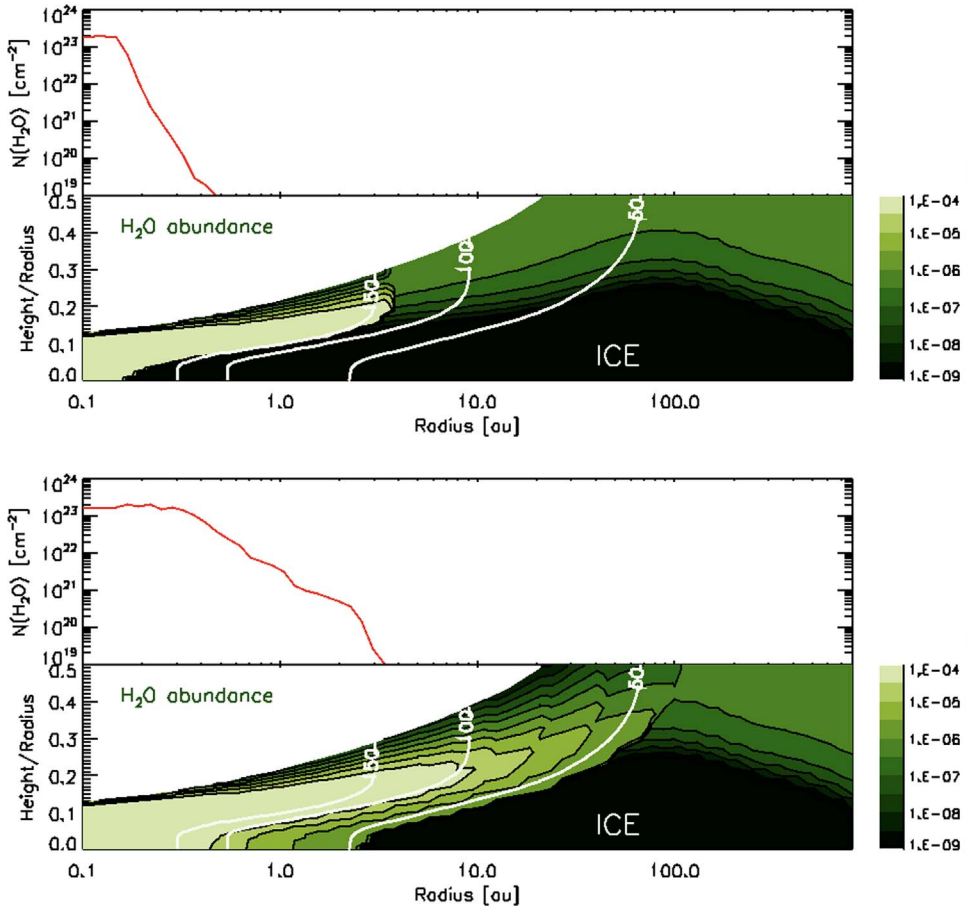

Figure 6. Astrochemical modeling results obtained assuming the water BE (5600 K) and ν (2×10^{12} s^{-1}) from Wakelam et al. (2017; top figure) and the BE and ν distributions computed in this work (bottom figure). Upper panels: gaseous water column density as a function of the radius. Lower panels: water abundance as a function of the radius and the height/radius. The white solid lines indicate where the dust temperature (assumed to be equal to that of the gas) is equal to 150, 100, and 50 K, respectively. In the black region, water is frozen onto the grain ice.

specifically, $\leqslant$0.2 au in the midplane. In the regions where the water ice does not sublimate, the gaseous water abundance is dominated by the equilibrium set by the photodesorption of the ices, where it is around 10^{-7}–10^{-6} (dark green region), and freezing onto the grain surfaces, where it drops to 10^{-8} (black region).

However, when the BE and ν distributions are used (lower panel in Figure 6), a larger region has a gaseous water abundance $\geqslant 3 \times 10^{-4}$ (light green region), where the dust temperature is $\geqslant$100 K. Most importantly, the region with a gaseous water abundance $\geqslant 1 \times 10^{-8}$ is much more extended, both on the disk plan and on the layers above it, whereas the iced water region is much reduced and starts at a radius of 2 au instead of 0.2 au in this model.

As a result, when the Wakelam et al. (2017) values are used, the gaseous H_2O column density $N(H_2O)$ is about 2×10^{23} cm^{-2} in the 0.1–0.16 au region, and then it drops by more than a factor of 100 at $\geqslant$0.3 au. When the BE and ν distributions are used, the $N(H_2O)$ peak is level up to 0.3 au, and then $N(H_2O)$ gently decreases and is a factor of 100 lower than the peak value at $\geqslant$1 au. This behavior is due to the BE distribution, in which about 80% of the BEs are less than the Wakelam et al. (2017) value.

It is worth remembering, at this point, that very few observations have been obtained of the cold water in protoplanetary disks (Hogerheijde et al. 2011; Podio et al. 2013), and all have been obtained with the far-IR satellite Herschel. Unfortunately, these observations, which targeted the ground water line at 538 μm, have a very large beam ($\sim$30″), hence they are incapable of constraining the spatial origin of the observed cold water vapor.

Interestingly, on the other side of the BE distribution, about 20% of the BEs are larger than the Wakelam et al. (2017) value. This implies that about 20% of the ice will need a larger dust temperature to sublimate than the one predicted by using the Wakelam et al. (2017) value ($\sim$150 K). Indeed, according to our modeling, the abundance of gaseous water is never larger than 9×10^{-5} when using the new BE distribution, namely $\sim$10% of water stays frozen in the ices. Even though it is a small fraction, this ice may remain trapped in the planetesimals, eventually forming rocky planets and asteroids inside the standard snowline, and, consequently, having an impact on their water content. A fortiori, this consideration applies to the terrestrial water. However, a focused and more sophisticated model, beyond the scope of this work, is needed to fully appreciate the real impact of the newly computed BE distribution on the quantity of water that could have been trapped in the planetesimals that formed our Earth.

7. Conclusions

In this work, we present an extended computational study of the BE of water molecules on the largest icy grain model so far published, formed by 200 water molecules. The grain model was recently built by Germain et al. (2022a) using their ACO-FROST code, which carries out ab initio computations.

We used a multistep procedure to obtain the BE over the whole set of adsorbing sites on the icy grain. First, we adsorbed the H_2O molecule and let the system relax over a 5 Å ice sphere surrounding the adsorbed water molecule via a GFN2 multiple optimization. We then refined the computations via the two-layer ONIOM(QM:SQM) methodology to obtain high-level-accuracy QM computations of the 5 Å ice sphere plus adsorbed H_2O, while the rest of the ice grain was computed via an SQM method. From the initial set of 486 adsorbing sites where we computed the H_2O BE, we extracted 144 unique sites, meaning each site had a different geometry.

The H_2O BE of the unique sites ranges from 14.2 kJ mol^{-1} (1705 K) to 61.6 kJ mol^{-1} (7390 K). The distribution of the computed BE fairly well follows a Gaussian distribution, whose peak is at 35.4 kJ mol^{-1} (4230 K) and standard deviation is 9.7 kJ mol^{-1} (1160 K). We also computed the prefactor, which enters in the desorption rate computation, as a function of BE. It ranges from $\sim 1 \times 10^{13}$ s^{-1} (at the lowest BE) to $\sim 6 \times 10^{14}$ s^{-1} (at the largest BE). These values are significantly different from those computed with the Hasegawa et al. (1992) method ($\sim 2 \times 10^{12}$ s^{-1}) and adopted by several astrochemical models.

The newly computed H_2O BE distribution compares fairly well with the BE derived from previous experimental TPD studies, when considering the methodological difference between the experimental and computational derivations. For example, the former measure the curve of desorption, from which BE and ν are derived, and cannot avoid the effects of the adsorbing structural deformation during the process. On the contrary, computations are unaffected by this problem and, given the large scales involved, they very likely better describe the situation in the ISM.

Finally, we developed a model to describe the water abundance across a generic protoplanetary disk and compute the position of the snowline. We found that the region of gaseous water was more extended (by about a factor of 10) when considering the newly computed BE and ν distributions than when computed with the single BE and ν values reported in Wakelam et al. (2017) and used in many astrochemical models. Mirroring the gaseous water distribution, the region where water is frozen is much larger (by the same factor). This behavior is due to the fact that about 80% of the BEs in the computed distribution are lower than the Wakelam et al. (2017) value (5600 K).

Furthermore, the 20% of BEs with a value larger than the Wakelam et al. (2017) one leads to the incomplete sublimation of the ices at temperatures where the single-BE modeling would predict the full release of frozen water into the gas phase. Specifically, about 10% of water remains frozen at $\sim$150 K and this might contribute to the enrichment of the water of rocky planets inside the classical snowline. Whether this has an impact on the origin of the terrestrial water would need a dedicated and sophisticated model.

8. Online Database

To easily handle the large data set of BE samples (atomic coordinates and BH(0) values), we have developed and made publicly available a website based on the molecule hyperactive JSmol plugin (Jmol: an open-source Java viewer for chemical structures in 3D).[7] This extended electronic version of the calculated results, including all the 144 sample structures at the ONIOM(B97-3c:xTB-GFN2) level, can be explored interactively on GitHub,[8] and is made available under a Creative Commons Attribution license on Zenodo: doi:10.5281/zenodo.7802771.

Acknowledgments

This project has received funding within the European Union's Horizon 2020 research and innovation program from the European Research Council (ERC) for the project "The Dawn of Organic Chemistry" (DOC), grant agreement No. 741002, and from the Marie Skłodowska-Curie Actions for the project "Astro-Chemical Origins" (ACO), grant agreement No. 811312. P.U. acknowledges the Italian Space Agency for co-funding the Life in Space Project (ASI N. 2019-3-U.O). CINES-OCCIGEN HPC is kindly acknowledged for the generous allowance of supercomputing time through the A0120811498 project. L.T. is grateful to Jacopo Lupi, Leonardo Miele, Simon Ferrada-Chamorro, and Stefano Ferrero for insightful discussions and to the LATEX community for the insights on the TikZ and PGFPlots packages. Finally, we wish to acknowledge the extremely useful discussions with Prof. F. Dulieu, Prof. Albert Rimola, Dr. P. Theule, and Prof. Gretobape and his group.

Software: ORCA (Neese 2018), VMD (Humphrey et al. 1996), xTB (Grimme et al. 2017), ACO-FROST (Germain et al. 2022b).

Appendix A
Adopted Formalism to Compute the Binding Energies

The equation adopted for the calculation of the ONIOM BEs is

$$\mathrm{BE} = -\Delta E = E_{\mathrm{ads}}^{\mathrm{iso}}(\mathrm{QM}) + E_{\mathrm{grn}}^{\mathrm{iso}}(\mathrm{QM{:}\ SQM}) - E_c(\mathrm{QM{:}\ SQM}). \quad \text{(A1)}$$

The BE can be decomposed into the pure electronic interaction (BE_e) corrected for BSSE and the deformation energy (δE_{def}) contributions. The BE_e is given by

$$\mathrm{BE}_e = E_{\mathrm{ads}}^{\mathrm{iso}//\mathrm{c}}(\mathcal{G}(\mathrm{grn})) + E_{\mathrm{grn}}^{\mathrm{iso}//\mathrm{c}}(\mathcal{G}(\mathrm{ads})) - E_c(\mathrm{QM}), \quad \text{(A2)}$$

where $E_{\mathrm{ads}}^{\mathrm{iso}//\mathrm{c}}(\mathcal{G}(\mathrm{grn}))$ and $E_{\mathrm{grn}}^{\mathrm{iso}//\mathrm{c}}(\mathcal{G}(\mathrm{ads}))$ are the energies of the isolated adsorbate and the grain in the geometries assumed in the complex (iso//c) in the presence of the ghost orbitals of the grain $\mathcal{G}$(grn) and the adsorbate $\mathcal{G}$(ads), respectively. As the BSSE is already taken into account in the GFN2 method, Equation (A2) only applies to the QM methods (vide infra) in

[7] http://www.jmol.org/
[8] https://tinaccil.github.io/Jmol_BE_H2O_visualization/

the model zone. The $\delta E_{\rm def}$ is defined as

$$\delta E_{\rm def} = \underbrace{(E_{\rm ads}^{\rm iso//c} - E_{\rm ads}^{\rm iso})}_{\delta E_{\rm def}^{\rm ads}} + \underbrace{(E_{\rm grn}^{\rm iso//c} - E_{\rm grn}^{\rm iso})}_{\delta E_{\rm def}^{\rm grn}}, \quad \text{(A3)}$$

where $\delta E_{\rm def}^{\rm ads}$ and $\delta E_{\rm def}^{\rm grn}$ are the deformation energy of the adsorbate and the surface, respectively. Vibrational frequencies were computed on the model zone to obtain the ZPEs (Equation (C2)), from which the ΔZPE resulted as

$$\Delta{\rm ZPE} = {\rm ZPE}_c - {\rm ZPE}_{\rm ads}^{\rm iso} - {\rm ZPE}_{\rm grn}^{\rm iso}. \quad \text{(A4)}$$

Including all the abovementioned contributions, Equation (A1) becomes

$${\rm BH}(0) = \underbrace{{\rm BE}_e - (\delta E_{\rm def}^{\rm grn} + \delta E_{\rm def}^{\rm ads})}_{\rm BE} - \Delta{\rm ZPE}. \quad \text{(A5)}$$

Appendix B
Decomposition Results

The results of the BH(0) decomposition, described above, are presented in Figure B1.

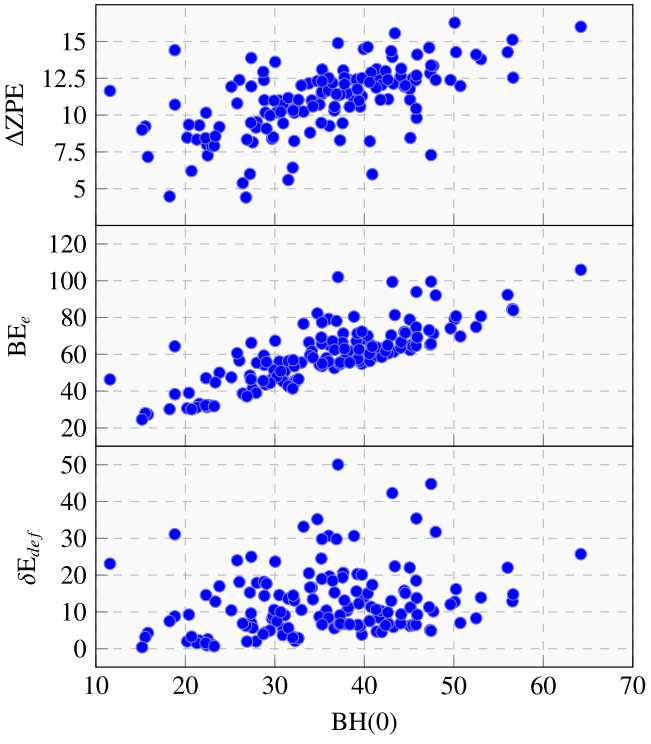

Figure B1. Correlation plots between BH(0) and ΔZPE, BE$_e$, and $\delta E_{\rm def}$. All the quantities are in units of kilojoules per mole. All BH(0) and BE$_e$ are BSSE-corrected. The BH(0) decomposition is illustrated briefly in Equation (3) and extensively in Appendix A.

Appendix C
Vibrational Partition Function and ZPE

The vibrational partition function $q_{\rm vib}$ in the harmonic oscillator assumption and for a nonlinear molecule is obtained as

$$q_{\rm vib} = \Pi_i^{3N-6} \frac{1}{1 - \exp\left(\frac{h\nu_i}{k_{\rm B}T}\right)}. \quad \text{(C1)}$$

In this case, we choose the first vibrational energy level to be the zero of energy.

The ZPE is obtained as

$${\rm ZPE} = \sum_i^{3N-6} \frac{h\nu_i}{2k_{\rm B}}. \quad \text{(C2)}$$

Appendix D
BE Thermal Correction

If thermal correction is taken into account, the BE at a given temperature T is defined as follows:

$${\rm BH}(T) = {\rm BH}(0) + 4RT + H_{\rm ads}^{\rm vib}(T) + H_{\rm grn}^{\rm vib}(T) - H_c^{\rm vib}(T), \quad \text{(D1)}$$

where the $4RT$ term (R is the ideal gas constant) comes from the classical rotational ($3/2RT$) and translational ($3/2RT$) contributions of the isolated adsorbate molecule (since the water is nonlinear) in the rigid rotor approximation and ($1RT$) from the volume of the work contribution to the enthalpy. The $H^{\rm vib}(T)$ is the vibrational thermal contribution to the enthalpy (without ZPE, since it is already taken into account in Equation (3)),

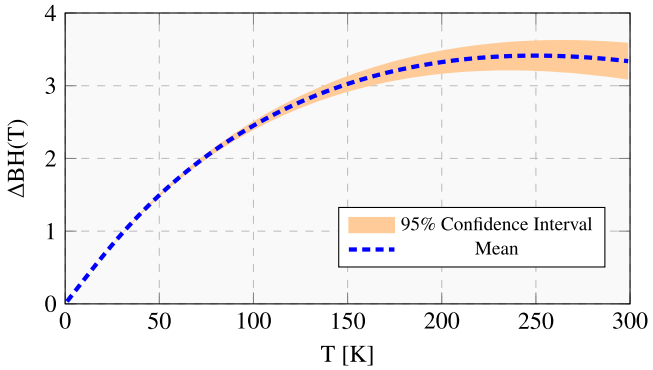

Figure D1. Thermal correction to the BE ($\Delta{\rm BH}(T) = {\rm BH}(T) - {\rm BH}(0)$) average value of the water adsorption samples and its confidence interval as a function of the temperature (Equation (D2)).

which in the case of the rigid rotor harmonic oscillator is

$$H^{\mathrm{vib}}(T) = R\sum_{i}^{3N-6}\frac{h\nu_i/k_{\mathrm{B}}}{\exp\left(\frac{h\nu_i}{k_{\mathrm{B}}\,T}\right)-1}, \tag{D2}$$

where h is the Plank constant, k_{B} is the Boltzmann constant, ν is the normal mode vibrational frequency, and N is the number of atoms of the system.

The thermal correction to the water BE distributions is reported in Figure D1.

Appendix E
Procedure to Obtain Unique Binding Energy Sites

To obtain unique BE sites, we developed a pruning procedure on the obtained ONIOM complex structures. This procedure is necessary because of the possibility of arriving at the same PES minima, starting from slightly different geometrical initial conditions. The redundancy cases are identified in the following way:

1. We calculate all the combinations between the BE's model zone structure that have the same number of atoms.
2. For each combination, we align the two structures and compute the RMSD and $|\Delta\mathrm{BE}|$.
3. If the two structures have an RMSD below 0.5 Å and $|\Delta\mathrm{BE}_e| < 0.75$ kJ mol^{-1}, they are considered identical.

The 0.5 Å and 0.75 kJ mol^{-1} threshold parameters were selected after an exhaustive inspection of the possible redundant BE sites.

Appendix F
B97-3c Benchmark

To validate the accuracy of the B97-3c function we adopted a twofold strategy: we compared the BE computed at the B97-3c level with respect to the less approximate B97D3 method (Grimme 2006; Grimme et al. 2011) coupled with the aug-cc-pVTZ basis, whose performance has already been extensively studied in our previous article (Tinacci et al. 2022). We limited the accuracy check to one H_2O/grain system as a test case, for which we compared structures and BH(0)s at the B97D3 level with different basis sets: (i) B97D3/aug-cc-pVTZ; (ii) B97D3/Def2-TZVP; and (iii) B97D3/Def2-TZVP with geometrical CounterPoise (Kruse & Grimme 2012) correction. All the DFT structures and their related ZPE corrections are obtained by starting from the ONIOM(B97-3c:xTB-GFN2) optimized structure. The usual single-point energy refinement at the DLPNO-CCSD(T)/aug-cc-pVTZ level is performed on the minimum obtained by each different setup. All BEs are corrected for their BSSE. The results of Figure F1 show that (i) the DLPNO-CCSD(T) BE is almost insensitive to the level of the optimized geometry, varying less than 2 kJ mol^{-1} across the considered cases; (ii) plain DFT results converge within 4 kJ mol^{-1} to the DLPNO-CCSD(T) results; and (iii) the B97-3c BE are off by a significant amount, while providing a robust and quick method for the geometry optimization.

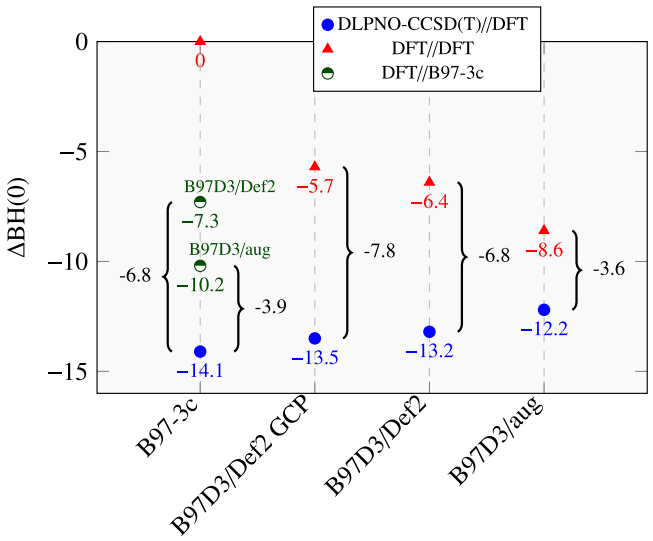

Figure F1. Difference ΔBH(0) between the BH(0) B97-3c reference value (67.8 kJ mol^{-1}) and the BH(0)s computed with the reported QM methods in the ONIOM(QM:xTB-GFN2) framework. All the BEs are BSSE- (when needed) and ZPE-corrected. By "Def2" and "aug," we mean the Def2-tZVP and aug-cc-pVTZ basis sets, respectively. All quantities are in kilojoules per mole.

Appendix G
Accuracy of the DLPNO-CCSD(T) Approximation

As pointed out in many articles in the last years, the DLPNO approximation to the CCSD(T) method could have significant additive errors with the size of the system (Pavošević et al. 2017; Al-Hamdani & Tkatchenko 2019; Sandler et al. 2021). The error in such a method, excluding the basis set incompleteness error and the algorithm to solve the triplets (T), resides primarily in the two parameters that control the size of the correlation space: T_{CutPairs} and T_{CutPNO}. T_{CutPairs} defines the pairs to be treated in the coupled cluster ("*Strong pairs*") or in the MP2 ("*Weak pairs*") fashion. T_{CutPNO} controls the dimension of the compact virtual space in the function of PNOs, for which the coupled cluster "*Strong pairs*" equations are solved.

In the article by Altun et al. (2020), they provided a solution by using an easy extrapolation procedure, similar but conceptually different to the two-point CBS strategy (Petersson et al. 1988; Feller 1993), to recover the Complete PNO Space (CPS) and correct the additive error on the electron pair truncation correlation energy. This extrapolation is applied to the correlation energies obtained with two T_{CutPNO} values. The equation to recover the correlation energy in the CPS extrapolation limit (E_{CPS}), for a given basis set, is

$$E_{\mathrm{CPS}} = E^X + 1.5\cdot(E^Y - E^X), \tag{G1}$$

where E^X and E^Y are the correlation energies with $T_{\mathrm{CutPNO}} = 10^X$ and 10^Y, respectively, and $Y = X + 1$. To calculate the total absolute energy, the E_{CPS} quantity must be added to the DLPNO-CCSD energy, which is indifferent to the PNO cutoff. Recently, the same authors also provided an extensive study, on many different chemical systems, on the impact of such a correction (Altun et al. 2021) and the dependence of the DLPNO-CCSD(T) error with respect to the RI approximation and BSSE. From what we can extract, for our purpose, from their conclusion: (i) the DLPNO-CCSD(T) error

Table 4
Accuracy of DLPNO-CCSD(T) Calculation Obtained from Structures and Related ZPE Corrections at the ONIOM(B97-3c:xTB-GFN2) Level

System	15 Water				7 Water							
Basis Set	aug and aug/C				aug and aug/C				aug and autoaux-max			
DLPNO-CCSD	(T_0)	(T_0)	(T_1)	(T_1)	(T_0)	(T_0)	(T_1)	(T_1)	(T_0)	(T_0)	(T_1)	(T_1)
$\log(T_{\mathrm{CutPNO}})$	−7	−6	−7	−6	−7	−6	−7	−6	−7	−6	−7	−6
BSSE	7.36	7.41	7.38	7.42	4.61	4.59	4.62	4.60	4.64	4.65	4.66	4.66
BH(0)	53.63	52.53	53.92	52.79	32.18	31.49	32.28	31.57	32.13	31.44	32.23	31.53
BH(0) CPS	54.15		54.46		32.54		32.65		32.47		32.58	
BH(0) B97-3c	67.81				36.37							
BH(0) CCSD(T)/aug	–				32.90 (BSSE 4.65)							

Note. The CCSD(T)/aug-cc-pVTZ was impossible to compute for the 15 water system for the choice of basis set and our computational resources.

does not depend on the BSSE error; (ii) with the TightPNO setting or with the CPS limit, the error linked to the basis set incompleteness is negligibly small; and (iii) the RI error for the absolute correlation energy for the aug-cc-pVTZ and aug-cc-pVTZ/C basis set is of the order of 0.5 kJ mol^{-1} for the water dimer. What is interestingly missing from their articles is the impact of such correction and error on our target quantity: the BE. In fact, due to its intrinsic subtractive nature, this quantity may have an error compensation. In order to investigate the error, we compute two water absorption cases, within the ONIOM framework, with 15 and six water molecules in the model zone each. In this benchmark, we study the accuracy and dependency between each of the following cases:

1. T_{CutPNO} at 10^{-6} and 10^{-7} with the related CPS(6/7) extrapolation.
2. The accurate iterative (T_1; Guo et al. 2018, 2020) algorithm and the approximate (T_0) correction (Riplinger et al. 2013).
3. The aug-cc-pVTZ/C and "*autoaux-max*" basis set (which is obtained by matching the maximum angular moment of the parent basis set; Stoychev et al. 2017).

The aug-cc-pVTZ primary basis set is always kept, as well as the TightPNO setting. Many works have reported that the approximated algorithm to compute the triplet excitation of the coupled cluster method—the default one in ORCA and here addressed as (T_0), which in the rest of the article is generally called (T)—can affect the accuracy of the calculations with respect to the nonapproximated one (i.e., (T_1); Guo et al. 2018, 2020). For this reason, we also test this issue. Altun et al. (2020) found the auxiliary "*autoaux-max*" basis set (Stoychev et al. 2017) coupled with the primary aug-cc-pVTZ ensured an RI that was error-free for the water dimer.

The benchmark is reported in Table 4. The BSSE has no relationship with the DLPNO-CCSD(T) correlation error, the treatment of the triplets, or the RI approximation, and it remains quasi-constant below ∼0.05 kJ mol^{-1} in all the configurations. The RI error in the 7 water system is negligible and of the order of ∼0.05 kJ mol^{-1}.

The BE seems to have a compensation error, so we decided to use the (T_0), aug-cc-pVTZ basis set and not CPS correction to achieve the maximum accuracy/cost ratio for our objectives. The CPS strategy should be useful in bigger systems than our own. In fact, in cases of larger systems, in which drastic decisions must be taken to save computational time, one can use BSSE evaluated with the (T_0), CPS(7/6) or CPS(6/5) schemes and a different combination of primary and secondary basis sets.

In summary: (i) the BE seems to have a compensation error; (ii) the BSSE, which has a non-negligible impact in the post–Hartree–Fock method, is not affected by the correlation error, the treatment of the triplets, or the RI approximation; (iii) the RI error is also negligible; and finally (iv) the CPS strategy is probably useful in larger systems than ours.

For these reasons, we have decided to keep our DLPNO-CCSD(T) setup, and do not apply the CPS strategy.

Appendix H
Deuteration Impact on Water Binding Energy

Figure H1 shows the correlation plots between ΔZPEs calculated for the water (ΔZPE(H2O): x-axis), and mono-deuterated (ΔZPE(HDO) (see main body for details).

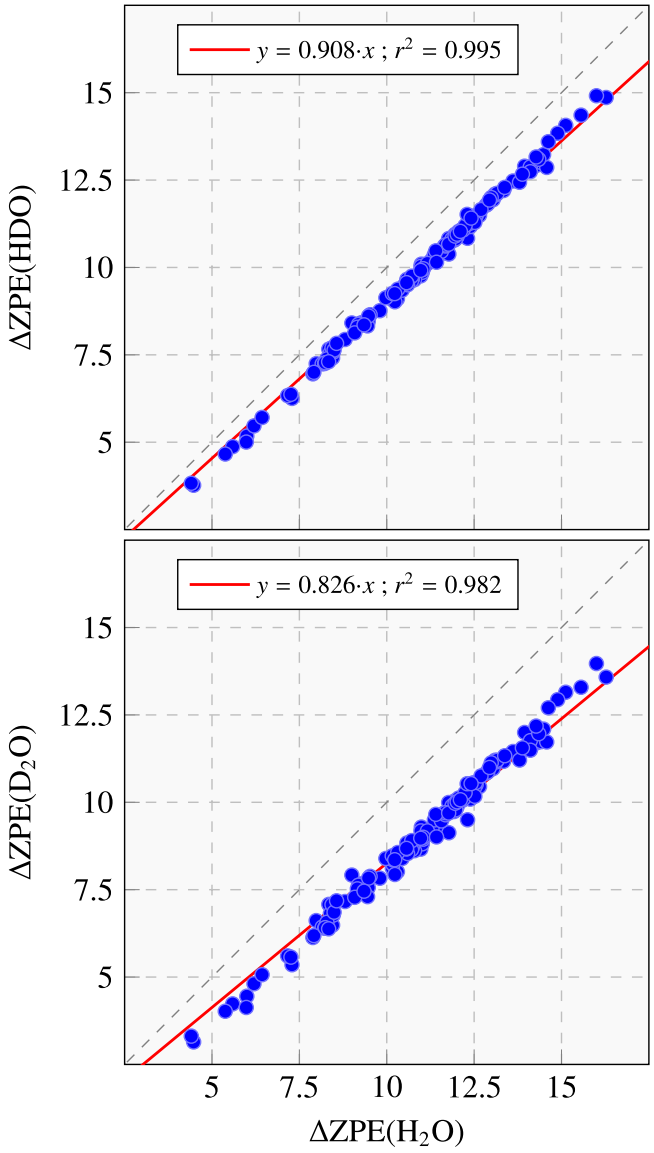

Figure H1. Correlation plots between ΔZPEs calculated for the water (ΔZPE (H_2O): x-axis), and monodeuterated (ΔZPE(HDO): upper panel) and bideuterated (ΔZPE(D_2O): bottom panel) water. For the monodeuterated case, the ΔZPE is the average between the two equivalent deuterated positions. All the quantities are in units of kilojoules per mole.

ORCID iDs

Lorenzo Tinacci https://orcid.org/0000-0001-9909-9570
Aurèle Germain https://orcid.org/0000-0001-7856-0516
Stefano Pantaleone https://orcid.org/0000-0002-2457-1065
Cecilia Ceccarelli https://orcid.org/0000-0001-9664-6292
Nadia Balucani https://orcid.org/0000-0001-5121-5683
Piero Ugliengo https://orcid.org/0000-0001-8886-9832

References

Aikawa, Y., & Herbst, E. 1999, A&A, 351, 233
Alexander, C. M. O. 2017, RSPTA, 375, 20150384
Al-Hamdani, Y. S., & Tkatchenko, A. 2019, JChPh, 150, 010901
Altun, A., Ghosh, S., Riplinger, C., Neese, F., & Bistoni, G. 2021, JPCA, 125, 9932
Altun, A., Neese, F., & Bistoni, G. 2020, JCTC, 16, 6142
Bannwarth, C., Ehlert, S., & Grimme, S. 2019, JCTC, 15, 1652
Bolina, A. S., Wolff, A. J., & Brown, W. A. 2005, JPCB, 109, 16836
Bolis, V., Busco, C., & Ugliengo, P. 2006, JPCB, 110, 14849
Boogert, A. C., Gerakines, P. A., & Whittet, D. C. 2015, ARA&A, 53, 541
Bovolenta, G., Bovino, S., Vöhringer-Martinez, E., et al. 2020, MolAs, 21, 100095
Bovolenta, G. M., Vogt-Geisse, S., Bovino, S., & Grassi, T. 2022, ApJS, 262, 17
Boys, S. F., & Bernardi, F. 1970, MolPh, 19, 553
Brandenburg, J. G., Bannwarth, C., Hansen, A., & Grimme, S. 2018, JChPh, 148, 064104
Brown, W. A., & Bolina, A. S. 2007, MNRAS, 374, 1006
Caselli, P., Keto, E., Pagani, L., et al. 2010, A&A, 521, L29
Ceccarelli, C., Caselli, P., Bockelée-Morvan, D., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 859
Ceccarelli, C., Caux, E., Loinard, L., et al. 1999, A&A, 342, L21
Ceccarelli, C., & Du, F. 2022, Eleme, 18, 155
Ceccarelli, C., Hollenbach, D. J., & Tielens, A. G. G. M. 1996, ApJ, 471, 400
Cernicharo, J., Gonzalez-Alfonso, E., Alcolea, J., Bachiller, R., & John, D. 1994, ApJL, 432, L59
Cevallos Soto, A., Tan, J. C., Hu, X., Hsu, C.-J., & Walsh, C. 2022, MNRAS, 517, 2285
Chakarov, D. V., Oesterlund, L., & Kasemo, B. 1995, Langm, 11, 1201
Chung, L. W., Sameera, W. M. C., Ramozzi, R., et al. 2015, ChRv, 115, 5678
Collings, M. P., Frankland, V., Lasne, J., et al. 2015, MNRAS, 449, 1826
Combes, F., & Wiklind, T. 1997, ApJL, 486, L79
Das, A., Sil, M., Gorai, P., Chakrabarti, S. K., & Loison, J. C. 2018, ApJS, 237, 9
Dominik, C., Ceccarelli, C., Hollenbach, D., & Kaufman, M. 2005, ApJL, 635, L85
Duflot, D., Toubin, C., & Monnerville, M. 2021, FrASS, 8, 24
Dulieu, F., Amiaud, L., Congiu, E., et al. 2010, A&A, 512, A30
Dulieu, F., Congiu, E., Noble, J., et al. 2013, NatSR, 3, 1338
Elitzur, M., & de Jong, T. 1978, A&A, 67, 323
Feller, D. 1993, JChPh, 98, 7059
Ferrero, S., Grieco, F., Ibrahim Mohamed, A., et al. 2022, MNRAS, 516, 2586
Ferrero, S., Zamirri, L., Ceccarelli, C., et al. 2020, ApJ, 904, 11
Fichthorn, K. A., & Miron, R. A. 2002, PhRvL, 89, 196103
Fischer, J., Luhman, M. L., Satyapal, S., et al. 1999, Ap&SS, 266, 91
Fraser, H. J., Collings, M. P., McCoustra, M. R. S., & Williams, D. A. 2001, MNRAS, 327, 1165
Freedman, D., & Diaconis, P. 1981, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 57, 453
Germain, A., Corno, M., & Ugliengo, P. 2021, Computational Science and Its Applications—ICCSA 2021 (Cham: Springer International Publishing), 632
Germain, A., Tinacci, L., Pantaleone, S., Ceccarelli, C., & Ugliengo, P. 2022a, ESC, 6, 1286
Germain, A., Tinacci, L., Pantaleone, S., Ceccarelli, C., & Ugliengo, P. 2022b, ACO-FROST: Generation of Icy Grain Models and Adsorption of Molecules, v1, Zenodo, doi:10.5281/zenodo.7385933
Germain, A., Tinacci, L., Pantaleone, S., Ceccarelli, C., & Ugliengo, P. 2022c, Amorphous Grain Model of Interstellar Icy Grains, v1, Zenodo, doi:10.5281/zenodo.7385781
Germain, A., & Ugliengo, P. 2020, Computational Science and Its Applications—ICCSA 2020 (Cham: Springer International Publishing), 745
Gibb, E. L., Whittet, D. C. B., Boogert, A. C. A., & Tielens, A. G. G. M. 2004, ApJS, 151, 35
González-Alfonso, E., Smith, H. A., Fischer, J., & Cernicharo, J. 2004, ApJ, 613, 247
Grassi, T., Bovino, S., Caselli, P., et al. 2020, A&A, 643, A155
Grimme, S. 2006, JCoCh, 27, 1787
Grimme, S., Bannwarth, C., & Shushkov, P. 2017, JCTC, 13, 1989
Grimme, S., Ehrlich, S., & Goerigk, L. 2011, JCoCh, 32, 1456
Guo, Y., Riplinger, C., Becker, U., et al. 2018, JChPh, 148, 011101
Guo, Y., Riplinger, C., Liakos, D. G., et al. 2020, JChPh, 152, 024116
Hartmann, L., Ciesla, F., Gressel, O., & Alexander, R. 2017, SSRv, 212, 813
Hasegawa, T. I., Herbst, E., & Leukng, C. M. 1992, ApJS, 82, 167
Haynes, D. R., Tro, N. J., & George, S. M. 1992, JPhCh, 96, 8502
He, J., & Vidali, G. 2014, FaDi, 168, 517
Herpin, F., Chavarría, L., van der Tak, F., et al. 2012, A&A, 542, A76
Hogerheijde, M. R., Bergin, E. A., Brinch, C., et al. 2011, Sci, 334, 338
Hollenbach, D., Kaufman, M. J., Bergin, E. A., & Melnick, G. J. 2009, ApJ, 690, 1497
Humphrey, W., Dalke, A., & Schulten, K. 1996, J. Mol. Graphics, 14, 33
Imanishi, M., Nakanishi, K., Izumi, T., & Baba, S. 2022, ApJ, 926, 159
Jenniskens, P., Blake, D. F., Wilson, M. A., & Pohorille, A. 1995, ApJ, 455, 389
Kaufman, M. J., & Neufeld, D. A. 1996, ApJ, 456, 611
Kendall, T., Jr. 1992, JChPh, 96, 6796

Kolasinski, K. 2002, Surface Science: Foundations of Catalysis and Nanoscience (Chichester: Wiley)
Kristensen, L. E., van Dishoeck, E. F., Bergin, E. A., et al. 2012, A&A, 542, A8
Kruse, H., & Grimme, S. 2012, JChPh, 136, 04B613
Lamberts, T., Samanta, P. K., Köhn, A., & Kästner, J. 2016, PCCP, 18, 33021
Langer, W. D., & Graedel, T. E. 1989, ApJS, 69, 241
Lécuyer, C., Gillet, P., & Robert, F. 1998, ChGeo, 145, 249
Leger, A., Klein, J., de Cheveigne, S., et al. 1979, A&A, 79, 256
Löfgren, P., Ahlström, P., Lausma, J., Kasemo, B., & Chakarov, D. 2003, Langm, 19, 265
Martín-Doménech, R., Caro, G. M., Bueno, J., & Goesmann, F. 2014, A&A, 564, A8
Mayhall, N. J., Raghavachari, K., & Hratchian, H. P. 2010, JChPh, 132, 114107
Melnick, G. J., & Bergin, E. A. 2005, AdSpR, 36, 1027
Minissale, M., Aikawa, Y., Bergin, E., et al. 2022, ESC, 6, 597
Molpeceres, G., Kästner, J., Herrero, V., Peláez, R., & Maté, B. 2022, A&A, 664, A169
Molpeceres, G., Rimola, A., Ceccarelli, C., et al. 2019, MNRAS, 482, 5389
Molpeceres, G., Zaverkin, V., & Kästner, J. 2020, MNRAS, 499, 1373
Morbidelli, A., Bitsch, B., Crida, A., et al. 2016, Icar, 267, 368
Morbidelli, A., Blanc, M., Alibert, Y., et al. 2019, The Delivery of Water to Protoplanets, Planets and Satellites (Berlin: Springer)
Neese, F. 2018, Wiley Interdisciplinary Reviews: Computational Molecular Science, 8, e1327
Oba, Y., Miyauchi, N., Hidaka, H., et al. 2009, ApJ, 701, 464
Öberg, K. I., Linnartz, H., Visser, R., & van Dishoeck, E. F. 2009, ApJ, 693, 1209
Öberg, K. I., & Wordsworth, R. 2019, AJ, 158, 194
Omont, A., Yang, C., Cox, P., et al. 2013, A&A, 551, A115
Pavošević, F., Peng, C., Pinski, P., et al. 2017, JChPh, 146, 174108
Pensabene, A., van der Werf, P., Decarli, R., et al. 2022, A&A, 667, A9
Petersson, a., Bennett, A., Tensfeldt, T. G., et al. 1988, JChPh, 89, 2193
Podio, L., Kamp, I., Codella, C., et al. 2013, ApJL, 766, L5
Potapov, A., Jäger, C., & Henning, T. 2018, ApJ, 865, 58
Riplinger, C., Sandhoefer, B., Hansen, A., & Neese, F. 2013, JChPh, 139, 134101
Rosu-Finsen, A., Chikani, B., & Salzmann, C. G. 2022, MNRAS, 517, 1919
Ruaud, M., & Gorti, U. 2019, ApJ, 885, 146
Rybicki, M., & Sauer, J. 2022, JCTC, 18, 5618
Sack, N. J., & Baragiola, R. A. 1993, PhRvB, 48, 9973
Sandford, S. A., & Allamandola, L. J. 1988, Icar, 76, 201
Sandler, I., Chen, J., Taylor, M., Sharma, S., & Ho, J. 2021, JPCA, 125, 1553
Smith, J. A., Livingston, F. E., & George, S. M. 2003, JPCB, 107, 3871
Speedy, R. J., Debenedetti, P. G., Smith, R. S., Huang, C., & Kay, B. D. 1998, JChPh, 105, 240
Sprowl, L. H., Campbell, C. T., & Arnadottir, L. 2016, J Phys. Chem. C, 120, 9719
Stoychev, G. L., Auer, A. A., & Neese, F. 2017, JCTC, 13, 554
Tait, S. L., Dohnálek, Z., Campbell, C. T., & Kay, B. D. 2005, JChPh, 122, 164708
Tielens, A. G. G. M., & Hagen, W. 1982, A&A, 114, 245
Tinacci, L., Germain, A., Pantaleone, S., et al. 2022, ESC, 6, 1514
Ulbricht, H., Zacharia, R., Cindir, N., & Hertel, T. 2006, Carbon, 44, 2931
van Dishoeck, E. F., Herbst, E., & Neufeld, D. A. 2013, ChRv, 113, 9043
van Dishoeck, E. F., Kristensen, L. E., Mottram, J. C., et al. 2021, A&A, 648, A24
Wakelam, V., Chapillon, E., Dutrey, A., et al. 2019, MNRAS, 484, 1563
Wakelam, V., Loison, J. C., Mereau, R., & Ruaud, M. 2017, MolAs, 6, 22
Weigend, F., Köhn, A., & Hättig, C. 2002, JChPh, 116, 3175
Westall, F., & Brack, A. 2018, SSRv, 214, 50
Yang, C., González-Alfonso, E., Omont, A., et al. 2020, A&A, 634, L3
Yang, C., Omont, A., Beelen, A., et al. 2016, A&A, 595, A80
Yocum, K. M., Smith, H. H., Todd, E. W., et al. 2019, JPCA, 123, 8702