



**HAL**  
open science

# Simulating core electron binding energies of halogenated species adsorbed on ice surfaces and in solution with relativistic quantum embedding calculations

Richard A Opoku, Céline Toubin, André Severo Pereira Gomes

## ► To cite this version:

Richard A Opoku, Céline Toubin, André Severo Pereira Gomes. Simulating core electron binding energies of halogenated species adsorbed on ice surfaces and in solution with relativistic quantum embedding calculations. *Physical Chemistry Chemical Physics*, 2022, 24, pp.14390-14407. 10.1039/D1CP05836C . hal-03500186v2

**HAL Id: hal-03500186**

**<https://hal.science/hal-03500186v2>**

Submitted on 18 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# PCCP

Physical Chemistry Chemical Physics

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: R. A. Opoku, C. Toubin and A. Severo Pereira Gomes, *Phys. Chem. Chem. Phys.*, 2022, DOI: 10.1039/D1CP05836C.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Cite this: DOI: 00.0000/xxxxxxxxxx

# Simulating core electron binding energies of halogenated species adsorbed on ice surfaces and in solution with relativistic quantum embedding calculations<sup>†</sup>

Richard A. Opoku,<sup>a</sup> Céline Toubin,<sup>a</sup> and André Severo Pereira Gomes<sup>\*a</sup>

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

In this work we investigate the effects of the environment on the X-ray photoelectron spectra of hydrogen chloride and the chloride ions adsorbed on ice surfaces, as well as of chloride ions in water droplets. In our approach, we combine a density functional theory (DFT) description of the ice surface with that of the halogen species with the recently developed relativistic core-valence separation equation of motion coupled cluster (CVS-EOM-IP-CCSD) via the frozen density embedding formalism (FDE), to determine the K and L<sub>1,2,3</sub> edges of chlorine. Our calculations, which incorporate temperature effects through snapshots from classical molecular dynamics simulations, are shown to reproduce experimental trends in the change of the core binding energies for Cl<sup>-</sup> upon moving from a liquid (water droplets) to an interfacial (ice quasi-liquid layer) environment. Our simulations yield water valence band binding energies in good agreement with experiment, and that vary little between the droplets and the ice surface. For the halide core binding energies there is an overall trend of overestimating experimental values, though good agreement between theory and experiment is found for Cl<sup>-</sup> in water droplets and on ice. For HCl on the other hand there are significant discrepancies between experimental and calculated core binding energies when we consider structural models which maintain the H-Cl bond more or less intact. An analysis of models that allow for pre-dissociated and dissociated structures suggests that experimentally observed chemical shifts in binding energies between Cl<sup>-</sup> and HCl would require that H<sup>+</sup> (in the form of H<sub>3</sub>O<sup>+</sup>) and Cl<sup>-</sup> are separated by roughly 4–6 Å.

## 1 Introduction

Ice is everywhere in the environment and its peculiar structure and properties make it a subject of intense scientific research. Studies connected to ice indicate that it hosts reactions that can influence climate, air quality, biology systems and initiate ozone destruction<sup>1–3</sup>. Hydrogen Bonding (HB) between ice and trace gases is the first step towards their interaction<sup>4</sup>. Investigations into the bound state of strong acids<sup>5</sup> interaction with ice indicate that strong acids lead to modification in the HB network of the liquid-like layer on ice surface. It has been shown however that, weak acids adsorption on ice surface does not produce any significant changes in the HB network of water ice.<sup>6</sup>

In this respect, the influence of strong acids, in particular, hydrogen chloride (HCl) and its dissociated ionic chloride by ice

has attracted a lot of attention over the years due to their link to ozone depletion<sup>7</sup>.

A better understanding of the chemical processes associated with how the reactions of these reservoir gases at the ice surface differ from that of the bulk ice surface is essential to their interpretation, and consequently to atmospheric science and environmental chemistry. Surface-specific spectroscopic approaches have been instrumental in gathering detailed information on structural and electronic properties of solvated halide/halide ions<sup>8,9</sup>, and among these X-ray Photoelectron Spectroscopy (XPS) stands out as a particularly powerful technique<sup>10</sup> due to its high specificity, and the great sensitivity of core binding energies (BEs) to small perturbations to the surroundings of the atoms of interest<sup>11–15</sup>.

The surface sensitivity and chemical selectivity of radiations from XPS has made it possible to investigate the loss of gas-phase molecules as well as their behavior and transformation in complex reactions or solvent mediums.<sup>11,16–22</sup>

This is illustrated in recent investigations of the electronic structure of halogen-related systems interacting with solvent environment<sup>4,14,23</sup>, for which the evidence from XPS suggests the dependence of chemical and solvent binding energy induced

<sup>a</sup>Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers, Atomes et Molécules, F-59000 Lille, France

\*E-mail: andre.gomes@univ-lille.fr

<sup>†</sup> Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 00.0000/00000000 and <https://zenodo.org/record/5729761>



shifts on the HB network configuration of the solvent to the halide systems. In a pioneering work by McNeill *et al.*<sup>24</sup> it has been shown that the quasiliquid layer (QLL) plays an essential role in influencing the sorption behavior of HCl and on the chemistry of environmental ice surfaces,<sup>24,25</sup> for which the evidence of XPS suggests that the dissociated form of HCl perturbs the HB network of the liquid-like layer on ice.<sup>4,26</sup> From these studies, it is observed that the ionization of HCl on ice surfaces follow a janus-like behavior, where molecular HCl is formed on the ice surface and its dissociated form is observed at the uppermost bulk layer of the ice. In addition, there is a long standing debate on whether the dissociation of HCl on ice surfaces is temperature dependent. While some studies show that the dissociation occurs only at high temperatures<sup>27,28</sup>, other studies indicate that this mechanism can occur even at low temperatures.<sup>4,29</sup>

As the physical and chemical processes at play (with respect to the interaction between adsorbed species and the ice surface, as well as the interaction between the incoming X-rays and the sample) are quite complex, it is difficult to make sense of the experimental results without the help of theoretical models, both in terms of geometric (the arrangement of the atoms) and electronic structure components.

From the electronic structure standpoint the problem consists of determining core binding energy for a particular atom and edge while incorporating the effects of the environment—which may go well beyond the immediate surroundings of the atoms of interest, and may be quite severely affected by the structural changes mentioned above. That requires first the treatment of electron correlation and relaxation effects, for which the equation of motion coupled cluster (EOMCC)<sup>30–36</sup>, combined with the physically motivated CVS approximation<sup>37</sup>, has proven to reliably target core states in an efficient manner with little modification to standard diagonalization approaches<sup>38–42</sup>. Second, it is essential to use relativistic Hamiltonians<sup>43</sup> in order to capture the changes in core BEs due to scalar relativistic and spin-orbit coupling effects (the latter being responsible for the splitting of the L, M and N edges). There, methods based on the transformation from 4- to 2-component approaches are particularly useful for correlated calculations such as those with CC approaches<sup>44–46</sup>.

Due to the relativistic correlated electronic structure methods' steep computational scaling with system size  $N$  ( $O(N^6)$  for CCSD-based methods), the incorporation of the environment surrounding the species of interest on the calculations is in general not possible beyond a few nearest neighbors. In this case embedding theories<sup>47–52</sup>, in which a system is partitioned into a collection of interacting subsystems, are a very cost-effective approach. Among the embedding approaches, classical embedding (QM/MM) models (continuum models, point-charge embedding, classical force fields etc) are computationally very efficient but at the cost of foregoing any prospect of extracting electronic information from the environment, and will be bound by the limitations of the classical models (e.g. the difficulties of continuum models to account for specific interactions such as hydrogen bonding). Purely quantum embedding approaches (QM/QM), on the other hand, may be more costly but with the advantage of permitting one to extract information from the electron density or wavefunctions of the en-

vironment and as such have been used to study absorption<sup>53</sup> and reaction energies<sup>54</sup>, electronically excited<sup>55–61</sup> and ionized<sup>62,63</sup> states of species of experimental interest.

The frozen density embedding (FDE) approach is a particularly interesting QM/QM method since it provides a framework to seamlessly combine CC and DFT approaches (CC-in-DFT) for both ground and excited states<sup>47,55,60,64</sup>. It has been shown to successfully tackle the calculation of valence electron BEs of halogens in water droplets<sup>62</sup>, while providing rather accurate valence water binding energies with no additional effort to represent the system of interest. Such an approach would be particularly interesting for addressing XPS spectra, as it would allow the calculation of the core spectra with correlated approaches for the species of interest while providing information on the valence band of the environment, which can then serve as an internal reference and help in comparisons to experiment.

To the best of our knowledge, there have been rather few studies employing QM/QM embedding for simulating core spectra: Parravicini and Jagau<sup>65</sup> have investigated projection-based CC-in-DFT embedding for the calculation of carbon K edge ionization energies (as well as valence excitation, ionization energies and resonances for first- and second-row model systems), employing non-relativistic Hamiltonians.

Thus, in this contribution, we aim to provide a computational protocol based on relativistic CC-in-DFT calculations, that can be used in a black-box manner to obtain absolute core binding energies (which are heavily dependent on the Hamiltonian and electronic structure approach employed) for species containing atoms beyond the first row, and in complex environments. To this end, we combine the basic ingredients of the computational protocol of Bouchafrat *et al.*<sup>62</sup> with the relativistic CVS-EOM-IP-CCSD method<sup>42</sup>, and investigate the performance of the resulting CVS-EOM-IP-CCSD-in-DFT method in the determination of chlorine (Cl) core electron BEs, and associated chemical shifts, for hydrogen chloride (HCl) and ionic chloride ( $\text{Cl}^-$ ) adsorption on ice surfaces. We shall also profit of this investigation and determine the ionic chloride BEs on a water droplet model<sup>62</sup>.

The manuscript is organized as follows: the basic theoretical aspects of DFT-in-DFT and CC-in-DFT embedding are outlined in section 2, with a description of the structural models (alongside the details of the calculations) provided in section 3. The discussion of our results, and conclusions are presented in sections 4 and 5, respectively.

## 2 Theoretical approaches

### 2.1 Frozen density embedding (FDE) method

The main idea of FDE<sup>47,66</sup> is the separation of the total electron density  $\rho_{tot}$  of a system into a number of density subsystems. Two subsystems are considered in our case and the whole system is represented as the sum of the density of the subsystem of interest,  $\rho_\alpha$  and subsystem of the environment,  $\rho_\beta$  (i.e.  $\rho_{tot} = \rho_\alpha + \rho_\beta$ ). The  $\rho_\beta$  is considered to be frozen in this approximation. The corresponding total energy of the whole system is based on the electron densities of the subsystems and can then be written as

$$E_{tot}[\rho_{tot}] = E_\alpha[\rho_\alpha] + E_\beta[\rho_\beta] + E_{int}[\rho_\alpha, \rho_\beta] \quad (1)$$





where  $E_{int}[\rho_\alpha, \rho_\beta]$  is the energy obtained from the interaction of the two subsystems, which is known as the interaction energy of the system. The interaction energy is given as

$$E_{int}[\rho_\alpha, \rho_\beta] = E_{int}^{NN} + \int \rho_\alpha(r) v_\beta^{nuc}(r) + \rho_\beta(r) v_\alpha^{nuc}(r) + \int \frac{\rho_\alpha(r) \rho_\beta(r')}{|r-r'|} d^3r d^3r' + E_{xc}^{nadd}[\rho_\alpha, \rho_\beta] + T_s^{nadd}[\rho_\alpha, \rho_\beta] \quad (2)$$

where  $E_{int}^{NN}$  is the nuclear repulsion energy between subsystems,  $v_\alpha^{nuc}$  and  $v_\beta^{nuc}$  are the electrostatic potential of the nuclei in subsystems  $\alpha$  and  $\beta$  respectively.  $E_{xc}^{nadd}$  and  $T_s^{nadd}$  are the non-additive contributions due to exchange-correlation and kinetic energies respectively are defined as

$$E_{xc}^{nadd}[\rho_\alpha, \rho_\beta] = E_{xc}[\rho_\alpha + \rho_\beta] - E_{xc}[\rho_\alpha] - E_{xc}[\rho_\beta] \quad (3)$$

$$T_s^{nadd}[\rho_\alpha, \rho_\beta] = T_s[\rho_\alpha + \rho_\beta] - T_s[\rho_\alpha] - T_s[\rho_\beta]$$

The non-additive kinetic energy and the non-additive exchange-correlation energy take into consideration non-classical contributions to the energy. The non-additive kinetic energy prevents spurious delocalization among the subsystems as observed by balancing the attractive interaction in the nuclear framework of one subsystem and the electron density of another subsystem. The FDE uses only the electron density in the calculation of interaction between subsystems without the sharing of orbital information among the subsystems. Minimization of the total energy of the system with respect to  $\rho_\alpha$  yields an Euler-Lagrangian equation that keeps the number of electrons in the subsystem of interest fixed.<sup>47</sup>

The application of the Euler-Lagrangian equation in the FDE allows the molecular system to be subdivided into smaller interacting fragments and each of them being treated at the most suitable level of theory. Although based on DFT, the FDE scheme also allows treatment of one of the subsystems with wave function method and the rest of the subsystems with DFT (WFT-in-DFT)<sup>57,67</sup> or treating all the subsystems with wave function (WFT-in-WFT)<sup>58</sup>. Several literatures have implemented such WFT-in-DFT, in particular coupling of CC with DFT to accurately probe the excitation energies<sup>47,60,61</sup> and ionization energies<sup>62</sup> of numerous molecules.

To obtain the electron density of the subsystem of interest in Kohn-Sham (KS) DFT, the total energy,  $E_{tot}[\rho_{tot}]$  is minimized concerning  $\rho_\alpha$ , while the electron density of the subsystem of the environment is kept frozen. It is performed under the restriction that the number of electrons in subsystem  $\alpha$  is fixed, with the orbitals of the embedded system generated from a set of KS-like equations,

$$[T_s(i) + v_{eff}^{KS}[\rho_\alpha] + v_{int}^\alpha[\rho_\alpha, \rho_\beta] - \epsilon_i] \phi_i^\alpha(r) = 0 \quad (4)$$

where  $T_s(i)$  and  $v_{eff}^{KS}[\rho_\alpha]$  are the KS kinetic energy and effective potential of the isolated subsystem of interest respectively. The embedding potential which describes the interaction between

subsystem  $\alpha$  and the frozen subsystem  $\beta$  is

$$v_{int}^\alpha[\rho_\alpha, \rho_\beta](r) = V_\beta^{nuc}(r) + \int \frac{\rho_\beta(r')}{|r-r'|} dr' + \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \Big|_{\rho=\rho_{tot}} - \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \Big|_{\rho=\rho_\alpha} + \frac{\delta T_s[\rho]}{\delta \rho(r)} \Big|_{\rho=\rho_{tot}} + \frac{\delta T_s[\rho]}{\delta \rho(r)} \Big|_{\rho=\rho_\alpha} \quad (5)$$

## 2.2 Core-Valence Separation (CVS) Equation-of-motion coupled cluster (CVS-EOM-CC) theory

In the CVS-EOM-IP-CCSD method, the BEs are obtained from the solution of the projected eigenvector and its corresponding eigenvalue equation<sup>38,39,42</sup>

$$P_c^v(\bar{H}P_c^v R_k^{IP}) = \Delta E_k P_c^v R_k^{IP} \quad (6)$$

where  $\Delta E_k$  is the ionization energy of the system,  $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$  is a similarity transformed Hamiltonian including equation 5,  $P_c^v$  is a projector introduced to restrict all elements of valence orbitals to zero and  $R_k^{IP}$  is the operator that transforms the coupled-cluster ground-state to electron detachment states.  $R_k^{IP}$  is given as

$$R_k^{IP} = \sum_i r_i \{a_i\} + \sum_{i>j,a} r_{ij}^a \{a_i^\dagger a_j a_i\} \quad (7)$$

## 3 Computational details

### 3.1 Electronic structure calculations

All DFT<sup>68</sup> and DFT-in-DFT<sup>69</sup> calculations have been performed with the 2017 version of the ADF code<sup>70</sup>, employing the scalar relativistic zeroth-order regular approximation (ZORA) Hamiltonian<sup>71</sup> and triple zeta basis sets with polarization function (TZP)<sup>72</sup>. In the case of single-point calculations and in the determination of embedding potentials, the statistical average of orbital potentials (SAOP) model potential<sup>73,74</sup> was used for the exchange-correlation potential of the subsystems, whereas the PBE<sup>75</sup> and PW91k<sup>76</sup> density functionals were employed for the non-additive exchange-correlation and kinetic energy contributions, respectively. In embedding calculations no frozen cores were employed. In the case of geometry optimizations, the PBE functional was used, along with the large core option. All integration grids were taken as the default in ADF. Embedding calculations have been performed via the PYADF scripting framework<sup>77</sup>.

All coupled-cluster calculations were carried out with the DIRAC electronic structure code<sup>78</sup> (with the DIRAC19<sup>79</sup> release and revisions ddbfa6a, 0757608, 323ab67, 2628039, 1e798e5, b9f45bd). The Dyal basis sets<sup>80-82</sup> of triple-zeta quality, complemented with two diffuse functions for each angular momenta as in<sup>62</sup> (d-aug-dyall.acv3z) were employed for chloride, while the Dunning aug-cc-pVTZ basis sets<sup>83</sup> have been employed for hydrogen and oxygen. The basis sets were kept uncontracted in all calculations. In order to estimate the complete basis set limit (CBS) of CVS-EOM-IP-CCSD calculations, we also carried out calculations with quadruple-zeta quality bases (d-aug-dyall.acv4z and aug-cc-pVQZ) for selected systems, and used a



two-point formula as carried out by Bouchafra *et al.*<sup>62</sup>.

Apart from the Dirac–Coulomb (<sup>4</sup>DC) Hamiltonian, we employed the molecular mean-field<sup>44</sup> approximation to the Dirac–Coulomb–Gaunt (<sup>2</sup>DCG<sup>M</sup>) Hamiltonian. In it the Gaunt-type integrals are explicitly taken into account only during the 4-component SCF step, as the transformation of these to MO basis is not implemented. Unless otherwise noted, we employed the usual approximation of the energy contribution from (*SS*|*SS*)-type two-electron integrals by a point-charge model.<sup>84</sup> In CC-in-DFT calculations, the embedding potential obtained (with ADF) at DFT-in-DFT level is included in DIRAC as an additional one-body operator to the Hamiltonian, following the setup outlined in<sup>55</sup>.

Unless otherwise noted, all occupied and virtual spinors were considered in the correlation treatment. The core binding energy calculations with CVS-EOM-IP-CCSD<sup>42</sup> were performed for the K, L<sub>1</sub>, L<sub>2</sub> and L<sub>3</sub> edges of the chlorine atom. The energies so obtained represent electronic states with main contributions arising from holes in the 1*s*, 2*s*, 2*p*<sub>1/2</sub> and 2*p*<sub>3/2</sub> spinors, respectively.

The datasets associated with this study are available at the Zenodo repository<sup>85</sup>.

### 3.2 MD-derived structures

The structures for Cl<sup>−</sup> in water droplets simulated at temperature of 300 K have been taken from Bouchafra *et al.*<sup>62</sup>, and originate from classical molecular dynamics (CMD) simulations employing polarizable force fields<sup>86</sup>. Here we have considered the same 100 snapshots as in the original reference. Each droplet contains 50 water molecules, and the halogen position has been constrained to be at the center of mass of the system.

Initial structures of the halogens adsorbed on the ice surfaces have been taken from Woittequand *et al.*<sup>87</sup>, which are based on CMD simulation of HCl adsorbed on the ice surfaces with a non-polarizable force field<sup>88</sup> at 210 K. We have considered 25 snapshots, each containing 216 water molecules. It should be noted that this set of structures account for the disorder at the air-ice interface associated with a thin ice quasi-liquid layer (QLL).

Due to the nature of the force field, structures for Cl<sup>−</sup> were not available for the same surface, and we have therefore started out from the HCl snapshots, removed a proton and proceeded to optimizations of the ion position while constraining the water molecules of the ice surface to keep their original positions. As such, the adsorption site is slightly altered with respect to the original HCl-ice system, but not the surface on which adsorption takes place.

Furthermore, to assess the importance of HCl-water interactions not captured by the classical force field, we have applied a constrained optimization to the HCl species as well, in a similar vein as outlined above, for all CMD snapshots. We have considered two situations: one in which only the position of HCl was allowed to change (thus allowing both changes in H-Cl bond distance and in relative position of H and Cl with respect to the surface), and another in which the atoms for the six waters closest to HCl were also allowed to change position.

We note that considering the charged system without a counter

ion implicitly assumes a model for a diluted solution/interface. From prior work with charged species in the literature such an approach is warranted if one takes into account the effects of the polarization of the solvent<sup>86,89–94</sup>, and has been shown to yield spectroscopic results<sup>62,95</sup> that closely match experimental measurements, provided of course the model mirrors the essential features of the experimental system.

### 3.3 Embedding models

For Cl<sup>−</sup> in water droplets a single embedding model is used, in which two subsystems are defined: the active subsystem (treated with CC), containing the halogen, and the environment (treated with DFT) composed of the 50 waters. Further details can be found in Bouchafra *et al.*<sup>62</sup>. For our discussion of the halogens adsorbed on ice we have considered three models: the first represents calculations without embedding (referred to as **SM**, for supermolecular model, in what follows). Furthermore, two embedding models are considered, the first (referred to as model **EM1** in the following) is similar to the droplet one in that only the species containing the halogen is contained in the active subsystem, and all water molecules make up the environment. In the second model (referred to as model **EM2** in the following), we include a number of water molecules (the nearest neighbors to the halogen species) in the active subsystem, and the remaining water molecules make up the environment. These models are pictorially represented in figure 1. In the following, **EM2<sup>V</sup>** will denote a model containing one water molecule in the active subsystem. Additional details can be found in the supplementary information.

## 4 Results and Discussion

Before proceeding to the discussion of our results for embedding systems, we consider it instructive to address first the performance of theoretical approaches for obtaining core binding energies for the isolated (gas-phase) species, since these provide us with a well-defined reference point with which to assess the behavior of embedded models in complement to a comparison to experimental results<sup>4,26</sup>.

### 4.1 Gas-phase calculations

Our results for gas-phase calculations are shown in table 1. Considering first the EOM calculations, our choice of focusing on the <sup>2</sup>DCG<sup>M</sup> Hamiltonian stems from the fact that the Gaunt interaction is essential for properly describing the K edge, while showing non-negligible effects for the L edges (see supplementary result for a comparison to two- and four-component results based on the Dirac–Coulomb Hamiltonian). Here, we also provide an investigation of basis set convergence with extended basis sets (including an extrapolation to the complete basis set limit), as well as a in-depth comparison to experimental results. For the SAOP model potential, on the other hand, we are not aware of any comparison for core ionization energies employing the equivalent of Koopmans theorem for DFT<sup>96</sup>—that is, by obtaining the core binding energy as the negative of the Kohn–Sham orbital energies (BE = −ε<sub>KS</sub>). A Koopmans' approach will be inherently less accurate



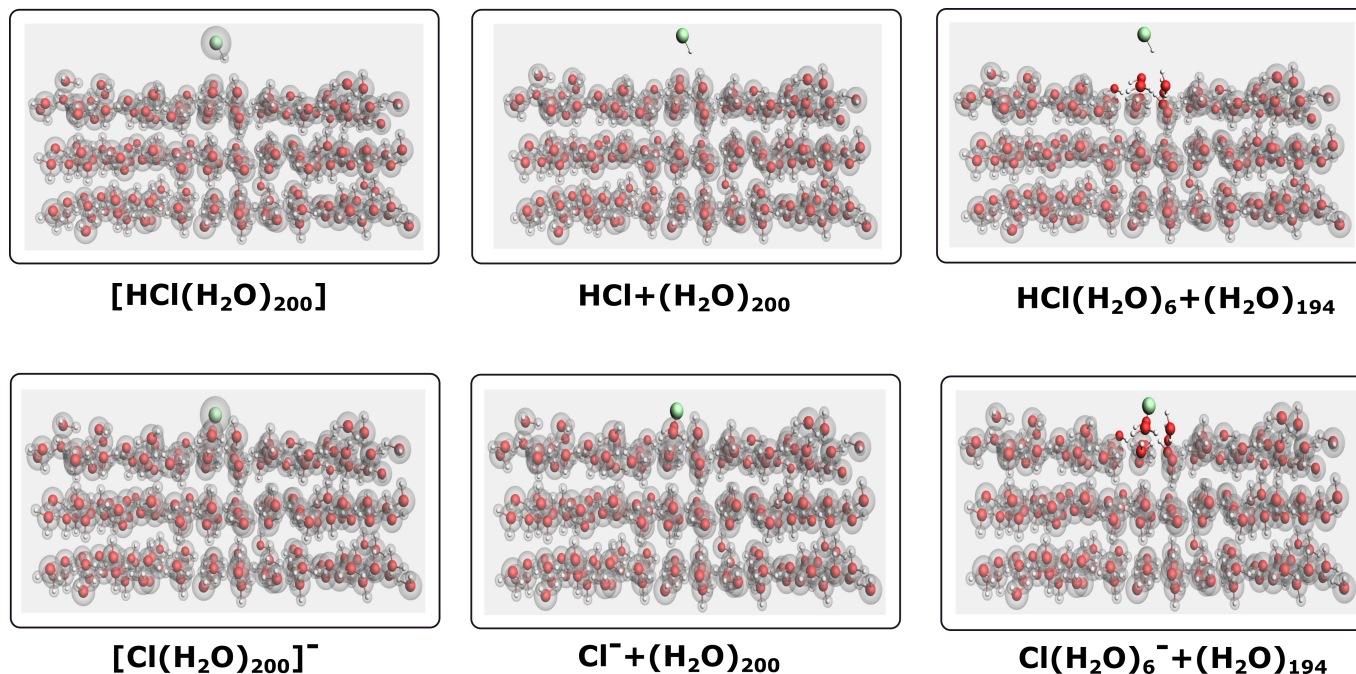


Fig. 1 Perspective views for cluster models of halogens adsorbed ice surfaces (represented by 200 water molecules): without a partition into subsystem (SM, left)), with an active subsystem containing only the halogen (EM1, center) and with an active subsystem containing the halogen and 6 nearest water molecules (EM2 right). Boxes A, B and C represent the system HCl-ice whereas boxes D, E and F represent the system Cl<sup>-</sup>-ice.

for core electron binding energies due to the lack of relaxation effects<sup>97</sup>, though it remains very convenient for a qualitative understanding of chemical shifts for a particular edge. In the case of SAOP, which has shown a very good performance for valence ionizations of halides and water in droplets<sup>62</sup>, it is interesting to verify by how much it deviates from EOM calculations for deeper ionizations.

The presentend SAOP binding energies correspond to those obtained with the scaled (spin-orbit) ZORA Hamiltonian (we recall that while scaled and unscaled ZORA energies differ significantly for deep cores, the eigenfunctions for both cases are the same<sup>98</sup>). A comparison of SO-ZORA SAOP and <sup>2</sup>DCG<sup>M</sup> EOM results shows, unexpectedly, a marked difference between the K, L<sub>1</sub>, and L<sub>2</sub>/L<sub>3</sub> edges. For L<sub>2</sub>/L<sub>3</sub> edges of both HCl and Cl<sup>-</sup> the difference between methods is roughly 16 eV, increases to roughly 27 eV for the L<sub>1</sub>, and reaches around 70 eV for the K edge, which represent differences in binding energies of roughly 8-9%, 10% and 2-3% for the respective edges. That the difference between SO-ZORA results and EOM is essentially the same for the two species will be useful for the comparison between structural models that follows.

From the table, we observe that our triple-zeta basis <sup>2</sup>DCG<sup>M</sup> EOM results overestimate the experimental L<sub>2</sub> and L<sub>3</sub> gas-phase HCl binding energies reported by Hayes and Brown<sup>99</sup> and Aitken *et al.*<sup>100</sup> by around 1 eV. A similar difference is observed for the L<sub>1</sub> edge of chloride in the NaCl crystal, with the L<sub>2</sub> and L<sub>3</sub> edges in this case showing deviations smaller than 1 eV. For the K edge there appears to be a slightly larger discrepancy (around 1.7 eV) between theory and the experimental values quoted by Thompson *et al.*<sup>101</sup>. In spite of the fact that for all edges these experimental results are not for a gas-phase chloride atom, we take the overall very good agreement to indicate that <sup>2</sup>DCG<sup>M</sup> EOM should show

an uniform accuracy for both species.

We note that there remain three potential sources of errors in our calculations: (a) basis set incompleteness; (b) QED and retardation effects; and (c) energy corrections due to higher-order excitations in the CC wavefunctions. For HCl there could be a fourth, the H-Cl bond distance, but as can be seen from results of a scan of this coordinate (see supplementary information), even large variation of bond lengths don't change energies by more than a 0.1-0.2 eV for the L edges, and around 0.3 eV for the K edge, meaning that the first three factors should be behind most of the discrepancy.

From the literature<sup>104</sup>, QED and retardation effects are expected to be well below 0.1 eV for chlorine. We are unable at this point to determine the effect of higher-order excitations with the current implementation in DIRAC. For assessing the basis set effects, we have calculated binding energies with quadruple zeta basis sets and with them and the triple-zeta results obtained the complete basis set limit (CBS) values shown in table 1.

Comparing the EOM CBS results to the experimental results of Hayes and Brown<sup>99</sup> and Aitken *et al.*<sup>100</sup> for the L<sub>2</sub> and L<sub>3</sub> edges of HCl, we see that agreement gets slightly worse than with the triple-zeta basis, and an interesting point is that for the K edge the effect of extrapolation is to reduce the binding energies whereas for the L edges the opposite is true. For Cl<sup>-</sup>, we see essentially the same variation between the triple zeta and the extrapolated values as for HCl. In both cases, as the L<sub>2</sub> and L<sub>3</sub> edges are affected by the same amounts, the spin-orbit splitting of the 2p<sub>1/2</sub> and 2p<sub>3/2</sub> remains at around 1.6 eV for both species, a value consistent with the gas-phase experimental values for HCl (in the NaCl crystal, this splitting is of 1.7 eV<sup>102</sup> whereas in KCl it is 1.6 eV<sup>103</sup>).





Table 1 CVS-EOM-IP-CCSD chlorine core binding energies (in eV) for HCl and  $\text{Cl}^-$  in gas-phase for the  $^2\text{DCG}^M$  Hamiltonian and employing triple-zeta basis sets as well as values extrapolated to the complete basis set limit (CBS). In addition to those, we present core binding energies obtained via the analogue of Koopmans theorem for DFT<sup>96</sup>, employing the SAOP model potential for the ZORA Hamiltonian. In parenthesis we present the differences with respect to the CVS-EOM-IP-CCSD  $^2\text{DCG}^M$  results with triple-zeta basis sets, which we take as reference. Apart from the energies for the individual edges, we provide the core binding energy shift ( $\Delta_{BE}$ , in eV) between HCl and  $\text{Cl}^-$  for the theoretical gas-phase values.

| Species       | Hamiltonian                     | Method | K                | L <sub>1</sub>  | L <sub>2</sub>  | L <sub>3</sub>  |
|---------------|---------------------------------|--------|------------------|-----------------|-----------------|-----------------|
| HCl           | SO-ZORA                         | SAOP   | 2764.58 (-69.32) | 253.94 (-26.75) | 194.08 (-15.94) | 192.41 (-15.98) |
|               | $^2\text{DCG}^M$                | EOM    | 2833.90 (0.00)   | 280.69 (0.00)   | 210.02 (0.00)   | 208.39 (0.00)   |
|               | $^2\text{DCG}^M$ , CBS          | EOM    | 2833.86 (-0.04)  | 280.79 (0.10)   | 210.18 (0.16)   | 208.55 (0.16)   |
|               | Exp. (gas phase) <sup>99</sup>  |        |                  |                 | 208.70          | 207.1           |
|               | Exp. (gas phase) <sup>100</sup> |        |                  |                 | 209.01          | 207.38          |
| $\text{Cl}^-$ | SO-ZORA                         | SAOP   | 2754.42 (-71.63) | 243.88 (-26.93) | 184.02 (-16.24) | 182.35 (-16.21) |
|               | $^2\text{DCG}^M$                | EOM    | 2824.17 (0.00)   | 270.73 (0.00)   | 200.10 (0.00)   | 198.47 (0.00)   |
|               | $^2\text{DCG}^M$ , CBS          | EOM    | 2824.13 (-0.04)  | 270.84 (0.11)   | 200.29 (0.20)   | 198.66 (0.19)   |
|               | Exp. (NaCl) <sup>102</sup>      |        |                  | 269.6           | 200.6           | 198.9           |
|               | Exp. (KCl) <sup>103</sup>       |        |                  |                 | 200.1           | 198.6           |
|               | Exp. (NaCl) <sup>101</sup>      |        | 2822.4           | 270             | 202             | 200             |
| $\Delta_{BE}$ | SO-ZORA                         | SAOP   | 10.16            | 10.06           | 10.06           | 10.06           |
|               | $^2\text{DCG}^M$                | EOM    | 9.73             | 9.96            | 9.92            | 9.92            |
|               | $^2\text{DCG}^M$ , CBS          | EOM    | 9.73             | 9.95            | 9.89            | 9.89            |

Finally, we see that the  $\Delta_{BE}$  value are nearly indistinguishable, in spite of the calculations having employed different Hamiltonians or correlated methods. This indicates that this is a robust quantity to (semi-quantitatively) characterize the chemical shift in the binding energies, even when the absolute binding energies are rather poor (as is the case of DFT orbital energies).

#### 4.2 Assessing the embedded models

Before proceeding to the calculation of halide binding energies considering the sampling of configurations from MD, it is necessary to make an initial assessment of the quality of the embedding methods, as well as the suitability of the structural models to which we apply the embedding methods. In the first case, the most straightforward evaluation comes from comparing how our embedding models (EM1 and EM2) can reproduce the reference supermolecular model (SM); in the second case, we shall be interested on characterizing the long-range interaction between the halogenated species and its environment.

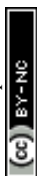
As discussed in sec. 4.1, the ZORA/SAOP model is sufficiently systematic to allow us to compare the behavior of both embedded HCl and  $\text{Cl}^-$  by using orbital energies as proxies for the extent to which our embedding models (EM1 and EM2) can reproduce the reference supermolecular model (SM) for the different core edges. From prior work<sup>55,62,105</sup> we expect the need for subsystem DFT calculations (in which both subsystem densities are optimized via freeze-thaw cycles) for  $\text{Cl}^-$ , whereas for the neutral HCl FDE calculations (in which the electron density for the environment—the ice surface—is constructed in the absence of the halogen species) relaxation of the environment would be less of an issue<sup>55,106</sup>, and therefore little should be gained from subsystem DFT calculations.

In figure 2 we present the comparison, for a selected snapshot, of the different models and how the core binding energies vary

as the number of water molecules is included in the Freeze-thaw procedure (for EM1 and EM2 we relax at most the 50 and 40 water molecules nearest to the halogen, respectively). The first important difference between the HCl and  $\text{Cl}^-$  systems is that, as expected, for  $\text{Cl}^-$  there is a much more important change between FDE calculations (zero relaxed water molecules) and subsystem DFT (at least one relaxed water molecule) ones, with FDE calculations with the EM1 model showing discrepancies of around 1.3 eV from SM for all edges.

By adding the six nearest water molecule to the active subsystem in EM2, we observe the difference to SM for FDE is reduced to around 0.8 eV. When 10 water molecules are relaxed, the subsystem DFT calculations with EM1 yield roughly the same results as FDE ones for EM2, and from this point onwards both EM1 and EM2 subsystem DFT calculations yield binding energies that differ by around 0.1 eV at most. It is important to note that even after 50 (40) water molecules relaxed, EM1 and EM2 still underestimate the SM binding energies by about 0.5 eV for the L edges and 0.6 eV for the K edge.

For HCl, on the other hand, we see very little improvement over FDE for the subsystem DFT calculations: varying the number of relaxed water molecules from zero to 40 or 50 introduces variations of at most 0.1-0.2 eV. This is in line with our expectation that FDE already provides a rather good representation of the environment for neutral subsystems. In contrast to  $\text{Cl}^-$ , we observe that EM2 shows slightly worse agreement to SM than EM1, and that for the L edges the embedding approaches are about 0.1-0.2 eV closer to the reference SM binding energies than for the K edge. In spite of that, embedded models still show a rather good agreement to SM, with EM2 typically underestimating the SM binding energies by 0.3 eV for the L edges (and 0.5 eV for the K edge), whereas EM1 reproduces SM binding energies nearly exactly for the L edge while underestimating the K edge binding energies by



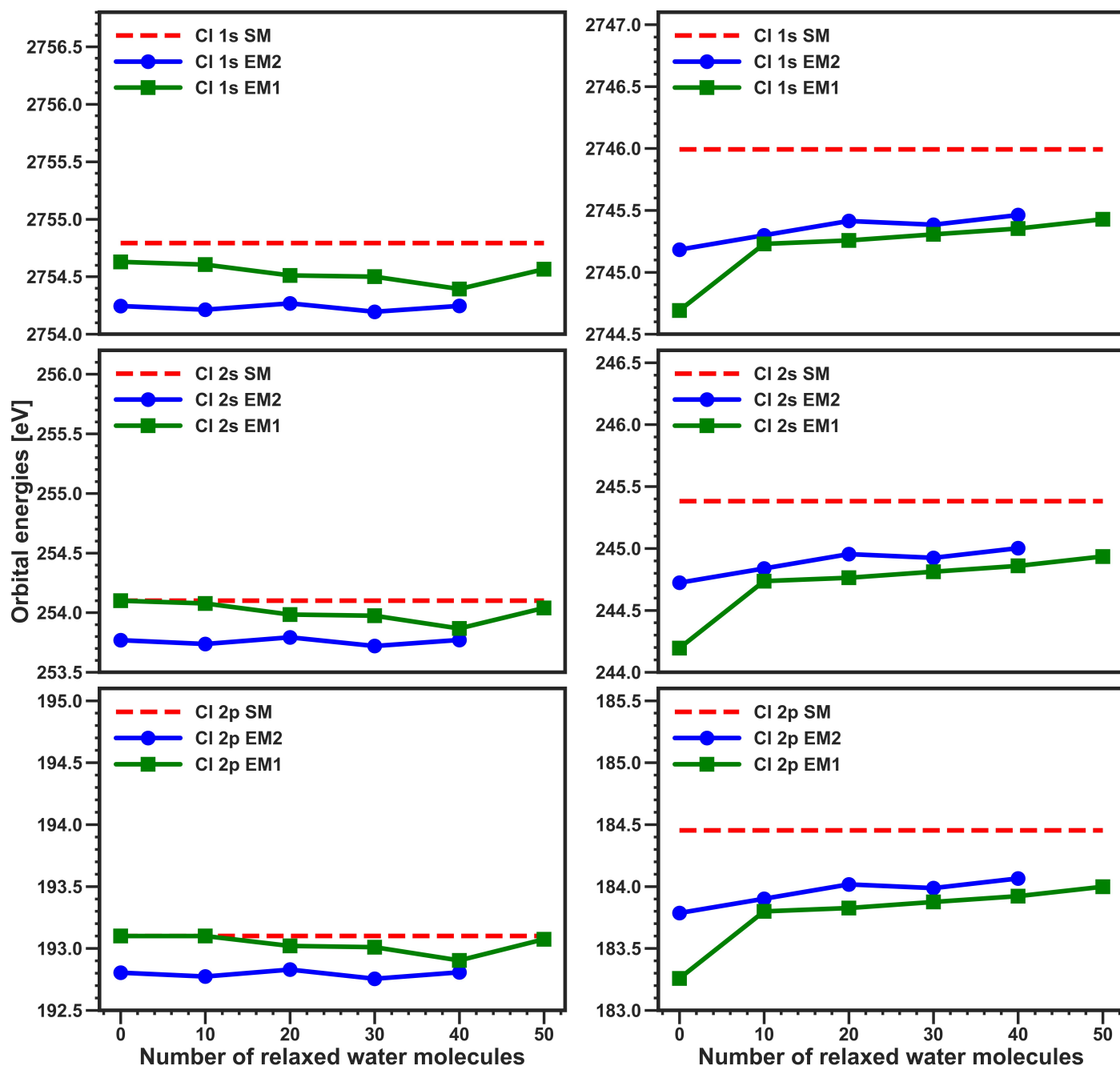


Fig. 2 Variation of the approximate K, L<sub>1</sub> and L<sub>2,3</sub> core binding energies of chlorine in HCl (left) and Cl<sup>-</sup> (right), obtained from scalar ZORA calculations with the SAOP model potential, with respect to the number of water molecules whose density is relaxed (in the ground state) via freeze-thaw cycles, for models EM1 (squares) and EM2 (circles). For comparison, the corresponding orbital energies obtained for model SM are provided as a reference (dashed line). The L<sub>2,3</sub> values are not split as calculations do not include spin-orbit coupling.





around 0.2 eV

Taken together, these results make us confident that, first and foremost, embedded models (and consequently, the underlying embedding potential) are indeed capable of reproducing SM calculations, and to do so in a manner that is roughly uniform for the K and L edges alike.

Second, that subsystem DFT derived embedding potentials introduce errors (due to the limited accuracy of the approximate kinetic energy density functionals employed to calculate the non-additive kinetic energy contributions<sup>48</sup>) that should result in small but non-negligible (0.5 eV or lower) underestimation of SM DFT binding energies.

In view of using them in CC-in-DFT calculations, the EM2 model has a significant disadvantage in that the active subsystem is significantly larger than EM1, and given the small difference in performance between the two, here we have opted to focus from now on on the EM1 model, and employing the EM2<sup>v</sup> model (that contains only one water molecule in the active subsystem instead of six) whenever assessing the suitability of the EM1 model in CC-in-DFT calculations.

In figure 3 we employ the EM1 and EM2<sup>v</sup> models, again for a single snapshot (and therefore disregard temperature effects introduced by considering several snapshots, as it will be done in the following), to verify the effect of long-range interactions between the halogens and the ice, through the truncation of the size of the water environment in the CC-in-DFT calculations.

We observe that for HCl there is no discernible difference between the embedded models, and that long-range effects seem to represent relatively small (0.2 eV) contributions, that are roughly uniform for the different edges, and tend to lower the core binding energies. Interestingly, the plots seem to indicate that long-range effects start to kick in after more than 100 water molecules have been taken into account. For Cl<sup>-</sup>, the situation is qualitatively slightly different, since we see a non-negligible difference between the EM1 and EM2<sup>v</sup> CC-in-DFT results, with the latter showing binding energies typically 0.2 eV lower than the former. That said, there appears to be a small decrease in binding energies between 100 and 150 water molecules (0.2 eV), as seen for HCl. We also note that irrespective of the model (EM1 and EM2<sup>v</sup>), the splitting between L<sub>2</sub> and L<sub>3</sub> edges remains around 1.6 eV.

Further evidence of the relative insensitivity of the results to the size of the cluster beyond 50 water molecules is given by the analysis of the system's dipole moment as a function of number of water molecules, as shown in table S4 in the supplementary materials, in that there are no significant changes in dipole for sufficiently large system.

From these results, we consider that employing the EM1 model is still advantageous from a computational point of view, since we consider that its smaller computational cost offsets the relatively modest improvement brought about by explicitly considering a water molecule in the active subsystem. Furthermore, due to the small changes upon considering a much larger environment, for the following we shall only consider models containing the halogen system and 50 water molecules.

### 4.3 Configurational averaging : ice and droplet models

Having established above that the EM1-based model containing 50 water molecules provides a very good balance between the ability to faithfully reproduce the reference calculations and the computational cost associated with CC-in-DFT calculations, we now turn to a discussion of the effects of the structural model for the environment and of the temperature, both associated with considering snapshots from classical molecular dynamics simulations. Table 2 summarizes our results.

Starting with the chloride ion in a droplet we observe that our calculated triple-zeta quality binding energies (BE(A), calculated from the average of binding energies over 100 snapshots of a simulation at 300K) show a slightly larger shift with respect to the gas-phase value ( $\Delta^s$ BE(A)) for the K edge (around 6.7 eV) than for the L edges (around 5.9 eV), which is consistent with the picture from our analysis of the single snapshot ZORA/SAOP results in section 4.2. We have not carried out quadruple zeta calculations for this system, due to the fact that the CBS corrections to the triple-zeta values for the chloride-ice surface (see supplementary information), at least for the L edges, are rather similar to the ones obtained for the gas-phase system. As such, we have decided to apply the gas-phase corrections for both K and L edges to the droplet system, given that for valence ionizations<sup>62</sup> CBS corrections from gas-phase or from averaged droplet binding energies were essentially the same.

A comparison of the droplet CBS-corrected values to the experimental results of Pelimanni *et al.*<sup>107</sup>, which have measured the L<sub>2</sub> and L<sub>3</sub> edges for KCl solutions at different concentrations and at somewhat higher temperatures (nozzle temperature of 373K), shows our results are in good agreement with experiment, as our results overestimate experiment by almost exactly 1 eV for each edge.

Our 2p spin-orbit splitting is consistent with that of experiment, at around 1.6 eV, a value that is close to the one seen in the gas phase (roughly the same differences are found in comparison to the experimental results of Partanen *et al.*<sup>108</sup>). There are much more significant discrepancies between our simulations and the experimental results of Kong *et al.*<sup>26</sup> obtained at somewhat lower temperatures (253K), not only in terms of the binding energy values (which are around 4 eV lower than our results) but of the 2p spin-orbit splitting (2.1 eV), which is 0.5 eV larger than both our simulation and other experimental results<sup>107,108</sup>.

The discrepancy between our results and those of Kong *et al.* could be due to temperature effects, since encapsulation of the halogens is driven by the temperature induced surface disorder, though the role of other parameters such as differences in calibration in the BE scale, cannot be dismissed out of hand. We note that Kong *et al.* have used as internal reference the oxygen K edge, but since obtaining such data is beyond the scope of this work (as it would require the construction and validation of new embedding models in order to carry out CC-in-DFT calculations on water molecules), we provide in Table 2 results for the valence band of water for the droplet model, obtained from SAOP orbital energies by Bouchafrat *et al.*<sup>62</sup>.

As discussed in the original work, the SAOP valence band of



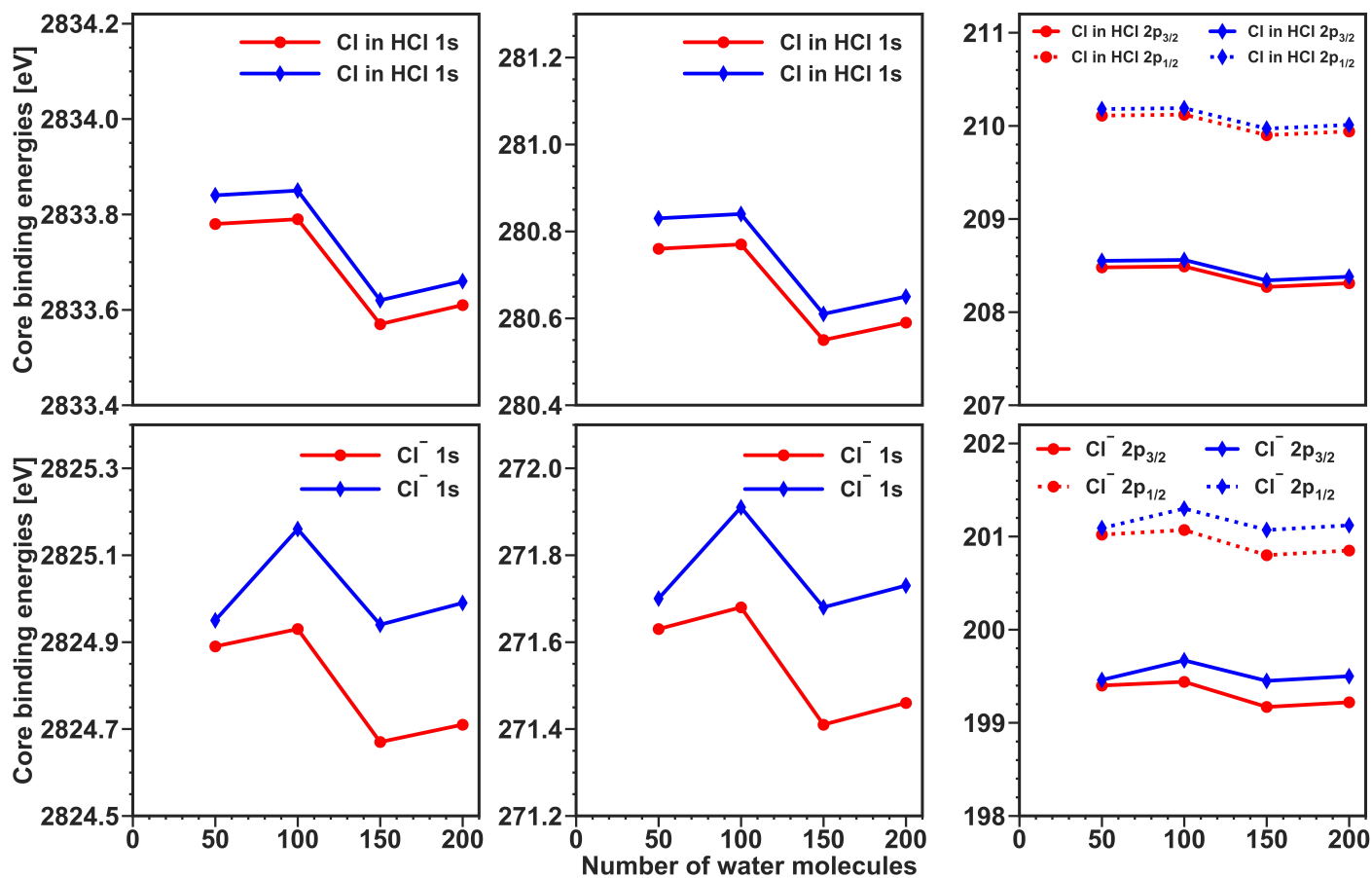


Fig. 3 CC-in-DFT K, L<sub>1</sub>, L<sub>2</sub> and L<sub>3</sub> triple-zeta binding energies of HCl and Cl<sup>-</sup> adsorbed on ice surfaces for a single snapshot, as a function of the number of water molecules in the environment (in addition to the 50 molecules nearest to the halogen system that are always taken into account). Blue lines represent the system with only the halogen species in the active subsystem, and red lines active systems containing halogen species and one explicit water molecule.



water is quite consistent with the available experimental results, with the 1 eV underestimation of experimental values having to do with the finite size of the water droplet Bouchafrat *et al.*. The good agreement between our theoretical values and the experimental values of Pelimanni *et al.* for the valence band of water, but also for the valence band of chloride, make us confident in the reliability of our embedded models and the CC-in-DFT protocol for core edges.

Considering now the chloride ion at the ice surface, our calculated triple-zeta quality binding energies (BE(B), calculated from the average of binding energies over 25 snapshots obtained by a combination of MD simulations at 210K for HCl, followed by an optimisation of the chlorine atom position) show similar behavior to that of droplets with respect to the free ion, with  $\Delta^s\text{BE(B)}$  values which are nearly the same for the K and L edges, again in line with the picture that the embedding potentials for these calculations affect the K and L edges in a roughly homogeneous manner. Here, however, we have a smaller shift for the K edge (3.53 eV) than for the L edges, and for the latter the differences between  $L_1$ ,  $L_2$  and  $L_3$  are of the order of 0.1 eV, that is, an order of magnitude more than for the water droplet.

In qualitative terms, our calculations reproduce well the trend of decreasing binding energies when going from solution–represented by the experimental results of Pelimanni *et al.*<sup>107</sup>–to the surface–represented by the experimental results of Parent *et al.*<sup>4</sup> for lower temperatures and Kong *et al.*<sup>26</sup> for higher temperatures (though for the latter, a near equivalence between the reported chloride binding energies for NaCl solution and ice surface 253K would indicate that chloride does behave as a free ion in both systems).

Quantitatively, our CBS corrected triple zeta results show differences of the order of 4 eV for the  $L_2$  and  $L_3$  edges with respect to the experimental results of Parent *et al.*<sup>4</sup>, obtained at 90K. At the same time the water valence band for the theoretical model for the ice surface is in good agreement with the same low-temperature experiment, with discrepancies of around 0.6 eV for the  $1b_1$  band. For the  $3a_1$  band discrepancies are of about 2 eV, but experimentally that is a broader band and therefore more difficult to provide an unambiguous comparison between theory and experiment. The difference of performance of our models for the ice surface valence binding energies and the core chloride binding energies could be an indication of the importance of temperature effects, that cannot be properly accounted for in our models since we only have data for 210K.

The discrepancies for core BEs are smaller with respect to the results of Kong *et al.*<sup>26</sup>, measured at 253K (and thus closer to the simulation conditions), but our values still overestimate the experimental results by 1.77 eV and 2.31 eV for the  $L_2$  and  $L_3$  edges, respectively. This is larger than the differences we observe for droplets between theory and the results of Pelimanni *et al.*<sup>107</sup>, but somewhat smaller than the differences between our droplet model and the results for NaCl solution from Kong *et al.*<sup>26</sup>. In our view, taken together the results for  $\text{Cl}^-$  on ice and water droplets seem to indicate a fairly systematic difference between our theoretical models and the experimental results of Kong *et al.*

On the simulation side, there is an important difference be-

tween the chloride-ice system with respect to the droplets, which is linked to the process and quality of the sampled structures, since sampling is intimately connected to the description of temperature effects. By inspecting figure 4, in which we show the K and L edge binding energies obtained for each snapshot around the mean value presented on table 2, we see narrow distributions around the mean for the droplet system for all edges considered (within envelopes of around 1 to 1.5 eV). For the chloride-ice system, the distributions are much wider, and of around 3 eV for the L edges, and almost 10 eV for the K edge.

This difference is in part expected, since in the droplet model the chloride ion is always completely surrounded by water molecules—and therefore one can consider that on average the ion has always a fairly constant degree of interaction with its environment—whereas for the ice model, the amount of water molecules with which the ion interacts greatly depends upon how much it has penetrated into the QLL. We speculate that, in our case, the sampled structures place the chloride ion deeper than it would be on average, and with that our results could be overestimating the chloride-surface interaction and, consequently, yielding an artificial increase in core binding energies, due to the fact that waters do not relax when the ion is introduced.

It may also be that our configurations are not properly representing the local environment of the chloride ion, as probed by the spin-orbit splitting of the 2p, though here the current experimental and theoretical data, in our view, do not allow for any definitive conclusions. On the one hand, Kong *et al.*<sup>26</sup> obtained an experimental difference between the  $L_2$  and  $L_3$  edges of 2.1 eV. On the other hand, our calculated splitting for chloride-ice is of roughly 1.7 eV, a value consistent with splitting between  $L_2$  and  $L_3$  edges in the NaCl crystal (see table 1) and slightly larger than the roughly 1.6 eV we obtained for the gas-phase ion, our droplet results and the experimental work of Pelimanni *et al.*<sup>107</sup> find for the solvated ion. It is also interesting to note that the 2.1 eV splitting is much larger than the 1.3 eV splitting observed by Parent *et al.*<sup>4</sup>, that is closer to our results.

In the case of the K edge, a problem with adequate sampling could be the reason for our large overestimation of the experimental binding energies, since from figure 4 we see that K edge energies are extremely sensitive to the configuration. At this stage we lack a better classical polarizable force field that can represent both the water-chloride and water-water interactions in the ice QLL. Due to this, the question of whether (and if so, how) better sampling would affect the K edge remains an open question.

Unlike the two chloride systems discussed above, for HCl a straightforward use of the molecular dynamics simulations snapshot yields results which are essentially the same as those for the isolated molecule. This indicates that in these snapshots there are, in effect, all but residual interactions between HCl and the ice ( $\Delta^s\text{BE(A)}$  values are very small and around 0.15 eV for all edges). By inspecting the top of figure 5, this becomes quite clear: in spite of averaging over 25 snapshots, there is essentially no spread in binding energy values, which would otherwise be the case if there were stronger interactions with the surface. This is consistent with the findings of Woittequand *et al.*<sup>111</sup> that the adsorption energy of HCl on ice (of the order -0.2 eV) is quite small



Table 2 Mean values of CC-in-DFT chlorine core binding energies (BE, in eV) averaged over structures from CMD simulations for models with 50 water molecules, and the difference of BEs and those calculated for the gas phase ( $\Delta^g\text{BE}$ , in eV). The molecular structures correspond to (A) the original CMD snapshots for water droplets<sup>62</sup> and ice surfaces<sup>87</sup>; (B) optimization of the halogen system coordinates, keeping the ice surface constrained to the CMD structure; (C) optimization of the halogen system coordinates and four nearest neighbor waters, keeping the remaining of the ice surface constrained to the CMD structure. For the ice surface systems, calculations correspond a temperature of 210 K. For water droplets, calculations correspond a temperature of 300 K. We also provide theoretical results (scalar ZORA SAOP) for valence bands ( $3a_1$  and  $1b_1$ ) of water for the ice surface, and for completeness we also provide for the water droplets the CC-in-DFT chlorine 3p and SAOP water  $3a_1$  and  $1b_1$  BEs from Bouchafra *et al.*<sup>62</sup>, obtained for the same snapshots as the chlorine BEs. We compare these results to experimental results by Kong *et al.*<sup>26</sup> (253 K), Parent *et al.*<sup>4</sup> (90 K), Partanen *et al.*<sup>108</sup> (393-423 K), Pelimanni *et al.*<sup>107</sup> (373 K), Kurahashi *et al.*<sup>109</sup> (280 K) and Winter *et al.*<sup>110</sup>.

| System           | Environment         | ionisation        | triple-zeta results |                        |         |                        |         |                        | BE(CBS <sup>†</sup> ) | Experiment  |
|------------------|---------------------|-------------------|---------------------|------------------------|---------|------------------------|---------|------------------------|-----------------------|---|
|                  |                     |                   | BE(A)               | $\Delta^g\text{BE(A)}$ | BE(B)   | $\Delta^g\text{BE(B)}$ | BE(C)   | $\Delta^g\text{BE(C)}$ |                       |   |
| HCl              | ice                 | K                 | 2834.03             | 0.13                   | 2833.40 | -0.50                  | 2833.29 | -0.61                  | 2834.33               | 2817.6 <sup>26</sup>  |
|                  |                     | L <sub>1</sub>    | 280.84              | 0.15                   | 280.15  | -0.54                  | 280.04  | -0.65                  | 281.19                |   |
|                  |                     | L <sub>2</sub>    | 210.18              | 0.15                   | 209.49  | -0.53                  | 209.38  | -0.64                  | 210.60                | 204.9 <sup>26</sup><br>202.2 <sup>4</sup>   |
|                  |                     | L <sub>3</sub>    | 208.54              | 0.15                   | 207.86  | -0.53                  | 207.75  | -0.64                  | 209.04                | 202.8 <sup>26</sup><br>200.9 <sup>4</sup>   |
| Cl <sup>-</sup>  |                     | K                 |                     |                        | 2827.70 | 3.53                   |         |                        | 2828.79               | 2815.4 <sup>26</sup>  |
|                  |                     | L <sub>1</sub>    |                     |                        | 274.90  | 4.17                   |         |                        | 275.07                |   |
|                  |                     | L <sub>2</sub>    |                     |                        | 204.41  | 4.31                   |         |                        | 204.47                | 202.7 <sup>26</sup><br>199.6 <sup>4</sup>   |
|                  |                     | L <sub>3</sub>    |                     |                        | 202.70  | 4.23                   |         |                        | 202.91                | 200.6 <sup>26</sup><br>198.3 <sup>4</sup>   |
|                  |                     | 3p                |                     |                        |         |                        |         |                        |                       | 10 <sup>4</sup>   |
| Cl <sup>-</sup>  | droplet             | K                 | 2829.97             | 6.66                   |         |                        |         |                        | 2829.93               |   |
|                  |                     | L <sub>1</sub>    | 276.63              | 5.90                   |         |                        |         |                        | 276.74                |   |
|                  |                     | L <sub>2</sub>    | 205.99              | 5.89                   |         |                        |         |                        | 206.19                | 205.0 <sup>107</sup><br>205.0 <sup>108</sup><br>202.7 <sup>26</sup>                         |
|                  |                     | L <sub>3</sub>    | 204.36              | 5.89                   |         |                        |         |                        | 204.55                | 203.4 <sup>107</sup><br>203.4 <sup>108</sup><br>200.6 <sup>26</sup>                         |
|                  |                     | 3p <sub>1/2</sub> | 9.9 <sup>62</sup>   |                        |         |                        |         |                        | 10.1 <sup>62</sup>    |   |
|                  |                     | 3p <sub>3/2</sub> | 9.7 <sup>62</sup>   |                        |         |                        |         |                        | 9.9 <sup>62</sup>     | 9.8 <sup>107</sup><br>9.5 <sup>109</sup><br>9.6 <sup>110</sup>                              |
| H <sub>2</sub> O | Cl <sup>-</sup> ice | 3a <sub>1</sub>   | 16.1                |                        |         |                        |         |                        |                       | 13.7 <sup>4</sup>   |
|                  |                     | 1b <sub>1</sub>   | 12.6                |                        |         |                        |         |                        |                       | 12 <sup>4</sup>   |
|                  | droplet             | 3a <sub>1</sub>   | 12.5 <sup>62</sup>  |                        |         |                        |         |                        |                       | 13.76 <sup>108</sup><br>13.78 <sup>109</sup><br>13.50 <sup>110</sup>                        |
|                  |                     | 1b <sub>1</sub>   | 10.4 <sup>62</sup>  |                        |         |                        |         |                        |                       | 11.4 <sup>107</sup><br>11.41 <sup>108</sup><br>11.31 <sup>109</sup><br>11.16 <sup>110</sup> |



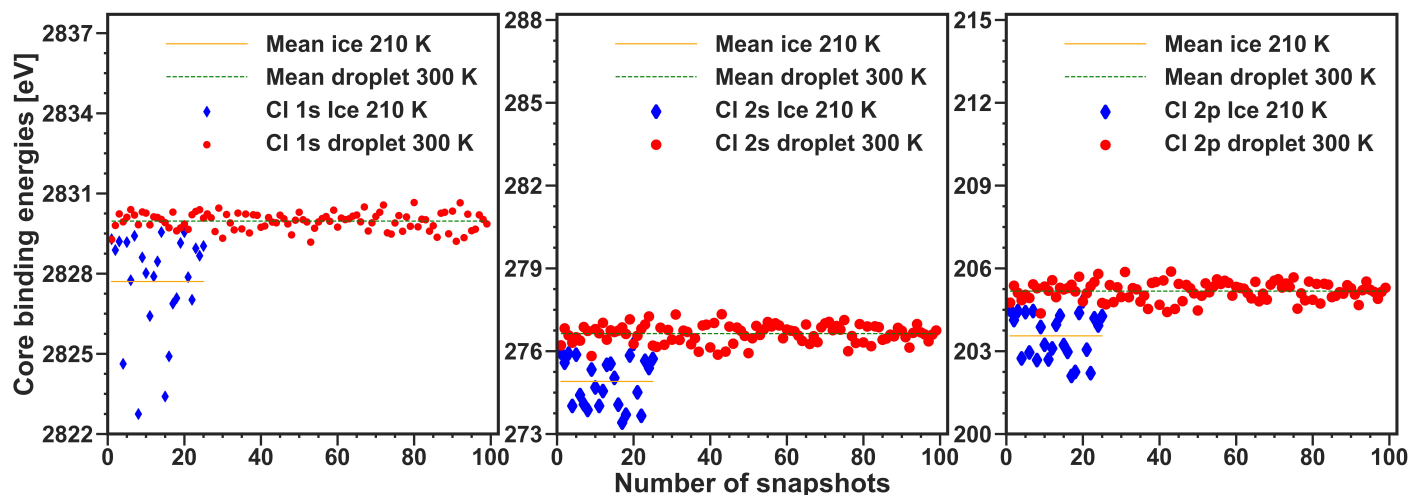


Fig. 4 CC-in-DFT K, L<sub>1</sub>, L<sub>2</sub> and L<sub>3</sub> triple-zeta binding energies of chloride adsorbed on ice surfaces at 210 K (results averaged over 25 snapshots) and in water droplets at 300 K (results averaged over 100 snapshots).

in absolute value.

If we take the snapshots as starting point for geometry optimizations of the HCl molecule, keeping the ice structure constrained to the original CMD configurations, we see small but non-negligible change in the binding energies (BE(B)) for the different edges, so that now instead of the slight increase of binding energies seen at first, we start to see a move towards lower binding energies ( $\Delta^s$ BE(B) of about -0.50 eV), that is, towards the experimental trend (HCl on ice binding energies being lower than gas-phase ones). Similarly to the chloride on ice system, there is a large spread in values (of around 2-3 eV for the K and L edges, see middle of figure 5).

Upon obtaining configurations in which we also optimize the waters nearest to the HCl molecule, we observe a further decrease in binding energies (BE(C)) that, though in the direction of experiment, is too small to bring our calculations to the same agreement with experiment as seen for the L edge of the chloride-ice system discussed above, and we see discrepancies of around 5.7-6 eV. The discrepancy between theory and experiment for HCl-ice K edge binding energies are also around 3 eV larger than for the chloride-ice system.

One possible issue with these simulations on ice is that, at 210K, the temperature of the simulations is somewhat lower than that of the experiment. This makes it worthwhile to explore the effect of the temperature on the ice structures, and see to which extent the changes would affect the binding energies.

To this end, we have carried out additional classical MD simulations for HCl under the same conditions as done for 210K, one at 235K and another at 250K, and selected a single snapshot of each to carry out exploratory CC-in-DFT calculations. In figures S1 and S2 we show respectively the structures for HCl and Cl<sup>-</sup> at the two temperatures.

From these figures, and keeping in mind the structures at 210K shown in figure 1, we see the progressive desorganization of the upper layers of the interface when passing from 210K to 235K (though the innermost two layers remain rather well structured), and a fairly substantial loss of structure going from 235K to 250K.

In spite of these significant changes in structure, between the different temperatures there are little changes to the binding energies, as can be seen from table S5 in the supplementary information. Even though we have only one structure, and therefore we cannot strictly speaking compare to the averaged results for 210K, the results in table S5 suggest nevertheless that temperature effects cannot play a major role in modifying the binding energies if the HCl molecule remains essentially bound (and with an internuclear distance not far from the gas-phase value), as we discuss in the following.

#### 4.4 A closer look on the HCl-water interaction

The contrast between chloride and HCl results, and the changes (albeit modest) in binding energies observed for HCl depending on the structural model for the HCl-ice surface interaction sites discussed above, call for a closer look at how the structural parameters affect the calculated binding energies, as shown in figure 6. Considering first the HCl internuclear distance (panel D), we see that for the original snapshots from CMD simulations of Woittequand *et al.*<sup>87</sup> (model A) one obtains essentially the same binding energies which, as discussed above, are nearly indistinguishable from the gas-phase ones.

Upon optimizing the HCl position while keeping the surface unchanged (model B), we see a significant change in that internuclear distances increase for all snapshots with respect to model A, to values between 1.28 and 1.38 Å; furthermore, we can identify three categories of points: those for internuclear distances around 1.28 Å, which are associated to larger core binding energies (right of the figure), those for internuclear distances between 1.32 and 1.36 Å, which are associated with lower core binding energies (left of the figure), and the third cluster for internuclear distances between 1.28 and 1.34 Å, but which exhibit roughly the core same binding energies (around 208 eV).

The optimization of the HCl and nearest water molecules (model C) accentuates somewhat the trend of increased internuclear distances in the region for lower core binding energies





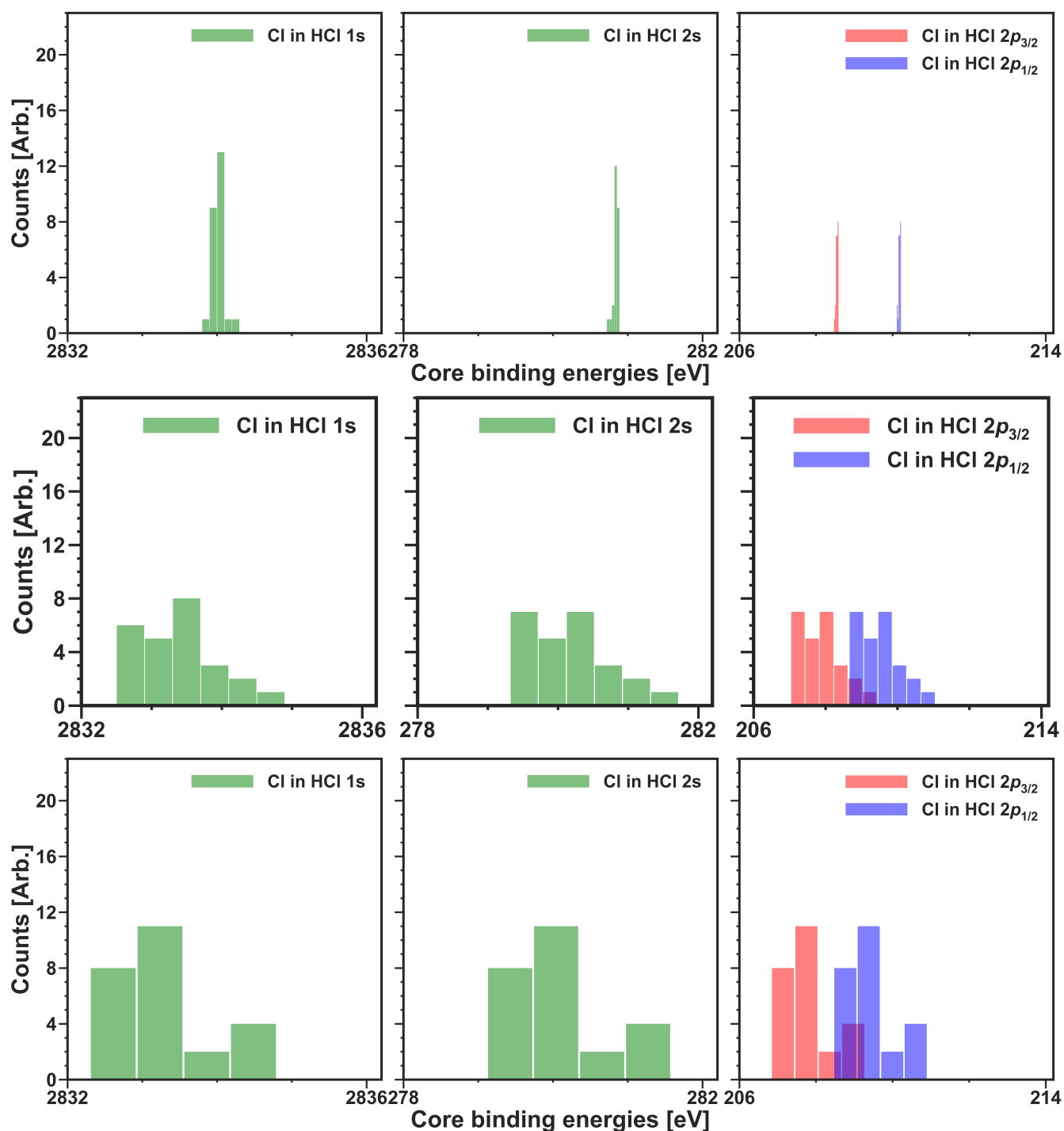


Fig. 5 CC-in-DFT chlorine K, L<sub>1</sub>, L<sub>2</sub> and L<sub>3</sub> triple-zeta binding energies for HCl adsorbed on ice surfaces at 210 K (results averaged over 25 snapshots) employing as structural models (A) the original CMD snapshots (top); (B) reoptimizing the HCl molecule while constraining the ice surface to retain the atomic positions of model A (middle); and (C) reoptimizing the HCl and four nearest water molecules, while constraining the rest of the ice surface to retain the atomic positions of model A (bottom)



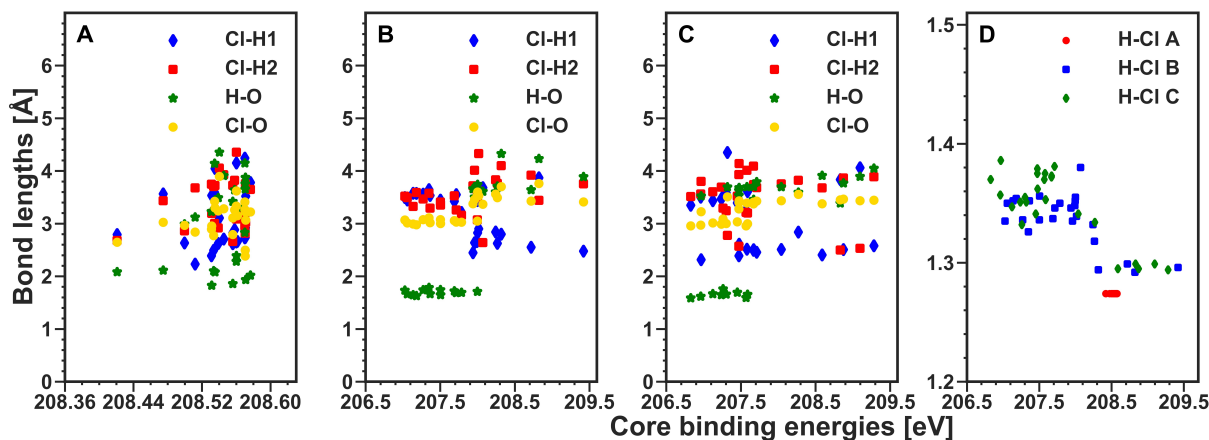


Fig. 6 Scalar ZORA chlorine 2p binding energies as a function of HCl-H<sub>2</sub>O intra- and inter-molecular distances of HCl adsorbed on ice surfaces at 210 K for the 25 snapshots, employing as structural models (A) the original CMD snapshots; (B) reoptimizing the HCl molecule while constraining the ice surface to retain the atomic positions of model A; and (C) reoptimizing the HCl and four nearest water molecules, while constraining the rest of the ice surface to retain the atomic positions of model A. In panel D the BEs with respect to the HCl bond lengths in models A, B and C are shown.

seen for model B. We observe more snapshots with internuclear distances larger than 1.36 Å, which are 1 to 2 eV lower than the core binding energies for model A (and we note that, contrary to the gas-phase results, relatively small changes in internuclear distance produce a significant shift of core binding energies). However, the average core binding energy only shows modest changes with respect to model A due to the fact that there remain several structures with core binding energies larger than 208.5 eV.

Apart from the H-Cl distance, we see significant changes in the distances between the hydrogen in HCl and the nearest oxygen atoms of the surface: while for model A, the large variation in this O-H distance does not significantly affect the binding energies, for models B and C we can distinguish two types of distances: longer ones (around 3.5 Å) for which the core binding energies are generally above 208 eV, and shorter ones (around 1.6 Å) associated with lower core binding energies.

Taken together, these observations suggest that the lowering of the core binding energies is closely connected to the concerted increase in the HCl internuclear distance and decrease of the oxygen surface atoms and the hydrogen of HCl, which would represent the initial stage of the (pre)dissociation of HCl mediated by the surface. Two such configurations can be more clearly visualized in figure 7, which depicts the spatial arrangement of HCl and its nearest four water molecules for two situations in which structural relaxation is taken into account. The figure also contains the core binding energies obtained for each microsolvated cluster. We observe that, already in such simplified models, binding energies are pretty sensitive to relatively small changes in structure. One can also identify in the figure cooperative effects coming from the elongation of certain O-H bonds in the water molecules as the HCl molecule gets closer (with the elongation of the H-Cl bond and the interaction of the hydrogen of HCl and the oxygen of the nearest water). A further investigation of the influence of predissociation of HCl would require more extensive CMD simulations for the surface, which are beyond the scope of this work.

To further explore this point, we have considered two additional models: (a) one in which a +1 point charge is placed at a

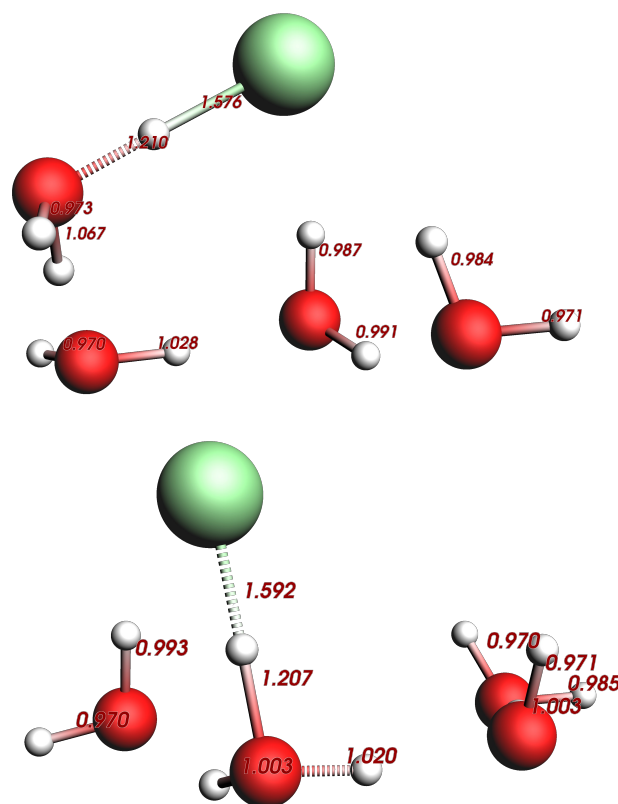


Fig. 7 Structures for a microsolvated HCl molecule originating from snapshots of model C, with nearest 4 waters shown. Internuclear distances (in Å) are indicated in the figure. Binding energies (in eV) for the [K, L<sub>1</sub>, L<sub>2</sub>, L<sub>3</sub>] edges, obtained with CC-in-DFT (for the microsolvated system) are respectively [2832.6, 279.3, 208.7, 207.1] (Top) and [2831.9, 278.7, 208.0, 206.4] (Bottom).



Table 3 Core binding energy shift ( $\Delta_{BE}$ , in eV) between  $\text{Cl}^-$  and the HCl systems as a function of the distance  $r$  (in Å) between the chlorine and the hydrogen atom, for (a) the electrostatic gas-phase model based on the pair of charges (chloride and a +1 point charge representing the hydrogen); and (b) the  $[\text{HCl}(\text{H}_2\text{O})_{50}]$  droplet model structure based on the  $[\text{Cl}(\text{H}_2\text{O})_{50}]^-$  droplet model discussed above, in which the additional hydrogen is found near the different heavy centers (chlorine, oxygens). See text and supplementary information for further details and approximate core electron binding energies.

| $r$    | Core Binding Energies |                |                                |                 |                |                                |
|--------|-----------------------|----------------|--------------------------------|-----------------|----------------|--------------------------------|
|        | droplet model         |                |                                | gas-phase model |                |                                |
|        | K                     | L <sub>1</sub> | L <sub>2</sub> /L <sub>3</sub> | K               | L <sub>1</sub> | L <sub>2</sub> /L <sub>3</sub> |
| 1.306  | 7.48                  | 7.53           | 7.69                           | 8.59            | 8.73           | 8.80                           |
| 2.559  | 3.05                  | 3.06           | 3.05                           | 5.70            | 5.71           | 5.75                           |
| 3.489  | 2.81                  | 2.81           | 2.82                           | 4.16            | 4.17           | 4.18                           |
| 5.690  | 2.32                  | 2.33           | 2.33                           | 2.53            | 2.53           | 2.54                           |
| 10.256 | 1.60                  | 1.64           | 1.64                           | 1.40            | 1.40           | 1.40                           |

given distance  $r$  from the chloride ion; and (b) one in which we employ one snapshot of the chloride droplet model to construct a HCl droplet model, by placing the added hydrogen atom near the chloride or a given oxygen (to simulate the  $\text{H}_3\text{O}^+$  species), and performing a constrained optimization (fixing all atoms but the hydrogens belonging to the same species the hydrogen has been attached to). The results for these models are found in table 3 and table S6 of the supplementary material.

Starting with the gas-phase model (a), we observe that at distances slightly larger than the gas-phase equilibrium ( $r = 1.306$  Å), the  $\Delta_{BE}$  values are rather close to the values for the molecular HCl system (8.6–8.8 eV vs roughly 10 eV in table 1). As  $r$  is increased to  $r = 2.559$  Å, distance already much larger than those sampled by our MD simulations and shown in figures 6 and 7, there's a significant drop in  $\Delta_{BE}$  to 5.7 eV, and then a relatively smooth decrease for larger distances. Interestingly, at around  $r = 5.690$  Å,  $\Delta_{BE}$  drops to around 2.5 eV, which is in the order of magnitude of the experimental chemical shift reported by Kong *et al.*<sup>26</sup> (2.2 eV), and also close to the value by Parent *et al.*<sup>4</sup> (2.6 eV). While not shown, we have investigated how far the +1 point charge would have to be in order for us to obtain the gas-phase  $\text{Cl}^-$  value, and at  $r = 100$  Å there are still small differences (around 0.1 eV).

For model (b) we have the same qualitative trend, but with an interesting difference: while the  $\Delta_{BE}$  value at  $r = 1.306$  Å is still relatively large (around 7.5 eV), comparing it with the gas-phase  $\Delta_{BE}$  value shown in 1, we can infer an environment effect of around 2.5 eV—which is already much larger than the effects shown in table 2, and in line with figures 6 and 7. For  $r = 2.559$  Å, due to the effect of the screening by the water molecules in the droplet (see figure S3 in the supplementary information), the  $\Delta_{BE}$  value (around 3 eV) is nearly 3 eV lower than the one from the gas-phase. Likewise, at  $r = 3.489$  Å  $\Delta_{BE}$  is already below 3 eV (and a little over 1 eV smaller than the gas-phase value).

For larger  $r$  values model (b) follows roughly the behavior of model (a), which may be due to the small size of the droplet. Contrary to the gas-phase model, our simulations do not allow us to probe much longer distances than 10 Å, but the relative small

difference to the anionic systems suggests in our view that the anionic systems would be indeed rather good models for diluted solutions. In subsequent work it will be interesting to investigate how screening will alter the rate of convergence towards the anionic result, now that polarizable force fields of similar accuracy to those employed here and that can handle counter-ions are starting to become available<sup>112</sup>.

Whatever the case, these results suggest that for a solvated system, chemical shifts compatible to those observed in experiment occur for values of  $r$  between 4 and 6 Å.

While these results only consider single structures, and therefore must be viewed as providing semi-quantitative evidence, they help to understand that the poor agreement with experiment for the simulations of HCl on ice shown in table 2 likely comes from an inadequate structural model, which assumes and retains a bound (molecular) picture for HCl, when it would seem that a more suitable situation resembles the existence of an ion pair, or some other intermediate situation. Unfortunately, we do not currently possess the adequate tools to further explore this problem, and in any case such an undertaking requires a dedicated study.

## 5 Conclusion

In this article we report the application of the relativistic CVS-EOM-IP-CCSD-in-DFT approach to obtain core electron binding energies for chlorinated species ( $\text{Cl}^-$  and HCl) for the air-ice interface which is of great interest for atmospheric chemistry and physics, as well as of  $\text{Cl}^-$  in aqueous solution (which allows us to differentiate isotropic and non isotropic solvation of the anion). In our coupled cluster calculations, we employ the molecular mean-field Dirac-Coulomb-Gaunt Hamiltonian, which accurately accounts for spin-same orbit and spin-other orbit interactions.

These calculations are based upon structural models considering, for both droplets and ice surfaces, the halides and the nearest 50 and 200 water molecules respectively. Based on these, embedding models in which all water molecules were treated at DFT level while the halide species were treated with coupled cluster have been assessed, and their relative accuracy verified against reference DFT calculations on the whole system. There, we have found that subsystem DFT calculations, in which both the halide and 50 nearest water molecules in the environment were relaxed in the presence of each other, were well-suited for both the neutral (HCl) and the charged ( $\text{Cl}^-$ ) subsystems, showing small and systematic errors for all edges.

The accuracy of our protocol has been shown for the water droplet case, for which we obtain L<sub>2</sub> and L<sub>3</sub> CBS-corrected core binding energies in very good agreement but overestimate the most recent experiments by Pelimanni *et al.*<sup>107</sup> in solution by around 1 eV for each edge, while obtaining nearly the same energy splitting (1.66 eV, due to spin-orbit coupling) as experiment (1.6 eV) between the L<sub>2</sub> and L<sub>3</sub> edges. We consider that remaining discrepancies are likely due, in part, to the different temperatures in which theoretical and experiment results have been obtained, and also due to the lack of corrections for higher excitations in our calculations.

Our theoretical values for the L edges of chloride on ice sur-



faces show systematic differences to the droplet model (binding energies roughly 1.7 eV smaller). They are also in fairly good agreement with experiment for the L<sub>2</sub> and L<sub>3</sub> edges, though with larger discrepancies than for water droplets, with theoretical values overestimating experiment by 1.8 and 2.3 eV respectively.

While that may partially arise from the shortcomings of our protocol to obtain structures for chloride (classical MD simulations on HCl followed by constrained geometry optimization of the chloride position, due to the lack of suitable force fields), it should be pointed out that our theoretical spin-orbit splitting between the L<sub>2</sub> and L<sub>3</sub> edges is consistent (1.56 eV) with that in the droplet model, with experiments in solution, and in the NaCl crystal. The experimental splitting on ice reported by Kong *et al.*<sup>26</sup>, on the other hand, is somewhat larger (2.1 eV), and merits to be further investigated from both a theoretical and experimental point of view.

For the K edge we observe that chloride binding energies on ice are 1.1 eV smaller than those in the droplet model, qualitatively in line with the results for the L edges and with the overall experimental trend of lowering the core binding energies when going from a solution to a surface for chloride. However, theory strongly overestimates (by 13.4 eV) the experimental K edge binding energies. We cannot at this stage pinpoint the factor(s) driving this discrepancy, though we note that in contrast to the droplet model, there are very significant variations in the K edge binding energies depending on the snapshot used. In our view, this calls for additional simulations once a better force field for the MD simulations is available, in order to confirm whether or not configuration sampling is a significant source of bias to the theoretical binding energies. However, at this point we are not equipped to carry out such simulations.

We also report estimates for the valence binding energy of water in the ice surfaces, which are in good agreement with experiments at low temperatures. For that we have followed the procedure previously employed for obtaining the water valence band for water in droplets, in which the use of the SAOP model potential to describe the environment yields valence binding energies as a by-product of the setup of the embedded coupled cluster calculations at no additional cost.

In contrast to the chloride results, our results for molecular HCl on ice surfaces show poor quantitative agreement with experiment. We obtain core binding energies that are significantly higher than experimental ones (nearly 6 eV for the L edges, and nearly 17 eV for the K edge), due to the fact that there are almost negligible environmental effects and, therefore, the results are essentially the same as those for gas-phase HCl.

We have assessed to a limited extent the importance of temperature effects on the binding energies, by comparing the CC-in-DFT core binding energies based on single structures taken from classical MD simulations of HCl at higher temperatures (235K and 250K) to those averaged over 25 snapshots at 210K. We found that in spite of the important desorganization of the ice structure as the temperature increases, there are no significant changes in binding energies for Cl in HCl.

From an analysis of small microsolvated clusters, and of two models to capture the variation of the core binding energies as

a function of the H-Cl distance (a simple gas-phase one, and droplet models for HCl constructed from the chloride droplet model shown to be reliable), we were able to trace the root cause of these lack of sensitivity to the environment to the inability of our structural models to account for the (pre)dissociation of HCl molecule upon interaction with water molecules around it, resulting in the transfer of the proton to the first and second solvation layers.

In spite of their limitations, and notably the lack of any configurational averaging, the gas-phase electrostatic and HCl droplet models provide evidence that, in order to be consistent with the experimentally observed chemical shifts between chloride and HCl of about 2.2-2.6 eV, the proton in the latter should likely be distant from the chlorine atom by around 5-6 Å, a situation which is compatible with it participating in the hydrogen bond network. As such, a perhaps more suitable description for HCl on ice is that of an ion pair rather than of a molecular species. In order to properly characterize this system it appears to us that a central ingredient would be a MD approach that could (at least) approximately account for such ion-pair formation, and ideally describe the exchange of hydrogens between different water molecules, but such studies fall outside the scope of this work.

## Conflicts of interest

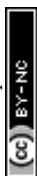
There are no conflicts to declare.

## Acknowledgements

We acknowledge support by the French government through the Program "Investissement d'avenir" through the Labex CaPPA (contract ANR-11-LABX-0005-01), the I-SITE ULNE project OVERSEE and MESONM International Associated Laboratory (LAI) (contract ANR-16-IDEX-0004), the Franco-German project CompRIXS (Agence nationale de la recherche ANR-19-CE29-0019, Deutsche Forschungsgemeinschaft JA 2329/6-1), CPER CLIMIBIO (European Regional Development Fund, Hauts de France council, French Ministry of Higher Education and Research) and French national supercomputing facilities (grants DARI A0070801859, A0090801859 and A0110801859).

## Notes and references

- 1 S. Solomon, *Reviews of Geophysics*, 1999, **37**, 275–316.
- 2 J. P. Abbatt, *Chemical reviews*, 2003, **103**, 4783–4800.
- 3 T. F. Kahan and D. Donaldson, *Environmental science & technology*, 2010, **44**, 3819–3824.
- 4 P. Parent, J. Lasne, G. Marcotte and C. Laffon, *Physical Chemistry Chemical Physics*, 2011, **13**, 7142–7148.
- 5 A. Křepelová, T. Huthwelker, H. Bluhm and M. Ammann, *ChemPhysChem*, 2010, **11**, 3859–3866.
- 6 J. T. Newberg and H. Bluhm, *Physical Chemistry Chemical Physics*, 2015, **17**, 23554–23558.
- 7 J. P. D. Abbatt, J. L. Thomas, K. Abrahamsson, C. Boxe, A. Granfors, A. E. Jones, M. D. King, A. Saiz-Lopez, P. B. Shepson, J. Sodeau, D. W. Toohey, C. Toubin, R. von Glasow, S. N. Wren and X. Yang, *Atmospheric Chemistry and Physics*, 2012, **12**, 6237–6271.





- 8 J. D. Smith, R. J. Saykally and P. L. Geissler, *Journal of the American Chemical Society*, 2007, **129**, 13847–13856.
- 9 P. N. Perera, B. Browder and D. Ben-Amotz, *The Journal of Physical Chemistry B*, 2009, **113**, 1805–1809.
- 10 S. Hüfner, *Photoelectron spectroscopy: principles and applications*, Springer Science & Business Media, 2013.
- 11 H. Tissot, G. Olivieri, J.-J. Gallet, F. Bournel, M. G. Silly, F. Sirotti and F. Rochet, *The Journal of Physical Chemistry C*, 2015, **119**, 9253–9259.
- 12 I. Gladich, S. Chen, M. Vazdar, A. Boucly, H. Yang, M. Ammann and L. Artiglia, *The journal of physical chemistry letters*, 2020, **11**, 3422–3429.
- 13 L. Partanen, M.-H. Mikkela, M. Huttula, M. Tchapyguine, C. Zhang, T. Andersson and O. Björneholm, *The Journal of chemical physics*, 2013, **138**, 044301.
- 14 S. K. Eriksson, I. Josefsson, N. Ottosson, G. Ohrwall, O. Björneholm, H. Siegbahn, A. Hagfeldt, M. Odelius and H. Rensmo, *The Journal of Physical Chemistry B*, 2014, **118**, 3164–3174.
- 15 K. H. Kim, J. H. Lee, J. Kim, S. Nozawa, T. Sato, A. Tomita, K. Ichiyana, H. Ki, J. Kim, S.-i. Adachi *et al.*, *Physical review letters*, 2013, **110**, 165505.
- 16 S. Yamamoto, T. Kendelewicz, J. T. Newberg, G. Ketteler, D. E. Starr, E. R. Mysak, K. J. Andersson, H. Ogasawara, H. Bluhm, M. Salmeron *et al.*, *The Journal of Physical Chemistry C*, 2010, **114**, 2256–2266.
- 17 J. T. Newberg, D. E. Starr, S. Yamamoto, S. Kaya, T. Kendelewicz, E. R. Mysak, S. Porsgaard, M. B. Salmeron, G. E. Brown Jr, A. Nilsson *et al.*, *Surface Science*, 2011, **605**, 89–94.
- 18 P. D. Dau, J. Su, H.-T. Liu, J.-B. Liu, D.-L. Huang, J. Li and L.-S. Wang, *Chemical Science*, 2012, **3**, 1137–1146.
- 19 A. Krepelova, T. Bartels-Rausch, M. A. Brown, H. Bluhm and M. Ammann, *The Journal of Physical Chemistry A*, 2013, **117**, 401–409.
- 20 G. Olivieri, A. Goel, A. Kleibert, D. Cvetko and M. A. Brown, *Physical Chemistry Chemical Physics*, 2016, **18**, 29506–29515.
- 21 A. P. Gaiduk, M. Govoni, R. Seidel, J. H. Skone, B. Winter and G. Galli, *Journal of the American Chemical Society*, 2016, **138**, 6912–6915.
- 22 D. E. Starr, M. Favaro, F. F. Abdi, H. Bluhm, E. J. Crumlin and R. van de Krol, *Journal of Electron Spectroscopy and Related Phenomena*, 2017, **221**, 106–115.
- 23 N. K. Jena, I. Josefsson, S. K. Eriksson, A. Hagfeldt, H. Siegbahn, O. Björneholm, H. Rensmo and M. Odelius, *Chemistry—A European Journal*, 2015, **21**, 4049–4055.
- 24 V. F. McNeill, T. Loerting, F. M. Geiger, B. L. Trout and M. J. Molina, *Proceedings of the National Academy of Sciences*, 2006, **103**, 9422–9427.
- 25 V. F. McNeill, F. M. Geiger, T. Loerting, B. L. Trout, L. T. Molina and M. J. Molina, *The Journal of Physical Chemistry A*, 2007, **111**, 6274–6284.
- 26 X. Kong, A. Waldner, F. Orlando, L. Artiglia, T. Huthwelker, M. Ammann and T. Bartels-Rausch, *The journal of physical chemistry letters*, 2017, **8**, 4757–4762.
- 27 K. Bolton, *Journal of Molecular Structure: THEOCHEM*, 2003, **632**, 145–156.
- 28 S.-C. Park and H. Kang, *The Journal of Physical Chemistry B*, 2005, **109**, 5124–5132.
- 29 P. Parent and C. Laffon, *The Journal of Physical Chemistry B*, 2005, **109**, 1547–1553.
- 30 J. Stanton and R. Bartlett, *J. Chem. Phys.*, 1993, **98**, 7029–7039.
- 31 K. Sneskov and O. Christiansen, *WIREs Comput. Mol. Sci.*, 2012, **2**, 566–584.
- 32 R. J. Bartlett, *WIREs Comput Mol Sci*, 2012, **2**, 126–138.
- 33 A. I. Krylov, *Ann. Rev. Phys. Chem.*, 2008, **59**, 433–462.
- 34 O. Christiansen, P. Jørgensen and C. Hättig, *Int. J. Quantum Chem.*, 1998, **98**, 1.
- 35 H. Koch and P. Jørgensen, *J. Chem. Phys.*, 1990, **93**, 3333–3344.
- 36 S. Coriani, F. Pawłowski, J. Olsen and P. Jørgensen, *J. Chem. Phys.*, 2016, **144**, 024102.
- 37 L. S. Cederbaum, W. Domcke and J. Schirmer, *Phys. Rev. A: At. Mol. Opt. Phys.*, 1980, **22**, 206–222.
- 38 S. Coriani and H. Koch, *J. Chem. Phys.*, 2015, **143**, 181103.
- 39 S. Coriani and H. Koch, *J. Chem. Phys.*, 2016, **145**, 149901.
- 40 A. Sadybekov and A. I. Krylov, *J. Chem. Phys.*, 2017, **147**, 014107.
- 41 M. L. Vidal, X. Feng, E. Epifanovsky, A. I. Krylov and S. Coriani, *J. Chem. Theory Comput.*, 2019, **15**, 3117–3133.
- 42 L. Halbert, M. L. Vidal, A. Shee, S. Coriani and A. S. P. Gomes, *J. Chem. Theory Comput.*, 2021, **17**, 3583.
- 43 T. Saue, *ChemPhysChem*, 2011, **12**, 3077–3094.
- 44 J. Sikkema, L. Visscher, T. Saue and M. Iliaš, *J. Chem. Phys.*, 2009, **131**, 124116.
- 45 X. Zheng and L. Cheng, *Journal of chemical theory and computation*, 2019, **15**, 4945–4955.
- 46 J. Lee, D. W. Small and M. Head-Gordon, *The Journal of chemical physics*, 2019, **151**, 214103.
- 47 A. S. P. Gomes and C. R. Jacob, *Annual Reports Section "C" (Physical Chemistry)*, 2012, **108**, 222–277.
- 48 C. R. Jacob and J. Neugebauer, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2014, **4**, 325–362.
- 49 T. A. Wesolowski, *Physical Review A*, 2008, **77**, 012504.
- 50 F. Libisch, C. Huang and E. A. Carter, *Accounts of chemical research*, 2014, **47**, 2768–2775.
- 51 T. A. Wesolowski, S. Shedge and X. Zhou, *Chemical reviews*, 2015, **115**, 5891–5928.
- 52 Q. Sun and G. K.-L. Chan, *Accounts of chemical research*, 2016, **49**, 2705–2712.
- 53 D. V. Chulhai and J. D. Goodpaster, *Journal of chemical theory and computation*, 2018, **14**, 1928–1942.
- 54 B. Hegely, P. R. Nagy and M. Kallay, *Journal of chemical theory and computation*, 2018, **14**, 4600–4615.
- 55 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Physical Chemistry Chemical Physics*, 2008, **10**, 5353–5362.





- 56 S. J. Bennie, B. F. Curchod, F. R. Manby and D. R. Glowacki, *The journal of physical chemistry letters*, 2017, **8**, 5559–5565.
- 57 S. Prager, A. Zech, T. A. Wesolowski and A. Dreuw, *Journal of chemical theory and computation*, 2017, **13**, 4711–4725.
- 58 S. Hofener and L. Visscher, *Journal of chemical theory and computation*, 2016, **12**, 549–557.
- 59 D. J. Coughtrie, R. Giereth, D. Kats, H.-J. Werner and A. Kohn, *Journal of Chemical Theory and Computation*, 2018, **14**, 693–709.
- 60 S. Höfener, A. S. P. Gomes and L. Visscher, *The Journal of chemical physics*, 2013, **139**, 104106.
- 61 C. Daday, C. König, J. Neugebauer and C. Filippi, *ChemPhysChem*, 2014, **15**, 3205–3217.
- 62 Y. Bouchafra, A. Shee, F. Réal, V. Vallet and A. S. P. Gomes, *Physical review letters*, 2018, **121**, 266001.
- 63 A. Sadybekov and A. I. Krylov, *The Journal of chemical physics*, 2017, **147**, 014107.
- 64 S. Höfener, A. Severo Pereira Gomes and L. Visscher, *The Journal of chemical physics*, 2012, **136**, 044104.
- 65 V. Parravicini and T.-C. Jagau, *Molecular Physics*, 2021, e1943029.
- 66 T. A. Wesolowski and A. Warshel, *The Journal of Physical Chemistry*, 1993, **97**, 8050–8053.
- 67 T. Dresselhaus, J. Neugebauer, S. Knecht, S. Keller, Y. Ma and M. Reiher, *The Journal of chemical physics*, 2015, **142**, 044111.
- 68 C. F. Guerra, J. Snijders, G. t. te Velde and E. J. Baerends, *Theoretical Chemistry Accounts*, 1998, **99**, 391–403.
- 69 C. R. Jacob, J. Neugebauer and L. Visscher, *Journal of computational chemistry*, 2008, **29**, 1011–1018.
- 70 *ADF2017, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands*, <https://www.scm.com>.
- 71 E. v. Lenthe, E.-J. Baerends and J. G. Snijders, *The Journal of chemical physics*, 1993, **99**, 4597–4610.
- 72 E. Van Lenthe and E. J. Baerends, *Journal of computational chemistry*, 2003, **24**, 1142–1156.
- 73 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Chem. Phys. Let.*, 1999, **302**, 199–207.
- 74 P. Schipper, O. Gritsenko, S. Van Gisbergen and E. Baerends, *The Journal of Chemical Physics*, 2000, **112**, 1344–1352.
- 75 J. P. Perdew, K. Burke and M. Ernzerhof, *Physical review letters*, 1996, **77**, 3865.
- 76 A. Lembarki and H. Chermette, *Physical Review A*, 1994, **50**, 5328.
- 77 Ch. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema and L. Visscher, *J. Comput. Chem.*, 2011, **32**, 2328–2338.
- 78 T. Saue, R. Bast, A. S. P. Gomes, H. J. A. Jensen, L. Visscher, I. A. Aucar, R. Di Remigio, K. G. Dyall, E. Eliav, E. Fasshauer, T. Fleig, L. Halbert, E. D. Hedegård, B. Helmich-Paris, M. Iliaš, C. R. Jacob, S. Knecht, J. K. Laerdahl, M. L. Vidal, M. K. Nayak, M. Olejniczak, J. M. H. Olsen, M. Pernpointner, B. Senjean, A. Shee, A. Sunaga and J. N. P. van Stralen, *J. Chem. Phys.*, 2020, **152**, 204104.
- 79 DIRAC, a relativistic ab initio electronic structure program, Release DIRAC19 (2019), written by A. S. P. Gomes, T. Saue, L. Visscher, H. J. Aa. Jensen, and R. Bast, with contributions from I. A. Aucar, V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, L. Halbert, E. D. Hedegård, B. Heimlich-Paris, T. Helgaker, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. di Remigio, K. Ruud, P. Sałek, B. Schimmelpfennig, B. Senjean, A. Shee, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, M. L. Vidal, S. Villaume, O. Visser, T. Winther, and S. Yamamoto (available at <http://dx.doi.org/10.5281/zenodo.3572669>, see also <http://www.diracprogram.org>).
- 80 K. G. Dyall, *Theor. Chem. Acc.*, 2002, **108**, 335–340.
- 81 K. G. Dyall, *Theor. Chem. Acc.*, 2003, **109**, 284–284.
- 82 K. G. Dyall, *Theor. Chem. Acc.*, 2016, **135**, 128.
- 83 D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.*, 1994, **100**, 2975.
- 84 L. Visscher, *Theor. Chem. Acc.*, 1997, **98**, 68–70.
- 85 R. A. Opoku, C. Toubin and A. S. P. Gomes, *Predictive simulations of core electron binding energies of halogenated species adsorbed on ice surfaces from relativistic quantum embedding calculations*, 2021, <https://doi.org/10.5281/zenodo.5729761>.
- 86 F. Réal, A. Severo Pereira Gomes, Y. O. Guerrero Martínez, T. Ayed, N. Galland, M. Masella and V. Vallet, *The Journal of chemical physics*, 2016, **144**, 124513.
- 87 S. Woittequand, D. DufLOT, M. Monnerville, B. Pouilly, C. Toubin, S. Briquez and H.-D. Meyer, *The Journal of chemical physics*, 2007, **127**, 164717.
- 88 C. Toubin, S. Picaud, P. Hoang, C. Girardet, B. Demirdjian, D. Ferry and J. Suzanne, *The Journal of chemical physics*, 2002, **116**, 5150–5157.
- 89 H. Yu, T. W. Whitfield, E. Harder, G. Lamoureux, I. Vorobyov, V. M. Anisimov, A. D. MacKerell and B. Roux, *Journal of Chemical Theory and Computation*, 2010, **6**, 774–786.
- 90 P. E. Videla, P. J. Rossky and D. Laria, *The Journal of Physical Chemistry B*, 2015, **119**, 11783–11790.
- 91 C. Houriez, F. Réal, V. Vallet, M. Mautner and M. Masella, *The Journal of Chemical Physics*, 2019, **151**, 174504.
- 92 E. Acher, M. Masella, V. Vallet and F. Réal, *Physical Chemistry Chemical Physics*, 2020, **22**, 2343–2350.
- 93 Z. Li, L. F. Song, P. Li and K. M. Merz, *Journal of Chemical Theory and Computation*, 2021, **17**, 2342–2354.
- 94 M. L. Berkowitz, *The Journal of Physical Chemistry B*, 2021, **125**, 13069–13076.
- 95 S. Skanthakumar, G. B. Jin, J. Lin, V. Vallet and L. Soderholm, *The Journal of Physical Chemistry B*, 2017, **121**, 8577–8584.
- 96 D. P. Chong, O. V. Gritsenko and E. J. Baerends, *The Journal of Chemical Physics*, 2002, **116**, 1760–1772.



- 97 N. P. Bellafont, F. Illas and P. S. Bagus, *Physical Chemistry Chemical Physics*, 2015, **17**, 4015–4019.
- 98 E. van Lenthe, E. J. Baerends and J. G. Snijders, *The Journal of Chemical Physics*, 1994, **101**, 9783–9792.
- 99 W. Hayes and F. C. Brown, *Physical Review A*, 1972, **6**, 21.
- 100 E. Aitken, M. Bahl, K. Bomben, J. Gimzewski, G. Nolan and T. Thomas, *Journal of the American Chemical Society*, 1980, **102**, 4873–4879.
- 101 *X-ray Data Booklet*, ed. A. Thompson, D. Attwood, E. Gullikson, M. Howells, K.-J. Kim, J. Kirz, J. Kortright, I. Lindau, Y. Liu, P. Pianetta, A. Robinson, J. Scofield, J. Underwood, G. Williams and H. Winick, Lawrence Berkeley National Laboratory, University of California, Berkeley, 2009.
- 102 B. V. Crist, *XPS handbook: elements & native oxides*, XPS International, 2019.
- 103 *Handbook of X-Ray Photoelectron Spectroscopy: A Reference Book of Standard Spectra for Identification and Interpretation of XPS Data*, ed. J. Moulder, W. Stickle, P. Sobol and K. Bomben, Physical Electronics, 1995.
- 104 V. A. Yerokhin, A. Surzhykov and A. Müller, *Phys. Rev. A*, 2017, **96**, 042505.
- 105 A. S. P. Gomes, C. R. Jacob, F. Réal, L. Visscher and V. Vallet, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15153–15162.
- 106 M. De Santis, L. Belpassi, C. R. Jacob, A. Severo Pereira Gomes, F. Tarantelli, L. Visscher and L. Storchi, *J. Chem. Theory Comput.*, 2020, **16**, 5695–5711.
- 107 E. Pelimanni, L. Hautala, A. Hans, A. Kivimäki, M. Kook, C. Küstner-Wetekam, L. Marder, M. Patanen and M. Huttula, *The Journal of Physical Chemistry A*, 2021, **125**, 4750–4759.
- 108 L. Partanen, M.-H. Mikkela, M. Huttula, M. Tchapyguine, C. Zhang, T. Andersson and O. Björneholm, *The Journal of Chemical Physics*, 2013, **138**, 044301.
- 109 N. Kurahashi, S. Karashima, Y. Tang, T. Horio, B. Abulimiti, Y.-I. Suzuki, Y. Ogi, M. Oura and T. Suzuki, *The Journal of Chemical Physics*, 2014, **140**, 174506.
- 110 B. Winter, R. Weber, I. V. Hertel, M. Faubel, P. Jungwirth, E. C. Brown and S. E. Bradforth, *Journal of the American Chemical Society*, 2005, **127**, 7203–7214.
- 111 S. Woittequand, C. Toubin, M. Monnerville, B. Pouilly, S. Briquez and S. Picaud, *Surface Science*, 2007, **601**, 3034–3041.
- 112 V. Vallet, J. Coles, F. Réal, C. Houriez and M. Masella, 2021, <https://doi.org/10.48550/arXiv.2109.15265>.

