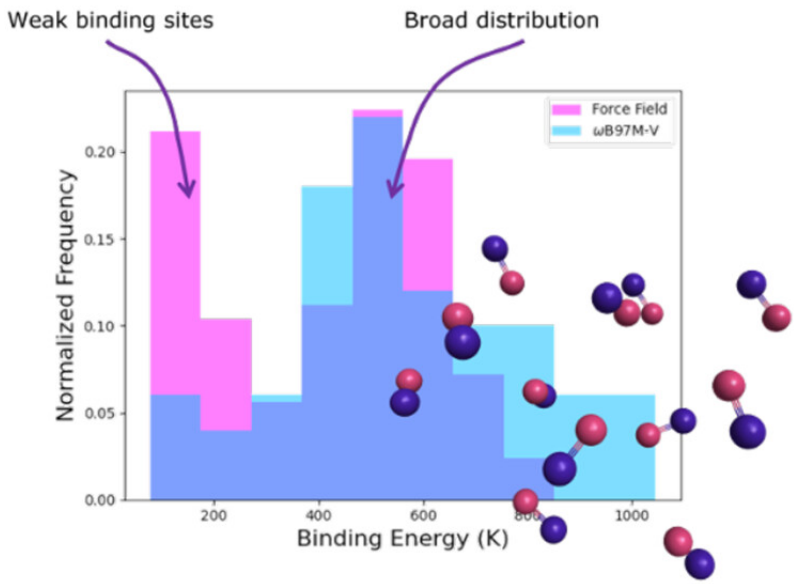
From adsorptation to dissipation: insights from computer simulations of solid H2O and CO

Ferrari, B.C.

Chapter 5

Floating in Space: How to Treat the Weak Interaction between CO Molecules in Interstellar Ices



This chapter is based on:

Ferrari, B. C., Molpeceres, G., Kästner, J., Aikawa, Y., van Hemert, M., Meyer, J., & Lamberts, T. (2023). Floating in Space: How to Treat the Weak Interaction between CO Molecules in Interstellar Ices. ACS Earth and Space Chemistry, 7(7), 1423-1432.

5.1 Abstract

In the interstellar medium, six molecules have been conclusively detected in the solid state in interstellar ices, and a few dozen have been hypothesized and modeled to be present in the solid state as well. The icy mantles covering micron-sized dust grains are, in fact, thought to be at the core of complex molecule formation as a consequence of the local high density of molecules that are simultaneously adsorbed. From a structural perspective, the icy mantle is considered to be layered, with an amorphous water-rich inner layer surrounding the dust grain, covered by an amorphous CO-rich outer layer. Moreover, recent studies have suggested that the CO-rich layer might be crystalline and possibly even segregated as a single crystal atop the ice mantle. If so, there are far-reaching consequences for the formation of more complex organic molecules, such as methanol and sugars, that use CO as a backbone. Validation of these claims requires further investigation, in particular on acquiring atomistic insight into surface processes, such as adsorption, diffusion, and reactivity on CO ices.
Here, we present the first detailed computational study towards treating the weak interaction of (pure) CO ices. We provide a benchmark of the performance of various density functional theory methods in treating the binding of pure CO ices. Furthermore, we perform an atomistic and in-depth study of the binding energy of CO on amorphous and crystalline CO ices using a pair-potential-based force field. We find that CO adsorption is represented by a large distribution of binding energies (200 – 1600 K) on amorphous CO, including a significant amount of weak binding sites ($<$ 350 K). Both increasing the cluster size and the number of neighbors increases the mean of the observed binding energy distribution. Finally, we find that CO binding energies are dominated by dispersion, and, as such, exchange-correlation functionals need to include a treatment of dispersion to accurately simulate surface processes on CO ices. In particular, we find the ωB97M-V functional to be a strong candidate for such simulations.

5.2 Introduction

Carbon monoxide (CO) is ubiquitous throughout the interstellar medium (ISM), and within molecular clouds it is one of the most abundant molecules. It was first detected in the gas phase in 1970[1] and in the solid phase in 1979[2]. Solid CO ice is typically identified by its features at 4.67 and 4.681 μm, that have been shown to correspond to CO embedded in an apolar and polar environment, respectively[3]. This aligns with the

general idea that interstellar ices are layered, with first a polar, mixed, but water-rich inner layer then an apolar, mixed, but CO-rich outer layer[4,5]. This is supported by the novel James Webb Space Telescope (JWST) observation of NIR38, which found that fitted CO ice profiles are dominated by a pure component with two weaker mixed components.[6]

Recent studies claim that the CO-rich ice is formed as a single crystal atop an amorphous water-rich ice[7]. Kouchi *et al.* showed that amorphous CO (am-CO) deposited on amorphous solid water (ASW) was highly susceptible to Ostwald ripening during crystallisation, resulting in only a few crystal islands growing with others feeding into the larger ones[7,8]. They also showed that both UV and electron irradiation of crystalline α-CO do not destroy the crystal structure. Complimentary to this work, He *et al.* have found that pure CO ice under interstellar conditions and timescales is likely to be crystalline[9].

During deposition, Kouchi *et al.*[7] found that CO only partially wets the ASW surface whereas Noble *et al.*[10] found complete wetting. In the latter study, wetting was determined via temperature programmed desorption (TPD), whereas, the former directly imaged the surface during deposition with transmission electron microscopy (TEM). In TEM imaging, the thickness of the sample is determined by the darkness of the image. Identifying the growth of a single monolayer through this method is quite challenging. To reconcile the differences in observed wetting, Kouchi *et al.*[7] evaluated inequalities based on binding energies to suggest that CO first completely wets the ASW surface, then partially wets the CO surface coating the ASW. One crucial assumption was that the binding energy for CO on α-CO is identical to the binding energy for CO on am-CO. It is important to emphasise that this particular assumption (*vide infra*) has not been verified by any experiments so far.

While the majority of studies on the binding energy of CO are experimental, they yield a wide range of values, due to the different empirically motivated pre-exponential factors used to fit the data from TPD curves[11–14]. TPD studies have difficulties distinguishing the different binding on amorphous versus crystalline CO, because of (a) the narrow desorption temperature range and (b) the dependence on the experimental heating rate. Temperature interval desorption (TID) experiments can circumvent the dependence on the heating rate[14], but are susceptible to deviations arising from restructuring. Thus, computational studies with an appropriate level of accuracy are pivotal to understand the binding energy of pure CO at an atomistic level.

Computational studies on CO binding energies have been sparse, because finding an appropriate balance between accuracy and computational cost is very challenging for

this weakly bound system.[15,16] High(er) levels of (*ab initio*) theory cannot describe systems large enough to mimic amorphous systems. Thus, to allow larger system sizes, other techniques, such as; force field based classical molecular dynamics[17,18], continuous time random-walk Monte Carlo simulations[19], and kinetic Monte Carlo simulations[20,21] have been applied in recent years.

Here, we report a benchmark of a previously parametrized[17] classical force field (FF) against coupled cluster theory and a variety of exchange-correlation functionals rooted in density functional theory (DFT) with the aim of determining a reliable and cost-effective way to treat surface processes on CO ices. To this end, we compare binding energies for CO on am-CO which is modeled by small CO clusters. We further report FF-based results for the effect of cluster size on the binding energy distributions, as well as, binding energies of CO on α-CO. Our results are placed in the astrochemical context, and aim to lay the groundwork to provide recommendations for future selections of techniques and cluster sizes for studies of adsorption or dynamics on interstellar CO ices.

5.3 Computational Details

5.3.1 Force Field

The force field uses a site-site pair potential that was parameterized by van Hemert *et al.*[17] based on CCSD(T) calculations for the CO dimer using the aug-cc-pVQZ basis set in combination with the Boys-Bernardi counterpoise correction. An explanation of the potential is given in Chapter 2. The force field has been reimplemented in Python in the form of an Atomic Simulation Environment (ASE) calculator[22,23]. Our implementation of the force field used for the calculations in this work does not employ any cutoffs for energy and force evaluations. The values for all parameters are summarized in Table 2.1. We estimate the (intermolecular) many-body effects beyond pair interactions to amount to less than 2% of the total interaction energy (see Section 1 of S.I. for details).

Simulation Procedures

We generated the amorphous clusters through an in-house "hit-and-stick" script. A randomly oriented CO molecule was spawned 8 Å away from the surface of the CO cluster in a random direction with $\sim$ 50 K of translational energy in the direction of the center of mass of the CO cluster, followed by a 5 ps microcanonical (NVE)

ensemble simulation using the ASE (version 3.22.1) implementation of the velocity Verlet algorithm[24,25]. Subsequently, a geometry optimization was run with the ASE implementation of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimizer[26] using a convergence criterion of 1×10^{-6} eV/Å as a maximum force per atom. This procedure was repeated until the cluster reached a specified number of molecules which was passed to it as an input. Using this technique, 220 clusters were generated with sizes ranging from 8 to 350 CO molecules.

Binding sites for amorphous clusters were sampled by first constructing an α-shape[27], the "shape" of a set of finite points in space, around the atoms.[1] The α parameter was optimized such that the α-shape would fully enclose all atoms while minimizing the volume enclosed. The vertices of the α-shape were then uniformly sampled and a new randomly orientated CO molecule was placed $3-5$ Å away from the norm of the sampled vertex, with the restriction that no atom in the new molecule was closer than 3 Å from any cluster atom. This ensured binding sites selected were equally spaced both from each other and the nearest CO molecule. The cluster and admolecule complex was then optimized to a maximum force per atom of 1×10^{-3} eV/Å. The convergence criteria here and for the "hit-and-stick" procedure were carefully selected based on a study of their effects on the binding energy distributions (see Section 2 of S.I. for details). After geometry optimization, the number of nearest neighbors was determined by counting all molecules with at least one atom within 3.6 Å distance of either C or O of the adsorbing CO molecule.

An α-CO crystal was created by using symmetry operators from the $P2_13$ space group with lattice parameters $a = b = c = 5.9638$ Å, $\alpha = \beta = \gamma = 90°$. We first cut out a cubic $(\alpha\text{-CO})_{864}$ finite cluster from this crystal, which was terminated by {100} faces, and optimized it with a tight (1×10^{-6} eV/Å) convergence criterion. To avoid edge effects except for where adsorption of additional CO molecules is being studied (i.e., at the "sides and at the bottom"), a $(\alpha\text{-CO})_{256}$ cluster is cut out from the center of one of the (100) faces of the $(\alpha\text{-CO})_{864}$ cluster. By freezing out the subsurface molecules they retain their crystalline character, and only the surface edges lose their crystalline character.

CO molecules were placed initially at distances of 3 Å above the crystal face. A total of 500 randomly sampled sites with a randomly orientated CO molecule were used to calculate the binding energies. Only the top layer and admolecule were free to move during optimization, the complex (CO + $\alpha\text{-CO}_{256}$) was optimized with the

[1] An α-shape is a set of lines that encloses a finite set of points. In this particular case, the α-shape is the surface of the amorphous clusters, with the atoms on the cluster surface being the vertices.

BFGS optimizer and a moderate (1×10^{-3} eV/Å) convergence criterion. A buffer region between the plane and crystal edge was left to ensure edge-sites would not be included in the binding energy distributions. The above procedure aims for a realistic simulation of adsorption on a α-CO crystal considering periodic boundary conditions were not applied.

Binding energies for both am-CO and α-CO are calculated via

$$E_{\mathrm{BE}} = (E_{\mathrm{clu}} + E_{\mathrm{mol}}) - E_{\mathrm{clu\,+\,mol}} \tag{5.1}$$

where E_{clu} is the energy of the optimized cluster, E_{mol} is the energy of a single optimized CO molecule, and $E_{\mathrm{clu\,+\,mol}}$ is the energy of the optimized complex (cluster with adsorbed CO). Vibrational zero-point energies (ZPE) were calculated within the harmonic approximation for am-CO, and their contribution to the binding energy is given via $\Delta\mathrm{ZPE} = (\mathrm{ZPE}_{\mathrm{clu}} + \mathrm{ZPE}_{\mathrm{mol}}) - \mathrm{ZPE}_{\mathrm{clu\,+\,mol}}$. Vibrational frequency calculations were carried out with the ASE (version 3.22.1) vibrations package, using finite displacements of $\pm 1 \times 10^{-4}$ Å. All systems studied in this work yielded a total of 4 or less imaginary modes, which are not related to the adsorbate motion. The sum of their absolute values corresponds to a ZPE of 0.2 meV or less, which was omitted when calculating the ZPE corrections.

5.3.2 Density Functional Theory

We compare the interaction energy ($\Delta E_{\mathrm{int}} = E_{\mathrm{dimer}} - 2 \cdot E_{\mathrm{CO}}$) of four randomly oriented CO–CO dimers optimized at the CCSD(T)/ma-def2-TZVP level of theory, with the ORCA program (version 5.0.3)[28,29], against a series of exchange and correlation density functionals. Given the importance of the dispersion interaction for weakly interacting systems[30], we treat dispersion in two ways: (A) *via* Grimme's dispersion correction, either the fourth generation D4[31], or D3[32] including three-body contributions, and (B) including a fraction of non-local (NL) correlation energy in the exchange and correlation energy. Under the latter formalism, the total exchange and correlation energy becomes

$$E_{\mathrm{XC}}^{\mathrm{NL}} = E_{\mathrm{X}} + E_{\mathrm{C}} + E_{\mathrm{C\text{-}NL}}, \tag{5.2}$$

where E_{X} and E_{C} denote exchange and correlation energies according to a particular functional, and $E_{\mathrm{C\text{-}NL}}$ accounts for non-local correlation following the formalism of Vydrov and Van Voorhis.[33] All DFT calculations used the ma-def2-TZVP[34,35] basis set, including a minimal set of diffuse functions to capture long-range interactions and

have been carried out with the ORCA program (version 5.0.3).[28,29] Due to a bug in ORCA version 5.0.3 associated with the D4 method, calculations using this method were corrected with ORCA version 5.0.4. All-but-one corrections were smaller than 10 K. To minimize numerical errors, a large integration grid (`defgrid3`) was used throughout the calculations. This is most important for meta-GGA functionals prone to bigger integration errors, however, to ensure consistency between calculations all functionals are integrated with the same numerical grid. For all functionals, except ωB97M-V, the calculations made use of the RIJK technique, e.g, resolution of the identity for the Coulomb and Exchange integrals. For the ωB97M-V functional, a *chain of spheres* evaluation of exchange (RIJCOSX in ORCA) was employed.

Furthermore, we compare DFT based binding energy distributions with FF based distributions. For the DFT based distributions, CO clusters with sizes of 8, 10, and 12 molecules have been generated as described previously in Molpeceres and Kästner.[36] This procedure differs from the "hit-and-stick" procedure mentioned in Section 5.3.1, because such an individual sticking procedure carried out at the DFT level would be too computationally expensive for the cluster sizes under consideration. Instead, initial structures were randomly generated using Packmol[37]. Subsequently, the structures are pre-conditioned using the generic GFN-FF method within the extended tight-binding (xTB) theoretical framework in the following way.[38,39] A long molecular dynamics (MD) simulation is run at 100 K for 100 ps for each structural model, to generate different starting configurations, applying a spherical wall potential to confine the CO molecules within the cluster and prevent evaporation of the ice. From this, we extracted an MD snapshot every 20 ps, which was quenched to 10 K for an additional 10 ps. These five initial structures per cluster size (15 in total) were subsequently optimized at the ωB97M-V/ma-def2-TZVP level of theory under strict convergence criteria: Max(E_{diff})= 1.0×10^{-7} E_h, Max(Step)= 3.0×10^{-4} Bohr, RMS(Step)= 2.0×10^{-4} Bohr, Max(F)= 7.5×10^{-5} E_h/Bohr, and RMS(F)= 5.0×10^{-5} E_h/Bohr.

Binding energy calculations were performed by placing additional CO molecules around the cluster, spanning a distorted Fibonacci lattice, with the centre of mass of each additional CO molecule at a (minimum) distance of 3 Å plus the maximum value of the Cartesian (x, y, and z) coordinate of any of the CO molecules in the cluster. More details for the sampling procedure can be found in[36]. To avoid surface restructuring, these initial configurations are relaxed with slightly less strict convergence criteria; Max(E_{diff})= 2.2×10^{-7} E_h, Max(Step)= 4.0×10^{-4} Bohr, RMS(Step)= 2.6×10^{-4} Bohr, Max(F)= 1.0×10^{-4} E_h/Bohr, and RMS(F)= 6.6×10^{-5} E_h/Bohr. For the $(CO)_{12}$ cluster, we fixed all the molecules present at a distance higher than 6.5 Å from

the admolecule's centre-of-mass, to ease the convergence of the geometry optimization procedure. Binding energies are calculated using Eqn. 5.1. We note that the aforementioned geometry optimization thresholds are less strict than those used for the FF calculations, owing to the fact that the energy and force calculations at the DFT level increase the computational time by about 5-6 orders of magnitude. Nevertheless, according to our detailed analysis in the S.I. (see Section 2), we do not expect this to have a significant effect on the calculated binding energies.

5.4 Results and Discussion

5.4.1 DFT Benchmark

The following functionals were tested: PBE-D3[40], B3LYP (without dispersion)[41], B3LYP-D4[31,41], B3LYP-NL[33,41], REVPBE0-D4[31,42], REVPBE0-NL[33,42], BHLYP-D4[31,43], M06-2X-D3[32,44], PW6B95-D4[31,45], PW6B95-NL[33,45] and ωB97M-V[46], which has the NL treatment in its original implementation. For each of these functionals, the dimer structure was re-optimized starting from the CCSD(T)/ma-def2-TZVP optimized geometries. The structures of the four dimer configurations are depicted in Fig. 5.1 (see Section 5 of S.I. for coordinates) and the calculated interaction energies are summarized in Table 5.1.

Overall, the degree of agreement between all DFT methods including dispersion and CCSD(T) is satisfactory, confirming previous calculations.[18,47,48] We observe that B3LYP without dispersion correction, yields results that are far from the reference value, while both D4 and NL corrections bring the values much closer to the reference. In general, NL-corrected functionals perform only slightly better compared to those with a D4 correction. Furthermore, PBE-D3, a commonly used method in computational solid state chemistry, clearly does not capture the interaction well, while the ωB97M-V functional outperforms all others. We further scrutinized the performance of the ωB97M-V functional by expressing the dimer potential explicitly in terms of radial and angular variables, similar to the method for the force field construction (see S.I. for details). We found good agreement between the ωB97M-V/ma-def2-TZVP potential and the CCSD(T)/aug-cc-pVQZ potential. Finally, note that a direct comparison to the energetics for these dimers predicted by the force field is not straightforward as a result of the use of different basis sets for the CCSD(T) calculations and the FF parametrisation.

The ωB97M-V/ma-def2-TZVP level of theory will be used for all future DFT

calculations discussed here and the applicability of this method to general adsorbates and reactions on ice clusters will be the subject of future work.

Table 5.1: Dimer interaction energy, E_{int}, for the reference method CCSD(T)/ma-def2-TZVP and the difference $\Delta E_{\text{int}} = E_{\text{int,CCSD(T)}} - E_{\text{int,DFT}}$ for different exchange and correlation functionals in Kelvin.

		Dimer 1	Dimer 2	Dimer 3	Dimer 4
CCSD(T)	E_{int}	*-158.4*	*-163.4*	*-149.3*	*-112.0*
PBE-D3		68.2	63.0	77.2	114.5
B3LYP		-153.2	-152.0	-143.1	-122.2
B3LYP-D4*		9.9	-18.4	12.2	3.6
B3LYP-NL		-15.2	-28.7	-0.2	-24.3
BHLYP-D4*		46.6	22.5	47.8	22.9
M06-2X-D3	ΔE_{int}	-19.2	-9.4	-18.0	-5.4
PW6B95-D4*		69.8	69.7	69.1	88.7
PW6B95-NL		59.0	67.1	55.3	76.3
REVPBE0-D4*		48.3	26.3	52.8	58.2
REVPBE0-NL		24.3	21.6	32.8	36.1
ωB97M-V		**0.7**	**1.6**	**-0.3**	**4.8**

5.4.2 Amorphous CO

Comparison between FF and DFT

We compared DFT and FF-based binding energy distributions for cluster sizes of 8, 10 and 12 CO molecules. Thanks to the computational efficiency of the FF method we have used larger sample sizes of 200, 250, and 300 respectively, in comparison to 60, 50, and 40 for the DFT calculations. The resulting distributions are depicted in Figure 5.2. Only for small clusters (16 or less CO molecules) the distributions were bi-modal, as such, only those distributions were fit with two Gaussian profiles. In both Figure 5.2 and Table 5.2, for the bi-modal cases only the descriptors for the higher BE distribution is shown. We find very good agreement between the DFT and FF binding energy distributions. The correspondence between FF and DFT results are underpinned by the descriptors of the distributions for all amorphous clusters studied, presented in Table 5.2. The standard deviations (σ, governing the width of the distribution) are in agreement even in cases where the means (μ, governing the average binding energy) differ between FF and DFT. However, the median (η) and the Median Absolute Deviation (MAD) on the other hand are more robust descriptors of the distributions. The median of the distribution is less sensitive to outliers and, in

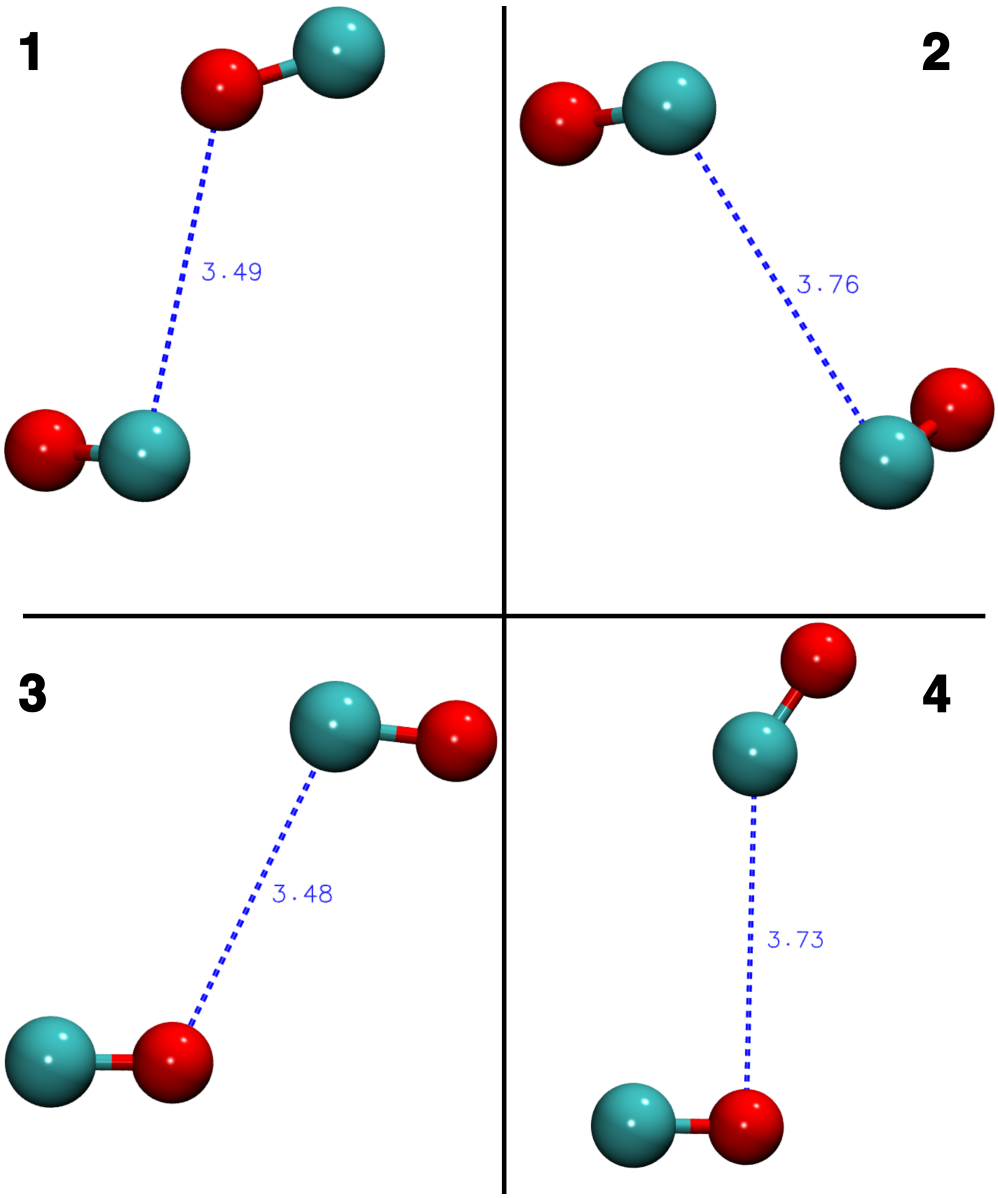


Figure 5.1: CO dimer configurations used to determine the interaction energies utilized in our benchmark study (Table 5.1). Full geometry coordinates are given in the SI.

line with earlier work, we recommend using the median to estimate binding energies in these weakly bound systems.[18]

Both methods find many weakly bound configurations, however, more so for the FF method. In particular, there is a significant amount of extremely weak (< 200 K) binding sites. Subsequent NVE MD simulations at the FF level (with time steps of 0.5 fs) revealed these binding sites are transient. Residual forces in those weakly bound configuration were sufficiently large so that the adsorbed CO diffuses to a stronger binding site during short timescales (∼ 1 ps). The relative abundance of these sites for FF simulations diminishes with increasing cluster size, see Figure 5.2d, as such, studies with small cluster sizes should scrutinize the relative abundance of these sites. For cluster sizes larger than 100 CO molecules these sites make 1% or less of the sample size, and for 350 CO molecules it is down to 1 transient site out of 550 samples. Since these sites disappear for larger cluster sizes, it implies they arise from adsorption at defect sites, where there is a reduced number of interactions.

Table 5.2: Mean (μ), standard deviation (σ), median (η), and median absolute deviation (MAD), smallest binding energy (E_{BE}^{min}) and largest binding energy (E_{BE}^{max}) of the CO-COn. All values of the table are in Kelvin.

Cluster Size	Method	μ		σ		η		MAD		E_{BE}^{min}		E_{BE}^{max}	
		a	b	a	b	a	b	a	b	a	b	a	b
8	FF	498	356	101	79	498	359	101	69	142	90	753	551
8	DFT	543		120		547		103		151		851	
10	FF	542	391	112	86	531	381	116	90	147	95	839	613
10	DFT	604		195		560		174		77		1044	
12	FF	562	409	130	105	565	397	115	92	147	95	1192	915
12	DFT	525		159		528		174		126		1049	
16	FF	611	451	148	125	611	459	152	129	153	97	954	753
24	FF	600	454	189	152	616	467	164	135	154	103	1088	795
32	FF	617	470	187	152	630	474	160	139	153	106	1161	917
48	FF	665	508	183	150	670	512	148	120	151	98	1280	1050
75	FF	676	523	175	145	689	526	137	112	157	106	1115	894
96	FF	717	550	183	148	730	552	144	119	172	129	1311	1030
150	FF	742	564	177	147	771	593	171	146	174	113	1205	899
200	FF	835	633	227	180	837	640	178	126	169	109	1610	1301
250	FF	755	598	181	154	785	624	161	139	167	107	1302	1088
350	FF	786	625	184	154	802	640	158	140	199	141	1506	1304

[a] Without ZPE corrections
[b] With ZPE corrections

Dispersion Contribution

Figure 5.3 shows the contribution of each intermolecular interaction on the binding energy of CO on amorphous CO clusters. The dispersion (V_{disp}) contribution dominates the binding energy, as expected, further underpinning the importance of the inclusion

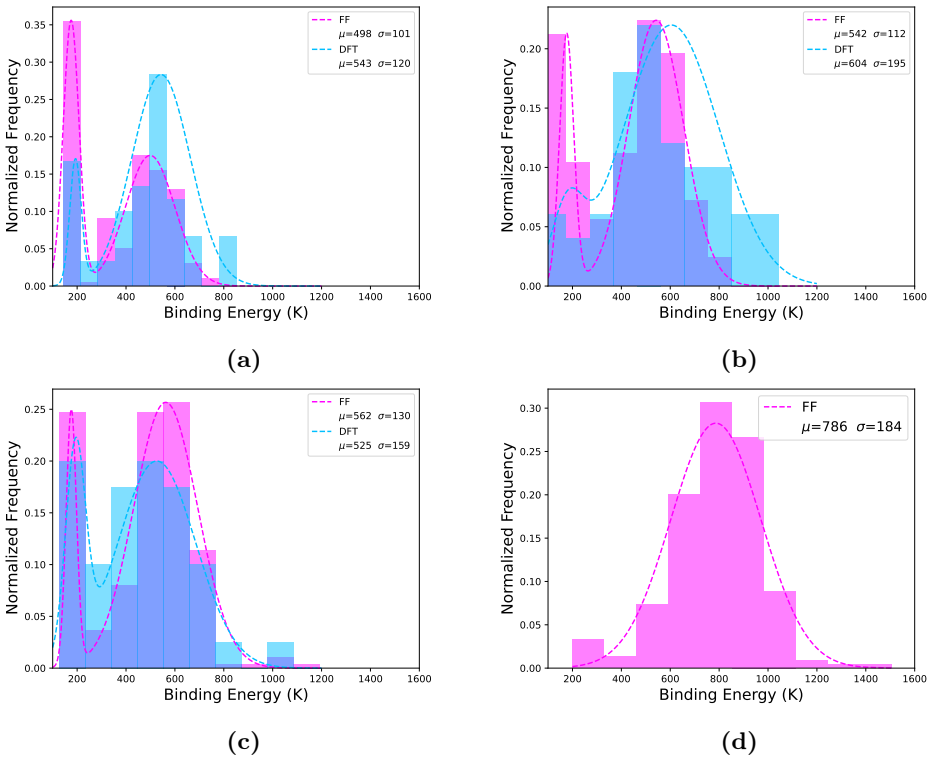


Figure 5.2: Distributions of binding energies calculated by DFT (blue) and FF (pink) methods for cluster sizes of (a) 8, (b) 10, (c) 12, and (d) 350 CO molecules. The mean (μ) and standard deviation (σ), both in K, for each distribution is shown in the plot legends. Dotted lines are probability density functions fitted to the binding energy distributions. For all distributions 10 equally spaced bins were used for plotting. Distributions consist of (a) 60 DFT samples and 200 FF samples (b) 50 DFT samples and 250 FF samples (c) 40 DFT samples and 300 FF samples (d) 550 FF samples. Note, overlapping blue and pink bars result in a third color within the plots.

of dispersion in exchange-correlation functionals used to simulate surface processes on CO ices. We also include the Buckingham potential ($V_{\text{exch}} + V_{\text{disp}}$) contribution since it illustrates the interplay between the exchange and dispersion contributions. Interestingly, the dispersion and exchange contributions vary significantly across binding arrangements, whereas, their combined interaction (Buckingham) seem to balance each other out to produce a contribution with less variation. Contrary to water clusters, here the electrostatic (V_{el}) contribution is the weakest (by absolute magnitude) contribution to the binding energy, with ~ 100 K of the binding energy resulting from electrostatic interactions.

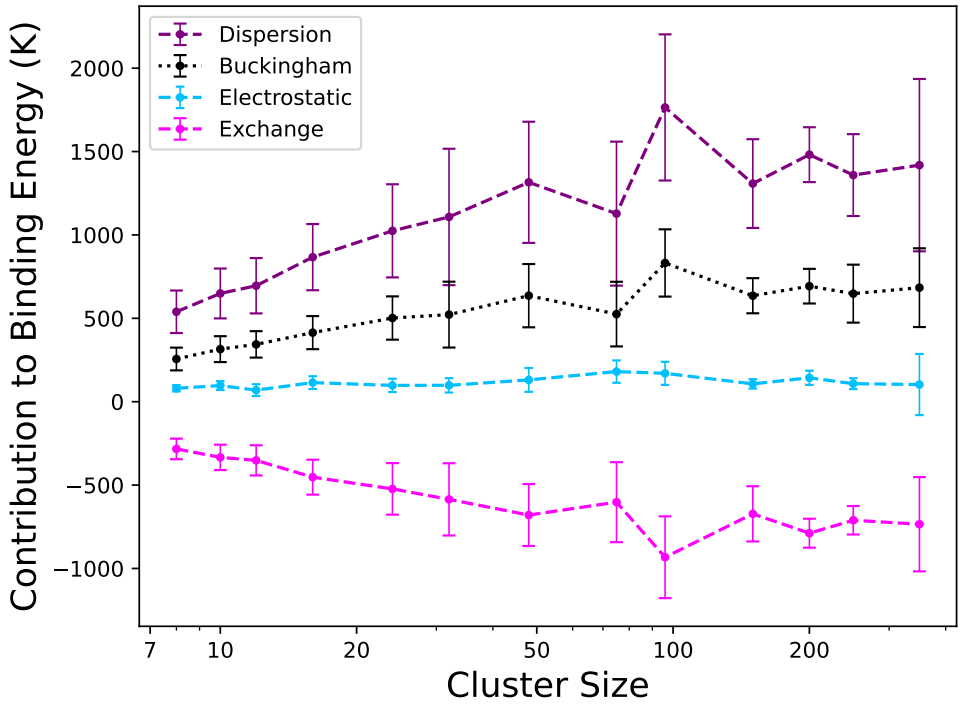


Figure 5.3: Average contribution to the binding energy of each part of the force field employed. Points are the mean of the contribution distributions, error-bars are the standard deviation and dashed/dotted lines are a guide for the eye.

$(CO)_x$ clusters have a key peculiarity in comparison with $(H_2O)_x$ ones, which has been the most commonly used substance to simulate interstellar ices. In water, the dominant interaction is brought about by the hydrogen bonds between molecules, which is directional and strong. On the contrary, in solid-state CO ices most of

the interaction energy stems from the dispersion interactions, as such, CO molecules tend to orient with less directionality than H_2O. As a consequence, while dual-level calculations with electronic energy refinement on a low-level geometry are a cheap and accurate way to describe reaction energetics, the physisorption of CO admolecules on CO clusters is not well captured.

Binding Energy Size Trend

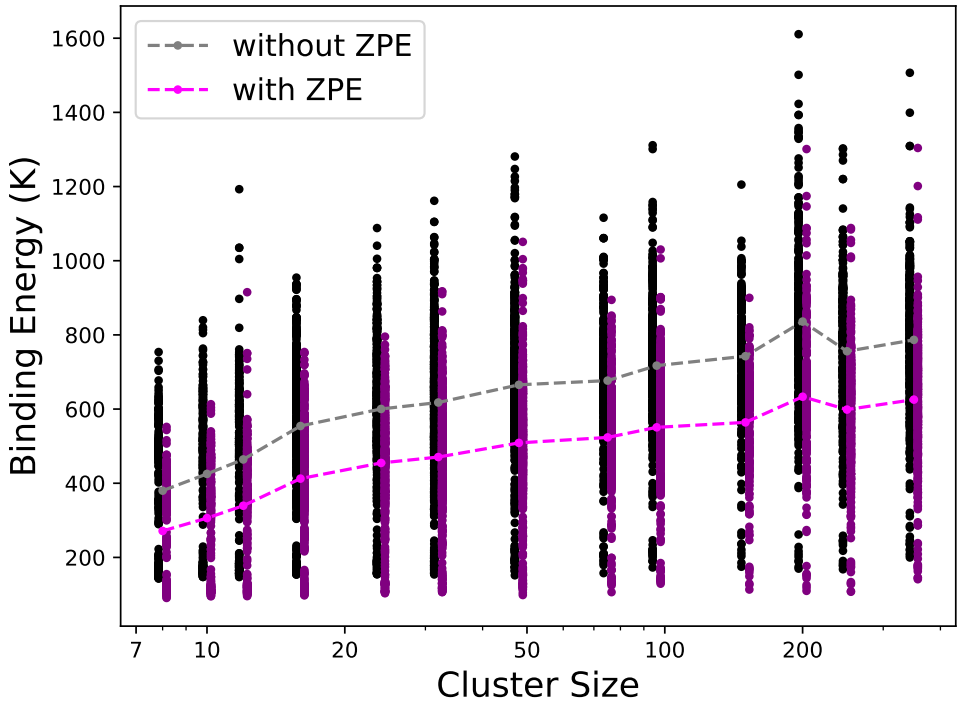


Figure 5.4: Black and purple dots are binding energy distributions, the grey and fuchsia dots are the median values of the distributions. Dashed lines do not indicate predicted trends.

It is important to study the cluster size dependence because the weak dispersion results in short-range interactions that do not dominate over the long-range ones. We find that the median binding energy for a cluster size of 350 molecules (802 K) is nearly twice the median binding energy for a cluster size of 8 molecules (426 K). Figure 5.4 shows the effect of cluster size on the binding energy distribution, where both ZPE corrected and uncorrected binding energies are shown. The binding energies are shown

as black and purple scatter points and demonstrate a wide range of binding sites on am-CO. Larger cluster sizes increase the overall binding energy, with the median value seeming to asymptotically approach the previously reported experimental values by Acharyya *et al.*[12] (858 ± 15 K) and Bisschop *et al.*[13] (855 ± 25 K).

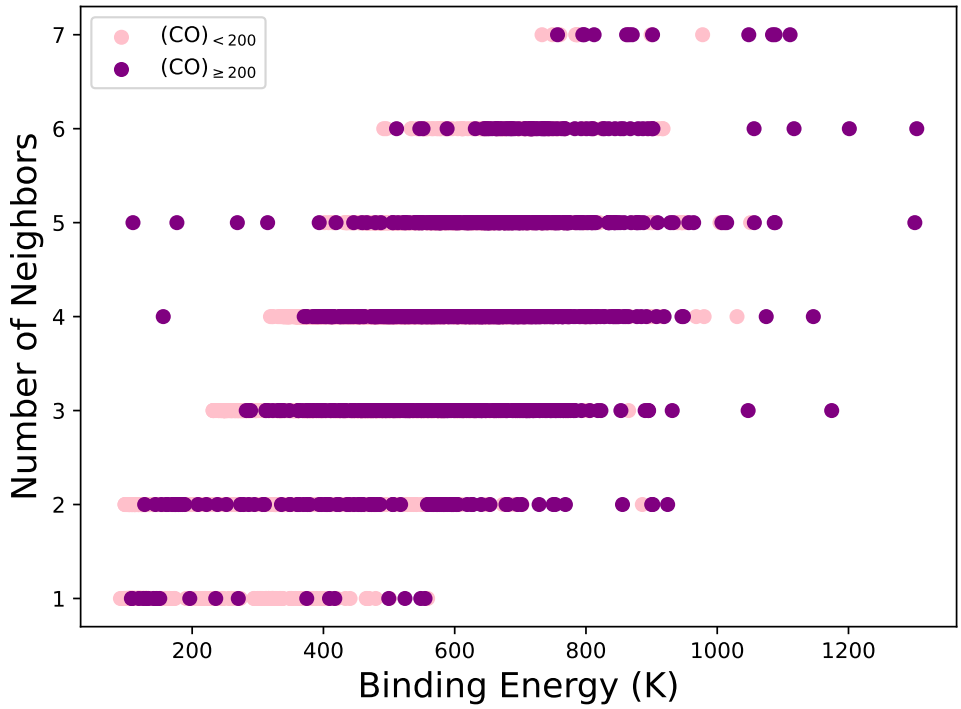


Figure 5.5: All FF calculated binding energies with ZPE corrections plotted as a function of number of nearest neighbors. Results for cluster sizes less than 200 molecules are pink and all larger clusters are purple.

In general, the largest binding energy at each cluster size increases with increasing size until 50 CO molecules. The $(CO)_{12}$ cluster is an exception due to one of the $(CO)_{13}$ complexes (CO + $(CO)_{12}$) having a uniquely symmetric orientation. This allowed the adsorbing CO to maximize its number of nearest neighbors, producing a binding energy significantly higher than on most other clusters. Larger clusters have maximum binding energies that depend less on the cluster size and depend more on the surface morphology, and the number of nearest neighbors. Figure 5.5 shows the relationship between binding energy and number of nearest neighbors. Increasing the number of neighbors shifts the distribution of binding energies to larger values, because

short range interactions contribute more to the binding energy.

The ZPE correction lowers the median binding energy by 20 – 30% and reduces the MAD by 15 – 20%. This reduction is greater than or equal to ZPE contributions found from previously reported studies of other molecules on water ice surfaces [48,49]. For a weakly interacting system, such as CO on CO, the ZPE corrections can influence binding significantly and should be carefully considered. The ZPE corrections are calculated within the harmonic approximation. This may result in an overestimation, because the weak CO−CO interactions on a relatively flat PES which are expected to be dominated by anharmonic character. We find that the low frequency modes dominate the ΔZPE with the largest contributions resulting from the frozen-out rotations and translation of the "newly" adsorbed CO molecule. Since these modes are more anharmonic in nature we expect anharmonic corrections to the ZPE to be significant. As such, the ZPE corrections reported here are an upper limit on the exact ZPE correction.

5.4.3 Crystalline CO

The binding energy distribution for CO on the (100) surface of α-CO is shown in Figure 5.6. The weakest binding sites ($\sim$ 500 K shown in red in Figure 5.6) correspond to sites directly on top of a CO molecule on the topmost CO layer, and the strongest binding sites ($\sim$ 800 K shown in pink in Figure 5.6) correspond to CO alignment with a subsurface α-CO molecule. Similar to the amorphous case, the weak binding sites ($<$ 650 K) are found to be transient, with CO diffusing into a site above a subsurface α-CO molecule when subsequent NVE simulations are performed. It then binds in orientations similar to those of the stronger binding sites mentioned above. Our calculations only consider the perfectly flat (100) CO crystal surface, and thus do not include the influence of step edges. Ignoring these effects, we find that the median binding energy of CO on α-$(CO)_{256}$ (742 K) is similar to that of CO on am-$(CO)_{250}$ (785 K); validating assumptions made in previous studies. [7]

In order to asses how the orientation of the adsorbing CO affects the binding energy we ran a density based spatial clustering of applications with noise (DBSCAN) [50] algorithm on the orientation of adsorbed molecules. The molecules were assigned a category based on their relative alignment with the second crystal layer, the categories were then used to color the distribution in Figure 5.6 (see S.I. for more details). A representative image for each category corresponding to the strongest binding sites is shown in Figure 5.7. The DBSCAN algorithm also separates "noise", points that are

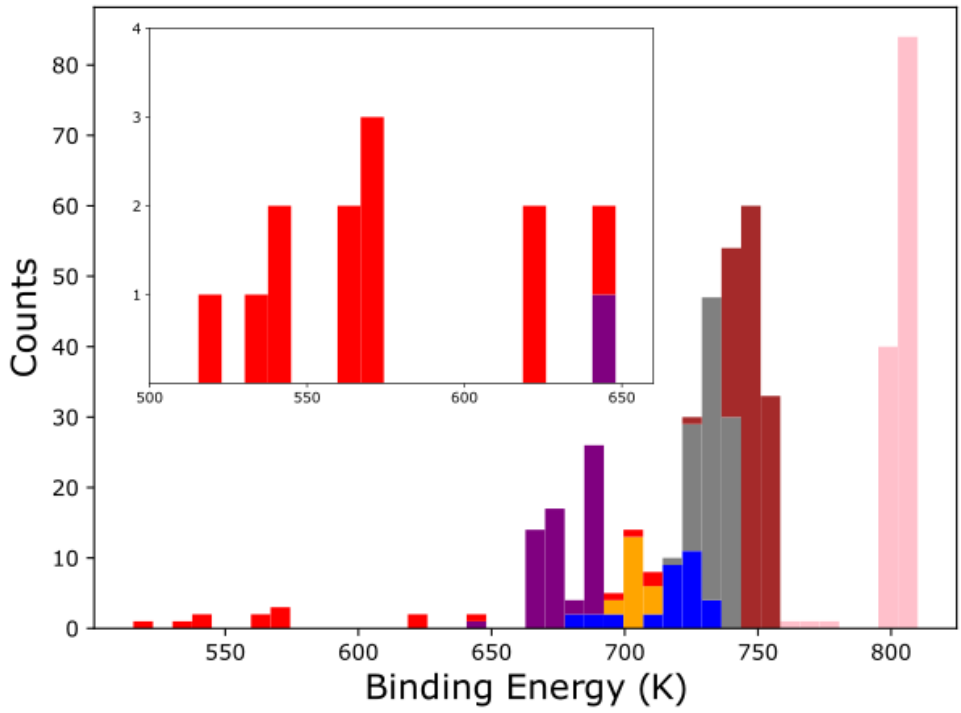


Figure 5.6: Distribution of binding energies calculated for a CO molecule adsorbed on the (100) surface of an α-CO crystal. Colors depict categories (see Figure 5.7) found by a clustering algorithm with reference to the orientation of the adsorbed CO. Inset is zoomed in on the low energy sites.

not similar to other points based on clustering space, represented by the red bars in Figure 5.6. In general, we can conclude that the more parallel the orientation of the adsorbed CO is with the subsurface α-CO below it, the higher the binding energy.

Many binding sites center around 740 K and they differ geometrically from the stronger sites around 800 K by the orientation of the adsorbed CO molecule, see Figure 5.7. Stronger binding occurs when the adsorbed molecule is approximately parallel to the sub-surface CO, with the adsorbed CO being inverted (in terms of C-down or O-down). Although being the strongest binding orientation, this does not correspond to full alignment with the crystal structure. Those orientations have a CO molecule parallel to the sub-surface CO, but non-inverted and the binding energies fall within the distribution centered around 740 K.

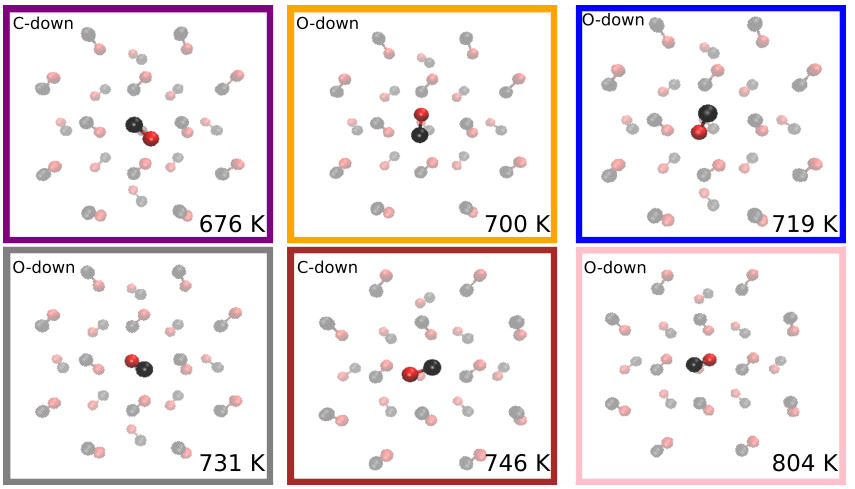


Figure 5.7: A representative image of each category for binding sites on the (100) surface of an α-CO crystal. Colored boxes indicate the category color in the histogram (see Figure 5.6), and text in upper left corner of each box indicates which atom is closest to the crystal face. Text in the lower right corner of each box is the median binding energy of the categories distribution. "Noise" points (red color in Figure 5.6) are not shown here since all orientations in the distribution differ significantly from each other. Note, all molecules shown here are adsorbed on a hollow site (over a CO molecule in the second layer).

5.5 Astrophysical Implications

We find that for large clusters (> 200 CO molecules) roughly 10% of the binding energy distributions falls below 600 K. This supports the idea that CO molecules are mobile even at low temperatures. This could promote the crystallization of the CO ice on interstellar timescales, which would support recent claims of interstellar CO being in crystalline form[7,9]. Whether or not the structural phase of CO ices has an effect on subsequent reactivity remains to be tested. It may lead to less accessible transition states for, e.g., the H + CO reaction, but given its intrinsic weakly interacting nature, it is possible that these effects are only observed in a single crystal without step edges or defects. As such, future studies on how the CO ice phase affects surface reactions would be of great interest to the astrochemical community.

The broad binding energy distributions reported herein are of particular interest to astrochemists modeling CO in pre-stellar cores. Within these regions there is evidence of a significant CO depletion in the gas-phase, between 74% and 94% of the gas-phase CO is observed to be frozen out.[51] However, models are unable to reproduce the observed gas-phase CO abundances and overestimate the CO depletion. For instance,

Keto and Caselli were only able to reproduce the observed value by increasing the cosmic-ray induced desorption rate for CO by a factor of 30 from what is expected.[52,53] Alternatively, Cazaux *et al.* showed that incorporating lower binding energy values for CO in models would decrease the CO depletion.[54] They found that using a binding energy of 350 K would lower the CO depletion by 10% and a value of 300 K would lower it by 100%.[54] Our results show that these binding energy values are within the distribution for CO. Additionally, we found that adsorption to defect sites are transient and in the limit of low CO coverage they will diffuse to stronger binding sites. At higher CO coverage, stronger binding sites will already be occupied and the admolecule will be unable to diffuse to a new site, resulting in a low temperature desorption event. Incorporating these concepts along with their probabilities into models could result in gas-phase CO abundances closer to those observed, without the need to invoke additional processes.

5.6 Conclusion

In summary, we find that CO binding is dominated by dispersion, and that many-body effects contribute minimally ($< 2\%$) to the interactions in CO ices. We also show that CO binding on amorphous CO occurs with a large range of binding energies (200 – 1600 K), and depends on both the cluster size and the number of nearest neighbors. Our results show a median binding energy of 802 K for $(CO)_{350}$, which is nearly twice as large as that for $(CO)_8$ (426 K). Accounting for ZPE within the harmonic approximation lowers the average binding energies by 20–30%. We expect this to be an upper limit because frustrated rotations and translations dominate the ZPE correction and the effect of these low-frequency modes might be overestimated by neglecting anharmonicity. This should be revisited in future studies.

We also report binding energy distributions (650 – 800 K) for CO binding on a flat α-CO crystal face, which we find to be orientation dependent. It also has a median value similar to that of amorphous CO, corroborating previous studies which assumed them to be equal. Lastly, we presented many dispersion corrected functionals that performed well, however, we find the ωB97M-V functional performs the best when treating the CO−CO dimer interaction and suggest this functional to be used for general adsorbates and reactions on ice clusters.

5.7 Data Availability Statement

All of the Python code used for the simulations is publicly available on GitHub at https://github.com/Cavenfish/CO_Project. XYZ files for all clusters used in this work can be found on Zenodo at https://zenodo.org/records/8068393.

5.A Appendix

5.A.1 Pair Interaction Validation

We estimate the (intermolecular) many-body effects beyond pair interactions to amount to less than 2% based on the procedure outlined below. We selected 20 trimer geometries and have performed CCSD(T)/aug-cc-pVQZ calculations using the same computational setup as used for the construction of the force field. Ten geometries were selected from the bulk of an unequilibrated $(CO)_{522}$ 'crystal' cluster, and the other ten were selected from the bulk of a $(CO)_{800}$ amorphous cluster equilibrated to 10 K.[17] For each geometry the total energy was calculated (E_{total}) and the energies of each pair ($E_{dimer,i}$) with the difference (Eqn. 5.A.3) attributed to non-additivity.

$$\Delta E = E_{\text{total}} - \sum_{i} E_{\text{dimer},i} \tag{5.A.3}$$

We find that ΔE can be both positive or negative and varies between 0 and 2% of E_{trimer}, indicating that a pair potential approach to study, e.g., binding of pure CO ices, is justified.

5.A.2 Geometry Optimization Threshold

Using the force field employed in thie work, we investigated the convergence threshold criteria for the creation of interstellar relevant CO clusters and for the subsequent binding energy calculations and concomitant distributions. While global sampling schemes provide a systematic way to obtain meta-stable states, instead, here we use a more pragmatic approach. This is also of more practical relevance to electronic structure calculations given their much higher computational cost. To this end, two $(CO)_{32}$ clusters were generated with the hit-and-stick method discussed in the manuscript. We tested six convergence criteria for the residual forces, namely 1×10^{-x} with $x = 1, 2, 3, 4, 5, 6$ eV/Å, leading to a total of 12 generated clusters. Subsequently, for each cluster ge-

ometry optimizations were performed for 60 binding sites and binding energies were computed for each convergence criterion equal to or looser than the one used for cluster generation, leading to 2520 simulations. The binding energies are calculated in the same method as described in the main article. The results are depicted in Fig. S5.A.8. As to be expected, the effect of the convergence criteria can be rationalized as follows: (a) during cluster generation, tightening the criteria results in a less disordered cluster and (b) during adsorption simulations, tightening the criteria forces the adsorbing molecule to overcome small barriers and progress toward stronger binding sites.

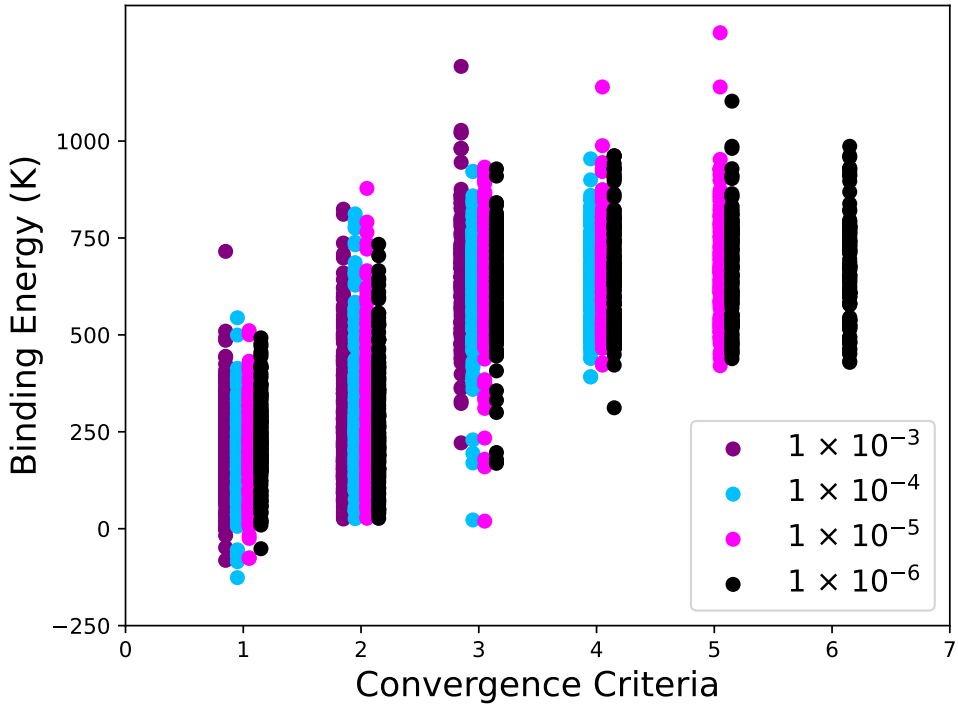


Figure 5.A.8: Binding energy distributions for simulations with varying optimization criteria, note that the points are artificially offset on x-axis for visibility. The x-axis denotes the convergence criteria for the binding energy (adsorption) simulations based on 1×10^{-x} eV/Å. The colors indicate the thresholds used for the cluster generation: (a) purple dots used 1×10^{-3} eV/Å (b) cyan dots used 1×10^{-4} eV/Å (c) magenta dots used 1×10^{-5} eV/Å (d) black dots used 1×10^{-6} eV/Å.

Tightening the criteria for the cluster creation overall leads to narrower binding energy distributions which can be seen by comparing within each set of calculations per criterion on the x-axis in Fig. S5.A.8. For instance, for a threshold of 1×10^{-3}

eV/Å on the adsorption simulations, the width of the black distribution (a threshold of 1×10^{-6} eV/Å for the cluster generation) is the smallest of all four colors. Tightening the criteria for the binding energy, *i.e.*, adsorption simulations, leads to an overall increase in binding energies and drastic reduction in the number of weak binding energies, because shallow minima will progress toward more stable local minima.

Computational chemistry studies usually focus on well-converged structures corresponding to local minima, if not even the global minimum. In this case it would mean to make use of threshold criteria of 1×10^{-6} eV/Å for both cluster formation and adsorption simulations. However, for the study of binding on cryogenic (10 K) interstellar ices, this would represent an unrealistic scenario: adsorbing molecules have near to no additional energy to reorient on the surface. In particular, transient binding sites for which the binding energy is sufficient to lead to adsorption, yet so low that subsequent diffusion leading to chemistry or fast desorption is expected, can not be captured with stringent convergence criteria. In fact, for CO, it has been shown that these weakly bound sites determine the gas-phase depletion rates and influence astronomical models[54]. At the same time, loose convergence criteria leads to negative binding energies representative of a repulsive force. Thus, for our work, keeping the interstellar conditions and questions in mind, we use 1×10^{-6} eV/Å as a criterion for generating the clusters and 1×10^{-3} eV/Å for computing the binding energies at the force field level. As detailed in the manuscript, different criteria are used for the electronic structure calculations in light of the concomitant computational costs.

5.A.3 Total Interaction Energy

In order to further elucidate the virtues and shortcomings of the ωB97M-V functional, we applied a spherical expansion analysis[55,56] previously used by van Hemert *et al.* for the original construction of the force field. We label the intermolecular distance (the distance between centers of mass of the two molecules) by R and use dimer jacobi angles (θ_A, θ_B and ϕ) for the orientation. The potential is then expressed as,

$$V(R, \theta_A, \theta_B, \phi) = \sum_{L_a, L_b, L} c_{L_a, L_b, L}(R)\ A_{L_a, L_b, L}(\theta_A, \theta_B, \phi) \tag{5.A.4}$$

with R-dependent expansion coefficients c (in energy units), dimensionless basis functions A and the angular momentum coupling labels given by L_a,L_b and L.

We focus here on the two best performing exchange-correlation functionals, ωB97M-V and B3LYP-D4 both including dispersion corrections, along with bare B3LYP to un-

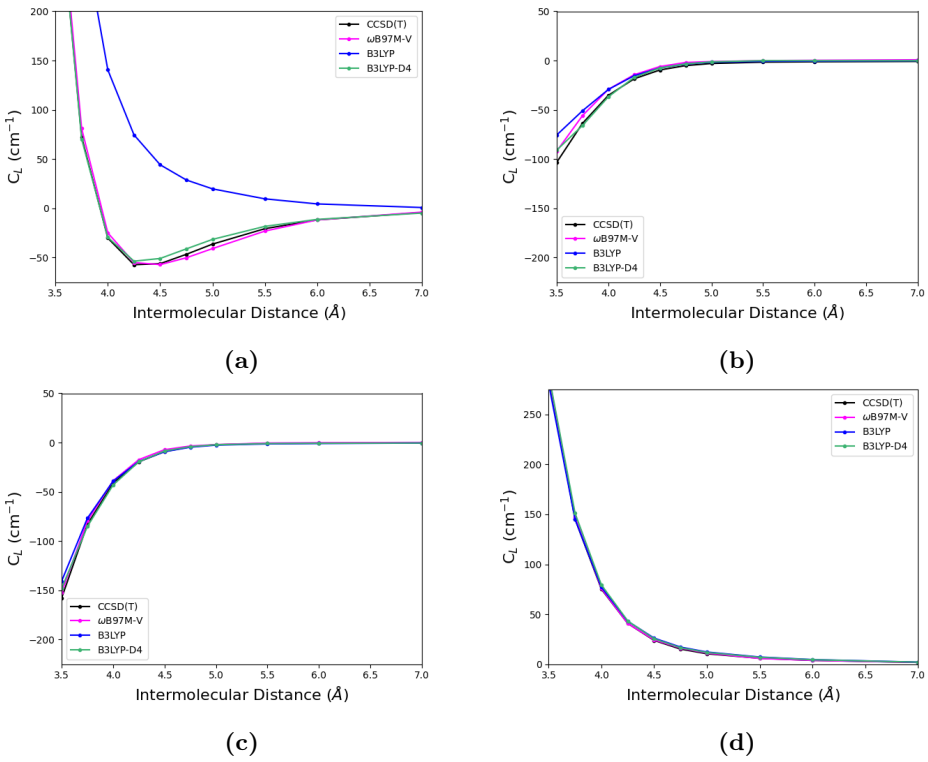


Figure 5.A.9: The intermolecular distance dependence of the dominant coefficients of the 140 term expansion. Shown here are (a) c_{000}, (b) c_{112}, (c) c_{123}, and (d) c_{224}.

derstand better the effects of dispersion. All functionals employed the ma-def2-TZVP basis set, and CCSD(T)/aug-cc-pVQZ acts as the reference. We use the Molpro program for all methods except ωB97M-V, which was run with ORCA. For 21 different intermolecular distances (R), the energy of 294 orientations was calculated. In Figure S5.A.9 we present the intermolecular distance dependence of the dominant coefficients of the truncated expansion including 140 terms.

Figures S5.A.9 (b), (c), and (d) show the 3 dominant anisotropy terms. At long range they determine the orientational dependence of (b) the dipole-dipole interaction, (c) the dipole-quadrupole interaction and (d) the quadrupole-quadrupole interaction. For R values larger than 10 Å, they show the expected R^{-n}, with n = 3, 4, and 5 for case (b), (c) and (d) respectively.

For the isotropic component, Figure S5.A.9 (a), the CCSD(T) reference curve (black) has its minimum at R=4.25 Å with -50 cm^{-1} energy. Both B3LYP-D4 (green)

and ωB97M-V (pink) reproduce the CCSD(T) curve closely, implying they incorporate dispersion in a similar way as CCSD(T). Bare B3LYP (blue) on the other hand demonstrates a lack of dispersion interaction, making it clear that B3LYP will always produce dimer equilibrium geometries with overestimated intermolecular distances.

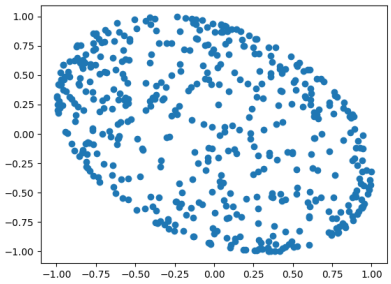
This further underpins that ωB97M-V is well suited to treat CO ices, and may be able to describe other dispersion dominated systems as well. Considering its quality compared to its computational cost, we propose it to be an excellent functional for treating CO-CO interactions both for geometries and energies.

5.A.4 DBSCAN

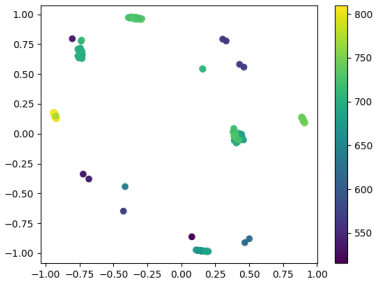
Our goal in clustering the adsorbing geometries is to correlate the binding energy distributions with the orientation of the admolecule. In order to do so, we need to carefully select the features that will be supplied to the DBSCAN algorithm[50] to cluster the data. Within the α-CO crystal structure, there are two orientations of CO molecules per layer along the (100) direction, which defined the surfaces studied in the main article. Incidentally, the two have bond vectors that are nearly orthogonal. The relative alignment between the admolecule and one of these two CO orientations can be determined by taking the scalar product of unit vectors along the orientation of the two molecules. This provides a vector of 2 features that represents the general orientation of the admolecule, each in a range from -1 to 1. Figure 5.A.10 (a) shows these relative orientation features plotted for all sampled geometries prior to geometry optimization, and Figure S5.A.10 (b) shows them after geometry optimization. We then clustered the binding geometries based on these two features, using the DBSCAN algorithm. Figure S5.A.10 (c) shows the post-geometry optimization features plotted with color coded groupings determined by the DBSCAN algorithm.

The DBSCAN algorithm requires the selection of two important parameters, ϵ which is the 'neighborhood' radius and N which is the minimum number of points within a 'neighborhood' needed to make a point a core point. In our study we varied both parameters across $0.05 - 0.15$ for ϵ and $3 - 6$ for N, there was little difference in the resulting clusters. This is due to the tight packing of the data points. As such, we used a small value for the 'neighborhood' radius ($\epsilon = 0.05$) and a moderate minimum number of points ($N = 5$). This ensured that noise points were properly accounted for within the clustering scheme.

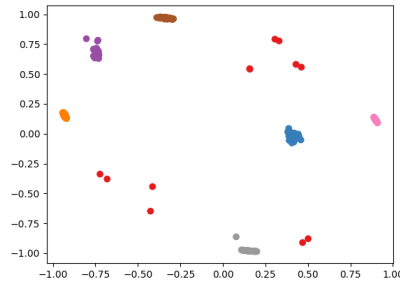
5.A.5 Dimer Geometries



(a)



(b)



(c)

Figure 5.A.10: Relative orientation features plotted, with $x-$ and $y-$ describing the relative alignment between the admolecule and the CO molecules in the second layer as described in the text. The plots show (a) all sampled geometries prior to optimizations, (b) all optimized geometries with color bar indicating the binding energy and (c) all optimized geometries after DBSCAN clustering where the colors indicate their grouping.

Dimer 1	x	y	z
C	-0.19045519661532	-0.16135156541274	-0.90552295962114
O	-0.24781144304162	-0.21210633747172	0.22651845170074
C	2.05554035880910	1.76473141430682	1.93645844013598
O	1.92163442084784	1.61431360857763	3.05282589778442
Dimer 2	x	y	z
C	-0.02861196368715	0.14226748736038	-0.25438691023046
O	0.02917690189452	0.20456101059696	0.87674186711502
C	-0.05833331699430	-1.15110586453745	-3.94238709981301
O	0.00624938878693	-2.06114797341989	-3.26806560707154
Dimer 3	x	y	z
C	0.33269802910679	-0.27970796365053	-0.42781900490899
O	0.22589212548614	-0.11896432978238	0.69006372841763
C	-2.77905130484776	1.47739207659477	-2.53763134041904
O	-2.08999886745526	1.87256599749566	-1.72746787668853
Dimer 4	x	y	z
C	-0.46825007893992	0.20316684833911	-0.60019754337303
O	-0.47660149193020	0.25745548182980	0.53294588147191
C	3.05170157807967	-1.08533323930954	-0.35691671835964
O	3.01979134766368	-2.21624085850747	-0.43999546289307

Table 5.A.3: Cartesian coordinates of dimer geometries.

5.B Bibliography

[1] RW Wilson, KB Jefferts, and AA Penzias. Carbon monoxide in the orion nebula. *The Astrophysical Journal*, 161:L43, 1970.

[2] BT Soifer, RC Puetter, RW Russell, SP Willner, PM Harvey, and FC Gillett. The 4-8 micron spectrum of the infrared source W33A. *The Astrophysical Journal*, 232:L53–L57, 1979.

[3] SA Sandford, LJ Allamandola, AGGM Tielens, and GJ Valero. Laboratory studies of the infrared spectral propertries of CO in astrophysical ices. *The Astrophysical Journal*, 329:498–510, 1988.

[4] H. M. Cuppen, E. M. Penteado, K. Isokoski, N. van der Marel, and H. Linnartz. CO ice mixed with CH_3OH: the answer to the non-detection of the 2152 cm^{-1} band? *Monthly Notices of the Royal Astronomical Society*, 417(4):2809–2816, 11 2011. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2011.19443.x. URL `https://doi.org/10.1111/j.1365-2966.2011.19443.x`.

[5] A.C. Adwin Boogert, Perry A. Gerakines, and Douglas C.B. Whittet. Observations of the icy universe. *Annual Review of Astronomy and Astrophysics*, 53(1): 541–581, 2015. doi: 10.1146/annurev-astro-082214-122348.

[6] Melissa K McClure, WRM Rocha, KM Pontoppidan, N Crouzet, Laurie EU Chu, E Dartois, T Lamberts, JA Noble, YJ Pendleton, G Perotti, *et al.* An ice age jwst inventory of dense molecular cloud ices. *Nature Astronomy*, pages 1–13, 2023.

[7] Akira Kouchi, Masashi Tsuge, Tetsuya Hama, Hiromasa Niinomi, Naoki Nakatani, Takashi Shimonishi, Yasuhiro Oba, Yuki Kimura, Sin-iti Sirono, Satoshi Okuzumi, *et al.* Formation of chiral CO polyhedral crystals on icy interstellar grains. *Monthly Notices of the Royal Astronomical Society*, 505(1):1530–1542, 2021.

[8] Akira Kouchi, Masashi Tsuge, Tetsuya Hama, Yasuhiro Oba, Satoshi Okuzumi, Sin-iti Sirono, Munetake Momose, Naoki Nakatani, Kenji Furuya, Takashi Shimonishi, *et al.* Transmission electron microscopy study of the morphology of ices composed of H_2O, CO_2, and CO on refractory grains. *The Astrophysical Journal*, 918(2):45, 2021.

[9] Jiao He, Francis E Toriello, Shahnewaz M Emtiaz, Thomas Henning, and Gianfranco Vidali. Phase transition of interstellar CO ice. *The Astrophysical Journal Letters*, 915(1):L23, 2021.

[10] JA Noble, E Congiu, F Dulieu, and HJ Fraser. Thermal desorption characteristics of CO, O_2 and CO_2 on non-porous water, crystalline water and silicate surfaces at submonolayer and multilayer coverages. *Monthly Notices of the Royal Astronomical Society*, 421(1):768–779, 2012.

[11] Scott A Sandford and Louis J Allamandola. The condensation and vaporization behavior of H_2O:CO ices and implications for interstellar grains and cometary activity. *Icarus*, 76(2):201–224, 1988.

[12] K Acharyya, GW Fuchs, HJ Fraser, EF Van Dishoeck, and H Linnartz. Desorption of CO and O_2 interstellar ice analogs. *Astronomy & Astrophysics*, 466(3):1005–1012, 2007.

[13] SE Bisschop, HJ Fraser, KI Öberg, EF Van Dishoeck, and S Schlemmer. Desorption rates and sticking coefficients for CO and N_2 interstellar ices. *Astronomy & Astrophysics*, 449(3):1297–1309, 2006.

[14] Lucas R Smith, Murthy S Gudipati, Rachel L Smith, and Robert D Lewis. Isotope effect on the sublimation curves and binding energies of ^{12}CO and ^{13}CO interstellar ice analogues. *Astronomy & Astrophysics*, 656:A82, 2021.

[15] Lorenzo Zamirri, Marta Corno, Albert Rimola, and Piero Ugliengo. Forsterite surfaces as models of interstellar core dust grains: computational study of carbon monoxide adsorption. *ACS Earth and Space Chemistry*, 1(7):384–398, 2017.

[16] Aurèle Germain, Marta Corno, and Piero Ugliengo. Computing binding energies of interstellar molecules by semiempirical quantum methods: Comparison between dft and gfn2 on crystalline ice. In *Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part V 21*, pages 632–645. Springer, 2021.

[17] Marc C. van Hemert, Junko Takahashi, and Ewine F. van Dishoeck. Molecular Dynamics Study of the Photodesorption of CO Ice. *Journal of Physical Chemistry A*, 119(24):6354–6369, jun 2015. ISSN 15205215. doi: 10.1021/acs.jpca.5b02611. URL `https://doi.org/10.1021/acs.jpca.5b02611`.

[18] G Molpeceres, V Zaverkin, N Watanabe, and J Kästner. Binding energies and sticking coefficients of H_2 on crystalline and amorphous CO ice. *Astronomy & Astrophysics*, 648:A84, 2021.

[19] GW Fuchs, HM Cuppen, S Ioppolo, C Romanzin, SE Bisschop, S Andersson, EF Van Dishoeck, and H Linnartz. Hydrogenation reactions in interstellar CO ice analogues-a combined experimental/theoretical approach. *Astronomy & Astrophysics*, 505(2):629–639, 2009.

[20] LJ Karssemeijer and HM Cuppen. Diffusion-desorption ratio of adsorbed co and co2 on water ice. *Astronomy & Astrophysics*, 569:A107, 2014.

[21] Trish Lauck, Leendertjan Karssemeijer, Katherine Shulenberger, Mahesh Rajappan, Karin I Öberg, and Herma M Cuppen. CO diffusion into amorphous H_2O ices. *The Astrophysical Journal*, 801(2):118, 2015.

[22] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017. URL `http://stacks.iop.org/0953-8984/29/i=27/a=273002`.

[23] S. R. Bahn and K. W. Jacobsen. An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.*, 4(3):56–66, MAY-JUN 2002. ISSN 1521-9615. doi: 10.1109/5992.998641.

[24] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982.

[25] MBBJM Tuckerman, Bruce J Berne, and Glenn J Martyna. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, 97(3):1990–2001, 1992.

[26] Roger Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2013.

[27] Herbert Edelsbrunner and Ernst P Mücke. Three-dimensional alpha shapes. *ACM Transactions On Graphics (TOG)*, 13(1):43–72, 1994.

[28] Frank Neese. The ORCA program system. *WIREs Computational Molecular Science*, 2(1):73–78, jan 2012. ISSN 1759-0876. doi: 10.1002/wcms.81. URL https://doi.org/10.1002/wcms.81.

[29] Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The ORCA quantum chemistry program package. *Journal of Chemical Physics*, 152(22):224108, jun 2020. ISSN 10897690. doi: 10.1063/5.0004608. URL https://doi.org/10.1063/5.0004608.

[30] Stefan Grimme. Density functional theory with london dispersion corrections. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(2):211–228, 2011.

[31] Eike Caldeweyher, Sebastian Ehlert, Andreas Hansen, Hagen Neugebauer, Sebastian Spicher, Christoph Bannwarth, and Stefan Grimme. A generally applicable atomic-charge dependent London dispersion correction. *Journal of Chemical Physics*, 150(15), 2019. ISSN 00219606. doi: 10.1063/1.5090222.

[32] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics*, 132(15):154104, 2010. ISSN 00219606. doi: 10.1063/1.3382344. URL https://doi.org/10.1063/1.3382344.

[33] Oleg A. Vydrov and Troy Van Voorhis. Nonlocal van der Waals density functional: The simpler the better. *Journal of Chemical Physics*, 133(24):244103, dec 2010. ISSN 00219606. doi: 10.1063/1.3521275. URL https://doi.org/10.1063/1.3521275.

[34] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005. ISSN 14639076. doi: 10.1039/b508541a. URL http://dx.doi.org/10.1039/B508541A.

[35] Jingjing Zheng, Xuefei Xu, and Donald G. Truhlar. Minimally augmented Karlsruhe basis sets. *Theoretical Chemistry Accounts*, 128(3):295–305, 2011. ISSN 1432881X. doi: 10.1007/s00214-010-0846-z. URL https://doi.org/10.1007/s00214-010-0846-z.

[36] Germán Molpeceres and Johannes Kästner. Computational Study of the Hydrogenation Sequence of the Phosphorous Atom on Interstellar Dust Grains. *Astrophys. J.*, 910(1):55, 2021. ISSN 0004-637X. doi: 10.3847/1538-4357/abe38c.

[37] L. Martinez, R. Andrade, E. G. Birgin, and J. M. Martínez. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comp. Chem.*, 30(13):2157–2164, 2009. ISSN 1096987X. doi: 10.1002/jcc.21224. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21224.

[38] Sebastian Spicher and Stefan Grimme. Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems. *Angew. Chem. Int. Ed.*, 59(36):15665–15673, 2020. ISSN 15213773. doi: 10.1002/anie.202004239. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202004239.

[39] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comp.*, 15(3):1652–1671, mar 2019. ISSN 1549-9618. doi: 10.1021/acs.jctc.8b01176. URL https://doi.org/10.1021/acs.jctc.8b01176.

[40] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, 1996. ISSN 10797114. doi: 10.1103/PhysRevLett.77.3865. URL http://www.ncbi.nlm.nih.gov/pubmed/10062328%5Cnhttp://link.aps.org/doi/10.1103/PhysRevLett.77.3865%5Cnhttp://link.aps.org/abstract/PRL/v77/p3865. ISBN: 9780596529321.

[41] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993. ISSN 00219606. doi: 10.1063/1.464913. URL http://scitation.aip.org/content/aip/journal/jcp/98/7/10.1063/1.464913.

[42] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, 110(13):6158–6170, mar 1999. ISSN 0021-9606. doi: 10.1063/1.478522. URL https://doi.org/10.1063/1.478522.

[43] Axel D. Becke. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.*, 98(2):1372–1377, 1993. ISSN 00219606. doi: 10.1063/1.464304. URL https://doi.org/10.1063/1.464304.

[44] Yan Zhao and Donald G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theor. Chem. Acc.*, 120 (1-3):215–241, jul 2008. ISSN 1432881X. doi: 10.1007/s00214-007-0310-x. URL http://link.springer.com/10.1007/s00214-007-0310-x.

[45] Yan Zhao and Donald G Truhlar. Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *J. Phys. Chem. A*, 109(25):5656–67, jun 2005. ISSN 1089-5639. doi: 10.1021/jp050536c. URL http://www.ncbi.nlm.nih.gov/pubmed/16833898.

[46] Narbe Mardirossian and Martin Head-Gordon. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.*, 144(21):214110, jun 2016. ISSN 0021-9606. doi: 10.1063/1.4952647. URL https://doi.org/10.1063/1.4952647.

[47] Lorenzo Zamirri, Silvia Casassa, Albert Rimola, Mireia Segado-Centellas, Cecilia Ceccarelli, and Piero Ugliengo. Ir spectral fingerprint of carbon monoxide in interstellar water–ice models. *Monthly Notices of the Royal Astronomical Society*, 480(2):1427–1444, 2018.

[48] Jessica Perrero, Joan Enrique-Romero, Stefano Ferrero, Cecilia Ceccarelli, Linda Podio, Claudio Codella, Albert Rimola, and Piero Ugliengo. Binding energies of interstellar relevant s-bearing species on water ice mantles: A quantum mechanical investigation. *arXiv preprint arXiv:2209.07255*, 2022.

[49] Lorenzo Tinacci, Aurele Germain, Stefano Pantaleone, Stefano Ferrero, Cecilia Ceccarelli, and Piero Ugliengo. Theoretical distribution of the ammonia binding energy at interstellar icy grains: a new computational framework. *ACS Earth and Space Chemistry*, 2022.

[50] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, *et al.* A density-based algorithm for discovering clusters in large spatial databases with noise. volume 96, pages 226–231, 1996.

[51] Aurore Bacmann, Bertrand Lefloch, Cecilia Ceccarelli, Alain Castets, Juergen Steinacker, and Laurent Loinard. The degree of CO depletion in pre-stellar cores. *Astronomy & Astrophysics*, 389(1):L6–L10, 2002.

[52] Eric Keto and Paola Caselli. The different structures of the two classes of starless cores. *The Astrophysical Journal*, 683(1):238, 2008.

[53] Eric Keto and Paola Caselli. Dynamics and depletion in thermally supercritical starless cores. *Monthly Notices of the Royal Astronomical Society*, 402(3):1625–1634, 2010.

[54] S Cazaux, R Martín-Doménech, YJ Chen, GM Muñoz Caro, and C González Díaz. CO depletion: A microscopic perspective. *The Astrophysical Journal*, 849 (2):80, 2017.

[55] Rut M Berns and Ad van der Avoird. N_2–N_2 interaction potential from abinitio calculations, with application to the structure of $(N_2)_2$. *The Journal of Chemical Physics*, 72(11):6107–6116, 1980.

[56] MC van Hemert. Potential energy surface for the study of inelastic collisions between nonrigid co and H_2. *The Journal of Chemical Physics*, 78(5):2345–2354, 1983.