Computational Estimation of the Binding Energies of $PO_x$ and $HPO_x$ ($x = 2, 3$) Species

Elettra L. Piacentino © and Karin I. Öberg ©

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Received 2022 May 16; revised 2022 September 14; accepted 2022 September 29; published 2022 November 9

Abstract

The distribution of molecules between the gas and solid phase during star and planet formation determines the trajectory of gas and grain surface chemistry, as well as the delivery of elements to nascent planets. This distribution is primarily set by the binding energies of different molecules to water ice surfaces. We computationally estimated the binding energies of 10 astrochemically relevant P-bearing species on water surfaces. We also validate our method for 20 species with known binding energies. We used Density Functional Theory (DFT) calculations (M06-2X/aug-cc-pVDZ) to calculate the energetics of molecules and water-molecule clusters (1–3 H$_2$O molecules) and from this determined the binding energy by comparing the complex and the separate molecule and cluster energies. We also explore whether these estimates can be improved by first calibrating our computational method using experimentally measured binding energies. Using the 20 reference molecules we find that the 2H$_2$O cluster size yields the best binding energy estimates and that the application of a calibration to the data may improve the results for some classes of molecules, including more-refractory species. Based on these calculations we find that small P-bearing molecules such as PH$_3$, PN, PO, HPO, PO$_2$, and POOH are relatively volatile and should desorb prior or concomitantly with water ice, while H$_2$PO, HPO$_2$, PO$_3$, and PO$_2$OH can strongly bind to any hydroxylated surface and will likely remain on the interstellar grains surface past the desorption of water ice. The depletion of P carriers on grains constitutes a pathway for the inclusion of phosphorous molecules in planets and planetesimals.

Unified Astronomy Thesaurus concepts: Interstellar molecules (849); Theoretical models (2107); Astrochemistry (75); Interstellar abundances (832)

1. Introduction

The relative abundance of chemical species in the ice and gas phase in astrochemical environments is of fundamental importance to predict the composition of planets and planetesimals (Öberg et al. 2009; Williams & Cieza 2011; Marboeuf et al. 2014). In the early stages of star formation the chemical distribution between ice and gas phase determines the chemical inventory that can be accessed for gas phase and surface chemistry on grains. In disks the balance between adsorption and desorption processes regulates which molecules can be incorporated into planets and planetesimals and which can only be delivered to planets through gas accretion. The temperature gradient present in disks results in the formation of condensation lines and therefore a radius dependent chemical distribution (Öberg et al. 2011; Bergin et al. 2015; Öberg & Bergin 2021). The location of these condensation lines depends on the strength of this physisorption interaction—the binding energy (BE)—between a molecule and a solid surface (i.e., water ice, silicate, and carbonaceous grains) and consequently its availability in condensed or gaseous phase at a particular disk radius.

Among the biogenic elements, the distribution of phosphorous (P) containing molecules between gas and solid phases is perhaps most uncertain. The availability of P is key to the formation of several biotic molecules (Pasek & Lauretta 2005) and it is quite abundant on Earth (P/H ~ 10$^{-7}$; Fagerbakke & Heldal 1996). In comparison the cosmic abundance of P relative to hydrogen is much lower (P/H ~ 2.57 × 10$^{-7}$, Asplund et al. 2009). Phosphorus carried by PN and PO has been detected in the gas phase around evolved stars (Tenenbaum et al. 2007; Ziurys et al. 2007; Milam et al. 2008) and in star-forming regions (massive: Fontani et al. 2016; Rivilla et al. 2018; low mass: Yamaguchi et al. 2011; Lebock et al. 2016; Bergner et al. 2019). In all these cases the P abundance is low accounting for a P/H of about 10$^{-10}$–10$^{-9}$ (Lebock et al. 2016; Rivilla et al. 2018; Bergner et al. 2019; Rivilla et al. 2020). In circumstellar envelopes, such as around IRC +10216, phosphorous, carried by PH$_3$ and HCP, has been observed with abundances of 10$^{-9}$ with respect to molecular hydrogen (Agúndez et al. 2012). This accounts for about 7% of the phosphorous elemental abundance (Agúndez et al. 2014). These evidences suggest that the undetected phosphorous is likely incorporated into grains and that there must therefore be additional less-volatile carriers of P in the interstellar medium (ISM).

The nature of less-volatile P-containing compounds in the ISM is currently unclear, but solar system studies may provide some clues. Analysis of CI chondrites have shown an elemental P abundance similar to the solar phosphorus abundance (Lodders 2003). In stony meteorites most of the phosphorous is carried by Ca, Mg-phosphate minerals, while reduced phosphorous is more common in Fe-rich meteorites (Pasek et al. 2004). Very recently volatile phosphorous, mainly carried by PO fragments, was detected on comet 67P/Churyumov-Gerasimenko during the Rosetta mission ((Altwegg et al. 2016; Rubin et al. 2019; Gardner et al. 2020). This might suggest that, also in the ISM, P is incorporated into a relatively refractory phase, which likely consists of species containing PO$_3$ moieties.

In this study we address the possible distributions of phosphorous oxides in the ISM and disks through a theoretical
In this work we aim to computationally constrain the BEs of three known ISM P carriers as well as seven proposed interstellar P carriers over ASW. We will use two approaches described in detail in Section 2: direct calculation and calculations calibrated against experimental data on C-, N-, O-, and S-bearing molecules. The results of the two methods are presented in Section 3 where we also discuss the reliability of these approaches. In Section 4 we present and comment on the application of these methods to P-bearing molecules. We include some astrophysical implications of our new BEs for P-bearing molecules (Section 4.3). Finally, in Section 5 we summarize our findings.

2. Methods

2.1. Computational Details and Cluster Size Choice

When computing BEs there are several aspect of the process, such as the computational tools, the molecular approximations, and data treatment, for which choices need to be made.

To model the binding interactions we chose to use electronic structure based methods. In particular, within the density functional theory (DFT) we chose to use the M06-2X functional for its good performances in modeling noncovalent interactions (Mardirossian & Head-Gordon 2017). All calculations are run using the Gaussian 16 suite of software (Frisch et al. 2016) at the M06-2X/aug-cc-pVDZ level of theory (Dunning 1989; Kendall et al. 1992; Zhao & Truhlar 2008) and included the optimization of the clusters geometry to a stationary point as well as vibrational frequency calculations for the identification of the energy minima. Additionally we tested the relative performances of MollerPlesset (MP) methods (Frisch et al. 1990), in particular at the MP2/aug-cc-pVDZ level of theory, for BE determination and we concluded that the two methods yields equivalent results (Appendix A). Previous work by Wakelam et al. (2017) and Ferrero et al. (2020) also chose the M06-2X functional for the estimation of BEs.

Concerning the description of the surface–molecule interaction, we chose to focus our work on cluster systems rather then on a periodic representation of the water surface due to its computational affordability. The use of small-cluster systems rather than a more extended periodic representation of the binding surface could lead to less accurate results. In the work from Ferrero et al. (2020) the authors show, however, that the BEs calculated using small water clusters are comparable to more expensive and/or complex periodic calculations. In our study the cluster size was purposefully kept small in the attempt to minimize the computational cost to enable easy scaling to larger molecular data sets. Inspired by the work of Das et al. (2018), who found an improvement in accuracy going from the monomeric to the tetrameric representation of ASW; while 5 and 6H2O clusters did not further improve the accuracy, we evaluate the impact that the size of the water cluster has on the BE estimation by using 1H2O, 2H2O, and 3H2O water cluster sizes. We also follow Wakelam et al. (2017) and explore whether directly calculated BEs can be improved through calibration against experimental values using a molecular training set. We extend this work by applying a calibration to the BEs calculated using 1–3 H2O water cluster sizes. To obtain a calibration set of molecules, we searched the literature for experimental data on BEs to ASW surfaces and identified

1 Data set is available at doi:10.5281/zenodo.6551710.
Table 1
Calculated and Experimentally Determined Binding Energies (in Kelvin) for Our Reference Molecules

| M06-2X | Direct Method | Calibrated Method | Exp. b |
|---|---|---|---|
| aug-cc-pVDZ | 1H2O | 2H2O | 3H2O | 1H2O | 2H2O | 3H2O | 1H2O | 2H2O | 3H2O |
| (1) | N2 | 530 | 998 | 953 | 1803 | 1723 | 1898 | 11251,2 | 11651,2,3,4 |
| (2) | CO | 504 | 1191 | 1107 | 1769 | 1867 | 2048 | 13702,5 | 23394,5,6 |
| (3) | CH4 | 672 | 1212 | 1168 | 1990 | 1882 | 2104 | 24007 | 24097 |
| (4) | CO2 | 1733 | 3104 | 2259 | 3395 | 3289 | 3170 | 23397 | 24097 |
| (5) | C2H6 | 960 | 1600 | 1321 | 2372 | 2171 | 2256 | 24957 | 25198,9 |
| (6) | C2H4 | 1498 | 2760 | 2294 | 3084 | 3034 | 3204 | 24007 | 24097 |
| (7) | H2S | 960 | 1600 | 1321 | 2372 | 2171 | 2256 | 24957 | 25198,9 |
| (8) | C3H6 | 1388 | 3458 | 2413 | 2938 | 3552 | 3320 | 30007 | 30007 |
| (9) | H2CO | 3076 | 4623 | 3154 | 3273 | 3663 | 4042 | 25198,9 | 25198,9 |
| (10) | CH3CN | 2445 | 5093 | 4148 | 2938 | 3552 | 3320 | 30007 | 30007 |
| (11) | CH3NC | 2352 | 4484 | 3680 | 3273 | 3663 | 4042 | 25198,9 | 25198,9 |
| (12) | NH3 | 3464 | 6183 | 3826 | 4520 | 5759 | 4696 | 541013 | 541013 |
| (13) | CH2CCH2 | 1792 | 2957 | 2145 | 3472 | 3180 | 3059 | 44007 | 44007 |
| (14) | CH2CH2 | 2280 | 3422 | 2251 | 4118 | 3525 | 3162 | 44007 | 44007 |
| (15) | HCl | 2836 | 5225 | 4743 | 4854 | 4866 | 5589 | 517012 | 517012 |
| (16) | CH2OH | 2584 | 6183 | 3826 | 4520 | 5759 | 4696 | 541013 | 541013 |
| (17) | NH4+ | 3464 | 6183 | 3826 | 4520 | 5759 | 4696 | 541013 | 541013 |
| (18) | CH3NC | 2352 | 4484 | 3680 | 4213 | 4315 | 4555 | 568618 | 568618 |
| (19) | H2O | 2633 | 6007 | 5416 | 4586 | 5448 | 6245 | 577516 | 577516 |
| (20) | CH3CN | 2445 | 5093 | 4148 | 4337 | 4769 | 5010 | 615015 | 615015 |

References: (1) Fayolle et al. (2016), (2) Smith et al. (2016), (3) Collings et al. (2003), (4) Noble et al. (2012a), (5) He et al. (2016), (6) Gálvez et al. (2007), (7) Behmard et al. (2019), (8) Wakelam et al. (2017), (9) Penteado et al. (2017), (10) Noble et al. (2012b), (11) Chaabouni et al. (2020), (12) Olarewaju (2011), (13) Bahr et al. (2008), (14) Hama & Watanabe (2013), (15) Bertin et al. (2017), (16) Fraser et al. (2001), (17) Fraser et al. (2001).

The Astrophysical Journal, 939:93 (12pp), 2022 November 10
Piacentino & Oberg

20 molecules with well-defined experimental BEs (Table 1). The same literature data were also used to benchmark the performance of our computational results by direct comparison to the experimental values.

Concerning the uncertainties of the experimental BEs, we apply a 10% uncertainty to the BEs of species for which an uncertainty value of <10% was reported in the literature and we conservatively consider an uncertainty of 30% when the uncertainties were not reported in the literature (see Table 1).

The direct method, which estimates the BEs from the energetics of 1H2O, 2H2O, and 3H2O water cluster sizes calculation is described in detail in Section 2.2. Section 2.3 describes the application of the calibrated method to the direct method data.

2.2. Direct Method

We first calculate the BEs at different cluster sizes directly. The electronic energy of each molecule, water cluster, and water-molecule complex was calculated at the M06-2X/aug-cc-pVDZ level of theory for the 1, 2, and 3 H2O cluster representation. The geometry of the water-molecule complexes is optimized starting from a noninteraction configuration with the molecule placed at a minimum distance >5 Å from the water cluster. This is to ensure that no molecule–cluster interactions are present in the input geometry. For most molecules we performed a single optimization calculation (single initial configuration), which yielded geometries in agreement with the criteria described below. For molecules that allow multiple unique interaction configurations we repeated the optimization with two to three initial configurations.

The following criteria were used to choose the representative geometry: In the 1H2O clusters the representative geometry was chosen as the one where the main interaction was between H2O and the molecule. This is because the ASW surface is more rich in hydrogen than oxygen atoms and therefore the H2O–molecule interaction is more likely to occur (Wakelam et al. 2017). Generally the H2O interacts with an atom of the molecule but in the cases of unsaturated hydrocarbons the interaction is set between the H2O and the double or triple bond on the carbon chain. Exceptions to this criteria are CO2, for which the main interaction is always set between the OOH and the OHH, which sees the main interaction occurring between the hydroxylic OHH and the OHH. These exceptions in the final geometry are a result of the optimization calculation and are likely due to the nature of the molecules. The 2H2O and 3H2O complex geometries are chosen in a similar way while allowing for the second and third interaction to occur between the OHH and the molecules. An example of the optimized geometries is shown in Figure 1 for the molecule PH3. A few more optimized geometries are shown in Appendix B.

In the direct method, the adsorption energies of the molecules in question is determined from the variation of the energy of an adsorbent molecule and an H2O cluster (in our case constituting 1–3 H2O water molecules) that arises when they are able to noncovalently coordinate with each other. The
The Astrophysical Journal, 939:93 (12pp), 2022 November 10

Piacentino & Öberg

Table 2

Binding Energies of Selected P-bearing in Kelvin

| M06-2X | Direct Method (K) | Calibrated Method (K) |
|--------|-------------------|-----------------------|
|        | 1H2O (±51.6%)     | 2H2O (±19.2%)         | 3H2O (±22.4%)         |
|        | 1H2O (±27.2%)     | 2H2O (±25.3%)         | 3H2O (±27.9%)         |
| PH₃    | 1117              | 2642                  | 1228                  |
|        | 2579              | 2945                  | 2166                  |
| PN     | 2168              | 5119                  | 2376                  |
|        | 3970              | 4787                  | 3284                  |
| PO     | 3791              | 9176                  | 9221                  |
|        | 6117              | 7804                  | 9951                  |
| HPO    | 4176              | 9964                  | 5304                  |
|        | 6626              | 8390                  | 6136                  |
| OPO    | 4327              | 10186                 | 11031                 |
|        | 6826              | 8555                  | 11714                 |
| PO₃OH  | 625               | 9512                  | 4551                  |
|        | 9382              | 8054                  | 5403                  |
| PO₂OH  | 7254              | 12290                 | 25502                 |
|        | 10698             | 10119                 | 25809                 |
| HPO₂   | 7457              | 12740                 | 15388                 |
|        | 10967             | 10454                 | 15958                 |
| H₃PO   | 9118              | 11112                 | 8585                  |
|        | 13164             | 9244                  | 9597                  |
| PO₃    | 11039             | 18343                 | 30441                 |
|        | 15706             | 14621                 | 30619                 |

Note.

* These structures react with the ASW to form a PO₄ moiety. The energy reported refer to a chemisorption events.

BE is calculated as follows:

\[
    \text{BE} = E_{\text{complex}} - (E_{\text{molecule}} + E_{H_2O\text{cluster}}).
\]

Where \( E_{\text{complex}} \) is the energy of one of the potential energy minima geometries of the water-molecule cluster when the species is physisorbed onto the surface of the cluster (2–5 Å), \( E_{\text{molecule}} \) is the energy of the species alone, and \( E_{H_2O\text{cluster}} \) is the calculated energy of the water cluster.

The calculated electronic energies are used directly without accounting for the zero-point energy (ZPE) and the basis set superposition error (BSSE) similarly to Wakelam et al. (2017). Wakelam et al. (2017) tested whether the inclusion of ZPE and BSSE significantly affected the accuracy of the resulting fit and found that, in the dimer case, the inclusion of the corrections slightly reduced the goodness of the fit. This suggests that, while the omission of either the ZPE or the BSSE may significantly affect the accuracy of the resulting fit, the omission of both corrections in the estimation of BEs yields values that better approximate the experimental BE (Das et al. 2018). This errors compensation is likely just a fortuitous but nonetheless advantageous balance. The direct method is evaluated for a reference set of 20 molecules (Table 1) and then applied to calculate the BEs of 10 P-bearing species (Table 2).

2.3. Calibrated Method

Wakelam et al. (2017) demonstrated that there is a systematic offset between calculated and experimentally determined BEs when the BEs are calculated using a single-water-molecule cluster. They also show that the accuracy of the calculated BE estimates can be improved if they are calibrated against experiments. Wakelam et al. (2017) calculated, using the equation in Section 2.2, the BEs of 16 molecules using a 1H₂O ASW representation to build a calibration curve against the experimental BE of each molecule. From the fit, they evaluated the interaction correlation between a 1H₂O representation and the ASW. Building upon the work of Wakelam et al. (2017), we extended the method by applying the calibrated method to the 1–3 H₂O systems to evaluate the effects of the increased cluster size on the performances of the calibrated method. We also evaluate whether the use of the calibrated method provides a significant improvement in the BE estimation over the direct method.

The calibration curves were built by fitting the 20 BEs (20 for each of the 3 water cluster systems) obtained using the direct method against the experimental BE’s values assuming a linear relationship between calculated and experimental values. The fit was then applied to the BE’s calculated with the direct method resulting in the calibrated method estimation of the BE values. The calibrated method was then applied to estimate the BEs of the P species.

3. Methods Validation

Table 1 summarizes the calculated direct method and estimated calibrated method BE values for the calibration molecules. Figure 2 shows the BEs calculated using the direct method and the calibrated method as well as the percentage deviation from the experimental values. Below we present these results in detail. In Section 4 we present the application of the two methods to our selection of phosphorous molecules.

3.1. Direct Method H₂O Cluster versus ASW Binding Energies

Figure 2 and Table 1 show that the accuracy of the calculated direct method BEs is highly variable both between different molecules, and when using different cluster sizes. Calculations accounting for 1H₂O interaction performed the worst, with deviations as high as 72% and a median deviation of 52%, while the 2 and 3H₂O clusters performed better with deviations between a few percent and 54% and median deviations of 19% and 23%, respectively. These deviations can be compared to typical experimental errors of 10%. The performance improvement when increasing the cluster size from one to two or three H₂O molecules is visualized in Figure 3, which shows histogram plots of the deviations in percentage from experimental values. In other words, there is a real increase in performance when increasing the cluster size from one to two, but not when increasing it to three for our sample of molecules. This suggests that increasing the cluster size beyond two H₂O molecules may be of limited value when the focus is to computationally determine the mean value of a molecule BE using the cluster approach. However, this needs to be confirmed for a larger and more diverse sample of molecules, as well as for a larger range of cluster sizes. While the small-cluster approach is a useful and computationally inexpensive tool to determine the mean BE values, it provides limited information on the BE distribution and therefore cannot be
applied in cases when the whole range of the BE distribution is of interest.

Interestingly the performance of the direct method approach appears to depend on the strength of the binding interaction. Figure 2 shows that for molecules up a volatility of \( \sim 3000 \text{ K} \) the median deviations are only \( \sim 15\% \) for the \( 2\text{H}_2\text{O} \) and \( 3\text{H}_2\text{O} \) cluster calculations, while there is almost a 1.5 factor increase when considering the less-volatile species. There are also evidences that the accuracy of the direct method approach depends on the chemical nature of the molecule. The BEs prediction for hydrocarbons (i.e., \( \text{C}_2\text{H}_6, \text{C}_3\text{H}_8, \text{C}_2\text{H}_2\text{CCH}, \text{CH}_3\text{CCH} \)) seem to be less accurate than for other molecules with a deviation from the experimental values between 25% and 53% for the \( 2\text{H}_2\text{O} \) cluster size (Figure 2, panels (a) and (c)).

Finally we note that there are a handful of molecules for which the deviation from experimental values increases between the \( 1\text{H}_2\text{O} \) and \( 2\text{H}_2\text{O} \) clusters; namely, \( \text{CO}_2, \text{H}_2\text{S}, \text{H}_2\text{CO}, \) and \( \text{CH}_3\text{NH}_2 \). These molecules warrant further investigation, since this may be revealing something interesting about their interactions with water ice. For now we simply note that in each of these cases the BE proceeds from being slightly underpredicted in the case of \( 1\text{H}_2\text{O} \) to overpredicted for the \( 2\text{H}_2\text{O} \) cluster.

3.2. Consideration on the Direct Binding Energies Calculation

A first measure of the accuracy of the direct method BE calculations is the deviation from experimental values. As reported above the deviation from experimental BEs decreases from 52% to 19% (Figure 3) when increasing the cluster size from 1 to \( 2\text{H}_2\text{O} \), and no further improvement is seen when increasing the cluster size from 2 to \( 3\text{H}_2\text{O} \) (22% median deviation). The improvement when going from 1 to \( 2/3\text{H}_2\text{O} \) clusters is in line with the results from Das et al. (2018), who found an improvement in the deviation from the experimental values from 40% to 25% when increasing the cluster size from 1 to \( 3\text{H}_2\text{O} \) molecules (\( 2\text{H}_2\text{O} \) cluster calculations were not included in the study). Das et al. (2018) also reported a consistent underestimation of the BE values compared to the
experimental values, which is present for both the monomeric and the trimeric representation of the ASW, additionally showing a trend of progressively less negative estimations as the cluster size increases. Our calculations for the 1 and 3H₂O clusters show a similar trend where the 3H₂O system generally underestimate the experimental BEs while providing a substantial improvement form the BEs derived from the 1H₂O system.

However, we find that the 2H₂O system does not fit in the same trend, underestimating only 45% of the BEs studied. The randomness of this distribution along with the good absolute median uncertainty of 19% suggests that, among the cluster size and geometries that are considered in this study, the 2H₂O description of the ASW surface provides the best BEs prediction.

The improvement in the results when using 2/3H₂O clusters versus 1H₂O is intuitively due to the accounting for additional interactions between the molecule and the water cluster. The presence of more water molecules allows for additional binding constrains between the molecule and the water cluster yielding a binding geometry that better resembles the binding configuration on the ice. We can test this intuition by interrogating the molecule–cluster systems in detail. We find that the addition of a second water molecule results in a geometry where the primary water molecule can bond more strongly with the molecule in question. Since in an ice system there are always neighboring molecules, accounting for this distortion is important to produce accurate BEs. Not accounting for this geometrical distortion yields underestimated BE: this is consistent with the underestimation (Figure 4) of the BEs in the 1H₂O system where not all the fundamental interaction between the species and the ASW can be taken into account.

Figure 5 shows an example (NH₃) of the effects on the molecules binding environment caused by the second H₂O. The main interaction between the ASW and NH₃ is between the nitrogen on NH₃ and one of the water hydrogens. In the case of 2H₂O the interaction distance is shorter (1.96 and 1.86 Å respectively in the 1 and 2H₂O systems). A consequent elongation of the O-H bond in the primary water molecule is also observed. The second water molecule does not interact as strongly with the NH₃ but provides an additional anchoring point (2.15 Å) resulting in the reduction of the N–(H-O)H₂O angle from 171° to 159°. A similar behavior is generally observed across the studied molecules.

Following the same intuition as above, we should observe an improved BE accuracy when increasing the H₂O cluster size from 2 to 3H₂O, but this is not what we find. The presence of the third water molecule introduces additional structural constraints (2.46 Å, Figure 5), which cause a weakening of
the primary and the secondary interaction between the molecule and the water. This results in a BE prediction for the 3H2O system very similar to the 2H2O system with the absolute error for ammonia going from 8% in the 2H2O to 13% in the 3H2O. The observed lack of increased precision when increasing the cluster size from 2 to 3 H2O molecules is surprising. It is typically expected for larger clusters to increase the accuracy of the BE estimation as more long-range interactions can be taken into account and less unique geometries become available. One possible explanation is that the use of a greater number of water molecules introduces significant freedom in regard to the arrangement of the water molecules themselves and this may produce geometries at odds with ASW. In our study we found that the BE estimations are dependent on the functional group that the molecule uses to bind to the water cluster. This is most apparent when comparing the isomers CH3CN and CH3NC (see Figure 9 in the Appendix B). For CH3CN and CH3NC the BE estimates reflect the interactions of the water with either the N or C atoms that are terminal to the molecule. In the case of CH3NC we have that the interaction distance is 2.18 Å, which becomes 2.07 Å for CH3CN, consequently increasing the BE of the molecule. This is consistent with the experimental BE and it reflects the affinity of the water for the functionalities present in the molecules. The binding distance is further shortened for molecules having a terminal oxygen; in the case of H2CO for example, the OCH–H2O distance is 1.98 Å.

We find a similar functionality dependency of the BE to water in cases where the main binding interaction occurs between the H2O and the double/triple bond in the molecules. In the case of CH2CCH2 and CH2CCH, for example, we find that CH2CCH binds strongly to the water cluster when compared to CH2CCH2. This is because the triple bond in CH2CCH constitutes a better binding functionality for hydrogen than the double bond in CH2CCH2. This effect of the molecule saturation on the BE prediction is observed also in the C2 and C3 series of hydrocarbons with smaller BEs as the saturation of the molecule increases (Table 1). This same trend has also been studied experimentally by Behmard et al. (2019) who observed a similar trend showing that the BE of C2 and C3 hydrocarbons decreases with the saturation of the molecules (i.e., BEC2,H6 > BEC2,H5 > BEC2,H4). In general, the application of the direct method to hydrocarbons seems to yield a worse approximation of their BEs compared to other molecules in the study. The median deviation for nonhydrocarbons is 16% and for hydrocarbons 23%, in the 2H2O direct method calculation. The underestimation of the BE of hydrocarbons may be associated to the poor natural affinity that this class of molecules has toward water. In such cases the size of the cluster may be more relevant than for other classes of molecules and the use of a periodic representation of the ASW may be of aid. The presence of a matrix may help producing a tighter packing of the water around the molecule, which would result in a binding structure that better resembles the experiments. The poor constrain that the direct method provides for hydrocarbons suggests that this class of molecules may also especially benefit from the calibrated method.

3.3. Calibrated Method Improvements over the Direct Method

We next evaluate whether the results obtained through the direct method can be improved upon application of the calibrated method using the same 20 molecules. Figure 4 shows the linear fit of the BE calculated using the direct method and the experimental values. The calculated BEs using 1 and 3H2O generally underestimate the BE. This effect is more pronounced and less uniform for the 1H2O compared to the 3H2O data set. The use of 2H2O appears to instead produce a close to random scatter around the expected values. We note, however, that as we go from more- to less-volatile molecules the 2H2O cluster method seems to be systematically over-predict the BEs.

We find that the calibrated method BE predictions are very similar to each other for all the cluster sizes. By contrast to the direct method, we find no improvement in the median deviation from the experimental value as the cluster size is increased (Figure 3). For 1, 2, and 3H2O clusters we find median deviations of 27%, 25%, and 28%, and deviation ranges of 1%–60%, 2%–60%, and 2%–69%, respectively. This implies that the calibrated method generally achieves a higher level of accuracy when considering single-water-molecule clusters, but a comparable level of accuracy when considering larger clusters in comparison with the direct method.

Using the calibrated method we see an opposite dependence of the results accuracy with the molecule volatility compared to the direct method (Figure 2, panel d)). For volatile molecules having BE <3000 K, the median deviation from the experimental values calculated using the calibrated method—2H2O cluster size is 37%. The median deviation is reduced to 19% for the less-volatile group of molecules, which implies that the calibrated method outperforms the direct method for the more-refractory molecules. It also appears to do better with the 2 and 3C hydrocarbons compared to the direct method.

Additionally we tested the importance of the number of calibration points used to estimate the correlation for uncertainties minimization; we performed 1000 random selections of 5, 10, and 15 molecules from our list of 20 molecules and for each set we used the resulting fits to determine the BEs of the molecules not included in the selections. We finally calculated the median of the errors across the 1000 draws for each of the 5, 10, and 15 molecule sets (Figure 6). The median deviation decreases as the calibration set increases, but the improvement is small: about 2%. Though it will be interesting to revisit the calibrated method approach with more experimentally determined values, we may already be close to the limit where the peculiarities of each molecule–H2O system dominates the calculated uncertainty.

3.4. Utility of Calibrating Calculated Binding Energies?

The application of the calibrated method is more or less advantageous depending on the cluster size used and the volatility of the molecules in the exam. Without a calibration the 1H2O BEs are too inaccurate to be useful, and we hence recommend that such a calibration is always used for 1H2O BE calculations. The application of the calibrated method to the 2 and 3H2O cluster sizes do not appear to contribute to the accuracy of the results except for some specific classes of molecules.

The improvements to the predicted BEs observed when applying the calibration to the less-volatile species in contrast with the loss of accuracy for highly volatile molecules (<3000 K) suggests that the calibrated method does not, on average, improve the results. However, we suspect that the calibration could improve the results if applied in a more targeted way. More experimental BE values are needed to separately calibrate volatile and nonvolatile species.
4. Application of the Methods to P Molecules

4.1. BE of P-bearing Molecules

We calculated the BEs of 10 phosphorous molecules (Table 2) using the direct method and the calibrated method. The uncertainties on the calculated BEs are derived from the median deviation from the experimental value obtained respectively for each of the methods (Figure 3). When estimating the uncertainties for the calibrated method of the P molecular BEs below, we use the median deviations for the fiducial 20-molecule calibration set. As expected from the calibration set, the direct method 3H2O systems estimations provides, for most of the P-bearing molecules, BEs values that are in between the values obtained using the 2H2O and the 3H2O direct method calculations (Table 2). Exceptions are OPO, and HPO2, for which the direct method 3H2O BE values are higher than the 2H2O prediction. Additionally, the geometry optimization of PO2OH and PO3 with 3H2O results in the coordination of the P species to one oxygen from the water cluster yielding a PO4 moiety (see Appendix C). This complexation prevents the calculation of the physisorption energy for these two P molecules for the 3H2O cluster system. We note that the complexation is observed only for the 3H2O cluster setting; in all other cluster size neither covalent interactions nor deformation of the water cluster geometries are observed.

Similarly to what was observed for the reference set of molecules we find that the also in the case of P-bearing molecules the application of the calibrated method estimates BEs values that are equivalent across the three cluster systems. We also find similarities between the fiducial set and the P-bearing set of molecules when comparing the two methods’ performance for each cluster size. In the case of 1H2O cluster size we find that, similarly to the fiducial set of molecules, the calibrated method estimations have higher BE values than the direct method by a factor of ~2.4 for low-desorbing species, and ~1.5 for the less-volatile species. The 2H2O cluster size yields a smaller discrepancy between the direct method and calibrated method prediction compared to the 1H2O cluster size with the 2H2O estimation yielding values within 15% of the direct method estimation. In the cases when the 3H2O clusters size did not yield additional coordination chemistry, the calibrated method estimation relates to the direct method estimation by a 1.8 and 1.2 factor respectively for volatile and less-volatile species.

In Figure 7 the BE values calculated for the P-bearing species are shown in relation to the BEs calculated for the fiducial set of molecules using the 2H2O system. With the exception of PH3, PO, and PN the estimated BEs of the P-bearing species are found to be quite high (BE ~ 8000 K). With a most of the molecules exceeding the range of the fiducial set of molecules (PO2OH, HPO2, H3PO, and PO3) with BE > 9000 K.

Considering the results obtained for the calibration set of molecules, we divided the P-bearing species into three groups based on their BEs. The first group consists of PH3 alone, which is the only highly volatile molecule with a BE below 3000 K when calculating energies for the P molecules. PH3 BE is estimated using the direct method. PN, PO constitute the second group, defined by BEs similar to or lower than water ice (Fraser et al. 2001). For these molecules we recommend using the calibrated method results. In the highest range of BEs we find the refractory group constituted by HPO, OPO, POOH, H3PO, PO2OH, HPO2, and PO3 for which the BEs exceed 8000 K. For these too we recommend using the calibrated method 2H2O but caution that the results are more uncertain since they extend beyond the calibration set.
range of the experimentally determined BEs that we evaluated our methods against. We therefore expect that our error estimations for these species are reliable. The \( \text{PH}_3 \) BE has been previously computed by Nguyen et al. (2021) at 1813–2690 K; this range is consistent with our calculation of 2642 ± 19%. Our values are also in agreement with the BE range calculated, using a 20-water-molecule cluster, by Molpeceres & Kastner (2021) who report 2189 K as the average BE and with the computational BE reported by Viana & da Silva (2015) of 3000 K calculated using a two-water-molecule cluster at the CCSD (T) and MP2 level of theory. Experimental works have shown that phosphine’s thermal sublimation occurs at around 60 K (Turner et al. 2015), consistent with a BE of ~1800 K using the formalism of Hollenbach et al. (2009). In summary, our computational method appears to be accurate for phosphine.

To our knowledge, there have not been experimental or computational studies on PO, PN. Next we turn to the P species with calibrated 2\( \text{H}_2\text{O} \) cluster BEs above 8000 K, which makes them effectively refractory in astrophysical environments. The calculated BEs for these species present two complications: First of all they fall outside of our calibration range and their BE error bars are therefore more uncertain. Second, in interstellar regions, they are not expected to desorb off water ice. However, we argue that the sublimation temperature on silicate grain should not substantially deviate from those expected on ASW. It has been shown that phosphine’s thermal sublimation occurs at around 60 K (Turner et al. 2015), consistent with a BE of ~1800 K using the formalism of Hollenbach et al. (2009). In summary, our computational method appears to be accurate for phosphine.

The abundance of phosphorous compounds detected in the gas phase at various stages of cloud evolution varies significantly. While only 1% of the expected phosphorous has been detected in star-forming regions (Rivilla et al. 2018; Bergner et al. 2019; Rivilla et al. 2020), phosphorous, in its ionic form, has been detected with solar abundances in diffuse clouds in the ISM (Jura & York 1978; Lebouteiller et al. 2006). This abundance discrepancy indicates that, during star and planet formation, the majority of the phosphorous is depleted on icy grains in semirefractory molecular carriers that have not yet been well constrained.

The majority of the P-bearing molecules that we explored in this work are found to be more refractory than water. This suggests that we should expect a semivolatile to refractory phosphorous reservoir that remains in the solid phase well after water sublimation. In disks, such species would remain solid interior to the water snow line.

The abundance of the PO-bearing species simulated in this work will largely depend on their condensed phase formation chemistry and on the specific environmental conditions. In the solar system, organic phosphonic acids (Cooper et al. 1992) as well as Ca-phosphate (Le Guillou et al. 2014) have been detected on the Murchison meteorite suggesting the possibility for the existence of a rich phosphorous chemistry in condensed phase. We are currently investigating different scenarios computationally to determine the fraction of locked P that is attainable at different ISM conditions (E. L. Piacentino et al., 2022 in preparation).

In comets, \( (\text{H}_3\text{PO}_4) \) species have been previously considered as possible phosphorous carriers (Rivilla et al. 2020). Although the investigation on the ROSINA data collected on comet 67P...
did not provide the direct detection of \((\text{H})_x\text{PO}_y\) species, it showed the presence of PO fragments that were attributed solely to PO molecules (Rivilla et al. 2020). As the fragmentation pattern of \((\text{H})_x\text{PO}_y\) species is not well constrained, it is also possible for the detected PO signal to include a contribution due to the fragmentation of bigger phosphorus molecules. In either case we agree with Rivilla et al.’s (2020) speculation that the PO signal is due to P-bearing molecules that were locked in the grains early on during star formation.

In our calculations we find that \(\text{PO}_2\text{OH}\) and \(\text{PO}_3\) can chemisorb on the water surface to coordinate with an additional oxygen atom. This can indicate the tendency of these PO-bearing species to further react with water molecules to form the likely more-refractory phosphate moiety. We speculate that the easiness with which phosphates may form could, depending on the environmental condition, lead to even a larger fraction of the phosphorous to be locked on grains in phosphate form.

In conclusion, while \((\text{H})_x\text{PO}_y\) molecules have not yet being directly detected, the clues that we have indicate that these high desorbing species are good candidates for phosphorous carriers during star and planet formation, and may be the starting point of phosphate formation through their interaction with water ice.

5. Summary and Conclusions

We explored the performances of a direct ab initio \(\text{H}_2\text{O}\) cluster calculation, and a calibrated version of the same, for BE estimations. We tested our methods using 20 molecules for which the experimental BEs are well constrained in the literature and then applied these methods to 10 P-bearing molecules. We found the following:

1. The direct method–2\(\text{H}_2\text{O}\) cluster-method combination performs better than any other method/cluster size that we evaluated. It provides a quick BE estimation that does not seem to carry systematical errors.
2. The application of the calibrated method improves the BE estimation for less-volatile molecules while it reduces the estimation accuracy for highly volatile molecules. This suggests that a targeted selection of the calibration set may be needed.
3. While the application of the calibrated method improves the results, the estimation of the BE for hydrocarbons seems to be difficult. This suggests that the use of a heterogeneous calibration set may not be optimal for all classes of molecules. A functional group-based study may help highlight the effect of molecular proprieties on the BE estimation.
4. The application of these computational methods to astrochemically interesting PO-containing species show that most of these species are more refractory than water.
5. The presence of semirefractory PO-containing species in disks beyond the water snow line could help explain the depletion of P in the ISM and it would supply a pathway for the inclusion of phosphorous in planets and planetesimals.

This work was supported by a grant from the Simons Foundation 686302, KÖ, and by an award from the Simons Foundation 321183FY19, KÖ.

Appendix A

Selection of Model Chemistry

We also tested the model chemistry impact on the methods performances comparing the M06-2X results to similarly obtained BE values using an ab initio, namely MP2 (Frisch et al. 1990), model in combination with the same double zeta basis set (Figure 8). We also compare the M06-2X/aug-cc-
pVDZ and M06-2X/aug-cc-pVTZ (Kendall et al. 1992) performances in evaluating the BE. As only a minimal variation in the BE calculated using the direct method at the M06-2X/aug-cc-pVTZ, M06-2X/aug-cc-pVDZ and MP2/aug-cc-pVDZ is observed (Figure 8), we chose, for clarity, to limit our study to solely the M06-2X/aug-cc-pVDZ model chemistry. We found that the BE estimations depend only slightly on the model chemistry used, but the variation is larger when the cluster size is increased from 1 to 2H₂O regardless of the model used. In a few cases—N₂, CO, HNCO, and SO₂—there are real differences for different model chemistries within the same cluster size, which we speculate are due to a molecular peculiarity that we have not further investigated. In either case, even for these molecules the differences are within the reported uncertainties (≤20%), justifying the use of a single model chemistry in the main section of the paper.

Appendix B

Binding geometries of small hydrocarbons are shown in Figure 9.

Figure 9. Optimized binding geometries (M06-2X/aug-cc-pVDZ) of small hydrocarbons with the 1 H₂O representation of the ASW surface. Bond lengths are in angstroms.
Figure 10 shows the optimized geometries of PO₃ and PO₂OH with the three water molecules cluster resulting in the formation of the PO₄ moiety.

**Figure 10.** Optimized binding geometries (M06-2X/aug-cc-pVDZ) of PO₃ and PO₂OH with the 3H₂O cluster. These local minima show the formation of additional coordination between the P species and one of the waters of the cluster. This does not exclude the existence of physisorption configurations and it does not necessarily imply reactivity. Additional investigation that are beyond the scope of this work are needed to better understand this behavior.

**Appendix C**

**PO₃ Optimization**

The Astrophysical Journal, 939:93 (12pp), 2022 November 10

Appendix C

**PO₃ Optimization**

Figure 10 shows the optimized geometries of PO₃ and PO₂OH with the three water molecules cluster resulting in the formation of the PO₄ moiety.

Ferreró, S., Zamirri, L., Ceccarelli, C., et al. 2020, ApJL, 904, 11

Fontani, F., Rivilla, V. M., Caselli, P., Vasyunin, A., & Palau, A. 2016, ApJL, 822, L28

Fraser, H. J., Collings, M. P., McCoastra, M. R. S., & Williams, D. A. 2001, MNRS, 327, 1165

Frisch, M. J., Head-Gordon, M., & Pople, J. A. 1990, CPL, 166, 275

Frisch, M. J., Trucks, G. W., Schlegel, H. B., et al. 2016, Gaussian 16 Revision C.01 (Wallingford, CT: Gaussian Inc.)

Gálvez, O., Ortega, I. K., Maté, B., et al. 2007, A&A, 472, 691

Gardner, E., Lehto, H. J., Lehto, K., et al. 2020, MNRS, 499, 1870

Germain, A., Tinucci, L., Pantaleone, S., Ceccarelli, C., & Ungiø, P. 2022, ESC, 6, 1286

Hama, T., & Watanabe, N. 2013, ChRv, 113, 883

He, J., Acharya, K., & Vidalí, G. 2016, ApJ, 825, 89

Hollenbach, D., Kaufman, M. J., Bergin, E. A., & Melnick, G. J. 2009, ApJL, 690, 1497

Jones, A. P., Köhler, M., Ysard, N., Boccio, M., & Verstraete, L. 2017, A&A, 602, A46

Jura, M., & York, D. G. 1978, ApJL, 219, 861

Karssemeijer, L. J., & Cuppen, H. M. 2014, A&A, 569, A107

Karssemeijer, L. J., de Wijs, G. A., & Cuppen, H. M. 2014, PCCP, 16, 15630

Kendall, R. A., Dunning, T. H. J., & Harrison, R. J. 1992, JChPh, 96, 6796

Landhess, H., Kosslick, H., Storck, W., & Fricke, R. 1997, Solid State Ionics, 101, 271

Le Guillou, C., Bernard, S., Braye, A. J., & Renusat, L. 2014, GeoCoA, 131, 368

Lebouteiller, V., Kuussivi, & Ferlet, R. 2006, in ASP Conf. Ser., 348, Astrophysics in the Far Ultraviolet: Five Years of Discovery with FUSE, ed. G. Sonneborn, H. W. Moss, & B. G. Andersson (San Francisco, CA: ASP), 480

Lefloch, B., Vastel, C., Viti, S., et al. 2016, MNRS, 462, 3937

Lodders, K. 2003, ApJL, 591, L20

Marboeuf, U., Thiabaud, A., Albert, Y., Cabral, N., & Benz, W. 2014, A&A, 570, A36

Mardirossian, N., & Head-Gordon, M. 2017, MolPh, 115, 2315

Milam, S. N., Hulens, D. T., Tenenbaum, E. D., et al. 2008, ApJL, 684, 618

Molpeceres, G., & Kästner, J. 2021, ApJL, 910, 55

Nguyen, T., Oba, Y., Sameera, W. M. C., Kouchi, A., & Watanabe, N. 2021, ApJL, 918, 73

Noble, J. A., Congiu, E., Dulleiu, F., & Fraser, H. J. 2012a, MNRS, 421, 768

Noble, J. A., Theule, P., Mispelaar, F., et al. 2012b, A&A, 543, A5

Öberg, K. I., & Bergin, E. A. 2021, PRL, 893, 1

Öberg, K. I., Fayolle, E. C., Cuppen, H. M., van Dishoeck, E. F., & Linnartz, H. 2009, A&A, 505, 183

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJL, 743, L16

Olanezwauij, B. O., Herring, J. C., Grieves, G. A., Aleksandrov, A., & Orlando, T. M. 2011, JPCA, 115, 5936

Pasek, M. A., & Lauretta, D. S. 2004, in 35th Lunar and Planetary Science Conference, ed. S. Mackwell & E. Stansbery (Houton, TX: LPI), 1703

Pentead, E. M., Walsh, C., & Cuppen, H. M. 2017, ApJL, 844, 71

Rivilla, V. M., Drozdovskaya, M. N., Altwegg, K., et al. 2020, MNRS, 492, 1180

Rivilla, V. M., Fontani, F., Beltrán, M., et al. 2018, IAU5, 332, 409

Rubin, M., Altwegg, K., Balsiger, H., et al. 2019, MNRS, 489, 594

Schable, M. J., & Baragiola, R. A. 2014, IJREG, 119, 2017

Shimonishi, T., Nakatani, K., Furuya, K., & Hama, T. 2018, ApJL, 855, 27

Smith, R. S., May, R. A., & Kay, B. D. 2016, JPCB, 120, 1979

Soukhanov, T., Thower, J. D., & Zacharias, H. 2017, MNRS, 472, 389

Tenenbaum, E. D., Wolff, N. J., & Ziurys, L. M. 2007, ApJL, 666, L29

Turner, A. M., Abplanalp, M. J., Chen, S. Y., et al. 2015, PCCP, 17, 27281

Viana, R. B., & da Silva, A. B. 2015, Comput. Theor. Chem., 1059, 35

Wakelam, V., Loison, J.-C., Mereau, R., & Ruaud, M. 2017, MolAs, 6, 22

Ziurys, L. M., Milam, S. N., Apponi, A. J., & Woolf, N. J. 2007, Natur, 447, 1094