



HAL
open science

A consistent picture of phosphate-divalent cation binding from models with implicit and explicit electronic polarization

Julie Puyo-Fourtine, Marie Juillé, Jérôme Hénin, Carine Clavaguéra, Elise Duboué-Dijon

► To cite this version:

Julie Puyo-Fourtine, Marie Juillé, Jérôme Hénin, Carine Clavaguéra, Elise Duboué-Dijon. A consistent picture of phosphate-divalent cation binding from models with implicit and explicit electronic polarization. *Journal of Physical Chemistry B*, 2022, 126 (22), pp.4022-4034. 10.1021/acs.jpcc.2c01158 . hal-03677547

HAL Id: hal-03677547

<https://hal.science/hal-03677547>

Submitted on 24 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.

A consistent picture of phosphate–divalent cation binding from models with implicit and explicit electronic polarization

Julie Puyo,^{†,‡} Marie Juillé,^{†,‡} Jérôme Hénin,^{†,‡} Carine Clavaguéra,[¶] and Elise Duboué-Dijon^{*,†,‡}

[†]*CNRS, Université Paris Cité, UPR9080, Laboratoire de Biochimie Théorique, 13 rue Pierre et Marie Curie, 75005, Paris, France*

[‡]*Institut de Biologie Physico-Chimique – Fondation Edmond de Rothschild, PSL Research University, Paris, France*

[¶]*Université Paris-Saclay, CNRS, Institut de Chimie Physique, UMR8000, 91405 Orsay, France*

E-mail: duboue-dijon@ibpc.fr

Abstract

The binding of divalent cations to the ubiquitous phosphate group is essential for a number of key biological processes, such as DNA compaction, RNA folding or interactions of some proteins with membranes. Yet, probing their binding sites, modes and associated binding free energy is a challenge for both experiments and simulations. In simulations, standard force fields strongly overestimate the interaction between phosphate groups and divalent cations. Here, we examine how different strategies to include electronic polarization effects in force fields—implicitly through the use of scaled charges or pair-specific Lennard-Jones parameters, or explicitly with the polarizable force fields Drude and AMOEBA—capture the interactions of a model phosphate compound, dimethylphosphate, with calcium and magnesium divalent cations. We show that both implicit and explicit approaches, when carefully parametrized, are successful in capturing the overall binding free energy, and that common trends emerge from the comparison of different simulation approaches. Overall, the binding is very moderate, slightly weaker for Ca^{2+} than

Mg^{2+} , and the solvent-shared ion pair is slightly more stable than the contact monodentate ion pair. The bidentate ion pair is higher in energy (or even fully unstable for Mg^{2+}). Our results thus suggest practical ways to capture the divalent cations with biomolecular phosphate groups in complex biochemical systems. In particular, the computational efficiency of implicit models makes them ideally suited for large-scale simulations of biological assemblies, with improved accuracy compared to state-of-the-art fixed-charge force fields.

Introduction

Phosphate groups are ubiquitous in biochemistry,¹ as they not only form the backbone of nucleic acids and the head group of phospholipids, but are also present in metabolites such as NADPH, ADP, or ATP, and are added to protein side chains as post-translational modifications. Their interactions with divalent cations such as Mg^{2+} or Ca^{2+} play a key role in many fundamental biochemical processes. For instance, Mg^{2+} ions have been shown to modulate nucleic acid structures and assist RNA folding.^{2–4} They are also essential for the full

catalytic activity of ribozymes.⁵ In the cell, calcium ions can modulate the interactions of phospholipid membranes with proteins, and the formation of Ca^{2+} -phosphate based structures is also essential for bone mineralization.⁶

However, despite their high biological relevance, the binding modes and associated energetics of biological phosphate compounds with divalent cations are not fully characterized yet. For instance, Mg^{2+} cations are usually found to interact with nucleic acid phosphate groups through their hydration water molecules (thus forming solvent-shared ion pairs or "outer-sphere" coordination),^{7,8} except at specific RNA binding sites, where direct "inner-sphere" interactions have been evidenced (Fig. 1).^{3,8-10} The conditions allowing for direct magnesium-phosphate interactions are still not fully understood, also because the identification of magnesium cations in biological structures is sometimes controversial.¹¹ There is thus a strong need to better understand phosphate-divalent cation interactions, to be able to predict their binding modes, sites, and energetics in biological systems.

A first step is to characterize the ion binding behavior to small model phosphate compounds. The thermodynamics of Mg^{2+} binding to dihydrogen phosphate is well characterized experimentally,¹² with a binding free energy of $\Delta G_{\text{bind}}(\text{Mg}^{2+}) = -7.1$ kJ/mol. The body of experimental data is thinner for Ca^{2+} , but tends to suggest a comparable, slightly weaker binding free energy to dihydrogen phosphate ($\Delta G_{\text{bind}}(\text{Ca}^{2+}) = -5.9$ kJ/mol¹³). The same trend between the two cations is observed for binding to more complex phosphate compounds (*e.g.* ATP, CTP, ADP).¹⁴ In contrast, the information regarding the structure of the corresponding ion pairs remains very partial. Very recently, 2D-IR experiments on dimethylphosphate (DMP) suggested a significant fraction of direct contact ion pairs with both Ca^{2+} and Mg^{2+} in 0.2 M concentrated solutions,¹⁵ but a decomposition of the overall binding energetics into the contribution of the different ion pairs is still missing.

In this context, molecular dynamics simulations can complement experimental data

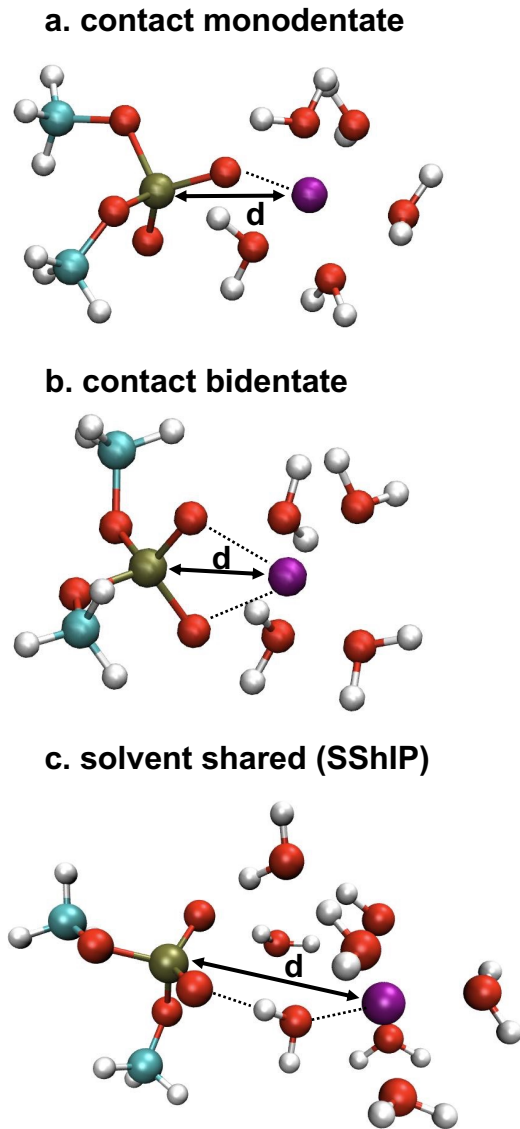


Figure 1: Molecular representation of the different cation-DMP ion pairing modes with the surrounding layer of water molecules. The P-cation distance can be used to distinguish the different ion pairs, with, for Mg^{2+} , typical values of 2.7–3.2 Å for the contact bidentate binding mode, 3.2–2.7 Å for the contact monodentate and 4.1–6.1 Å for the SShIP. Snapshots were prepared using the VMD software.¹⁶

and directly investigate the relative binding free energies corresponding to different binding modes.¹⁷⁻¹⁹ However, the interactions of divalent cations to phosphate groups have proved very challenging to capture in simulations, because of polarization and possible charge transfer effects. In addition, the very slow (μs)

kinetics of water exchange in the magnesium cation hydration shell poses acute sampling problems. Standard fixed-charge force fields have been shown to exhibit strong overbinding artefacts,^{20,21} in line with similar observations pertaining to the binding of divalent cations to other negatively charged biomolecular groups.^{22–24} Several approaches have been suggested to improve the simulated binding behavior, with either explicit or implicit inclusion of electronic polarization.

Polarizable force fields explicitly account for electronic polarization, which can be done in different ways.^{25–27} The binding of Mg^{2+} ions with phosphate groups has been investigated in the past using two different polarizable force fields on slightly different model compounds: AMOEBA, which uses permanent electrostatic multipole moments up to the quadrupole and includes explicit dipole polarization, on dihydrogenphosphate (DHP),¹⁸ and the Drude polarizable model on methylphosphate.¹⁹ They were both successful in reproducing the overall binding free energy, and predicted solvent shared ion pairs (SShIP) to be predominant. These studies were however limited to Mg^{2+} binding, and did not compare it with Ca^{2+} .

Another strategy is to account for induced polarization implicitly, within the standard fixed-charge functional form of the force-field. Two such mean-field approaches have been suggested in the literature, either through the use of pair specific Lennard-Jones (LJ) parameters^{20,28–31} (the so-called NBFIX corrections in CHARMM), or through scaling the total charge of the ion (Electronic Continuum Correction, or ECC).^{23,32–34} The strength of both approaches is that they allow the use of the same software as standard non polarizable force fields, with no additional computational cost. Pair-specific parameters have been recently derived to describe the interactions of Mg^{2+} and Ca^{2+} with dimethylphosphate, enabling simulations to reproduce the overall binding free energy and hydration kinetics.^{20,35} The scaled-charge ECC description has shown promising results for the interaction of polyvalent phosphate species with cations³⁶ but has yet to be tested on monovalent phosphate compounds.

In this work, our goal is to gain molecular insight into Mg^{2+} and Ca^{2+} binding to a model phosphate, DMP, using molecular dynamics simulations. We aim to compare how the relative binding of the two cations and the relative free energy of different binding modes are described by different force fields—non polarizable and with implicit or explicit electronic polarization. We examine how ion pairing free energies and binding modes are captured by a scaled-charge ECC force field, which we compare with a recently proposed force field based on pair-specific Lennard Jones parameters.²⁰ For comparison, we have used two polarizable force fields—the Drude polarizable force field²⁶ and AMOEBA,^{37,38}—which have been recently used to study ion-phosphate interactions,^{18,19} and are both available for nucleic acids and proteins, so that they can be used for simulation of complex biomolecular systems. This makes it possible to identify robust trends in phosphate-cation binding and compare the merits of different simulation strategies, which will pave the way for a better understanding of more complex, biologically-relevant systems.

Computational details

Molecular Dynamics simulations

Molecular dynamics simulations of DMP with a divalent cation—either Mg^{2+} or Ca^{2+} —were performed with different force fields to compare the ion pairing behavior of the two cations and assess its sensitivity to the employed force field. The simulation box was composed of one DMP anion, one cation (Mg^{2+} or Ca^{2+}) and 1723 water molecules.

We performed simulations using force fields from different families, with either implicit or explicit descriptions of electronic polarization. The employed parameters and computational protocol are detailed below. All the parameter files and typical input files for the different softwares are provided in an online repository (<https://zenodo.org/record/5971002#.YkrBDDw6-YA>).

Non polarizable force fields. Several parametrizations of divalent cations using a non polarizable force field and standard integer charges (+2 for divalent cations) are available in the literature. In the main body of this paper, the simulations with standard "non polarizable" force fields were performed with one of the most commonly used parametrizations, namely the 12-6 Lennard-Jones parameters suggested by the Merz group for both Mg^{2+} and Ca^{2+} , choosing the SPC/E compatible versions (labelled "Mg_m" and "Ca_m" in Table 1).³⁹ In the supporting information, we provide additional comparison with simulations using the Mg^{2+} force field suggested by D. Tobias and coworkers,⁴⁰ and the Ca^{2+} force field developed in the Netz group.⁴¹ The dimethyl phosphate was described using the General Amber Force Field (GAFF), which presents similar charges as the phosphate backbone in current Amber nucleic acid force fields (see Supporting Information). These ion force fields were combined with the SPC/E model⁴² for water.

All simulations with non polarizable force fields were performed with the Gromacs2019.6 software^{43,44} with a 2 fs integration step, using periodic boundary conditions, in the constant temperature/constant pressure (NpT) ensemble. We used a Parinello-Rahman barostat⁴⁵ with a 5 ps coupling time and the velocity rescaling thermostat⁴⁶ at 300 K with a 1 ps coupling time. Long-range electrostatic interactions were treated with Particle Mesh Ewald (PME),⁴⁷ with a 12 Å cutoff. Hydrogen-containing bonds were constrained using the LINCS algorithm,⁴⁸ and the SETTLE algorithm for water molecules.⁴⁹

Electronic Continuum Correction. The Electronic Continuum Correction (ECC) is a mean-field theory that prescribes the scaling of ionic charges by a factor $\frac{1}{\sqrt{\epsilon_{\text{el}}}} \simeq 0.75$ (where ϵ_{el} is the high frequency part of water dielectric constant) to implicitly account for electronic polarization.^{33,34} We used previously developed scaled-charge force fields with a 0.75 scaling factor for Mg^{2+} ⁵⁰ (labelled "Mg_s") and Ca^{2+} ⁵¹ ("Ca_s"), which we combined with a scaled-

charge version of the DMP GAFF ("GAFF_s" in Table 1) force field, where we simply scaled all atomic charges by the prescribed 0.75 factor (such a simple strategy was used successfully with acetate previously⁵¹). These ion force fields were used together with the SPC/E model⁴² for water.

Previous studies have shown that a 0.75 scaling factor often tends to underestimate the strength of the ion pairs.^{24,34} This probably originates from the fact that part of the electronic polarization is effectively wrapped into the employed empirical water model, whose dielectric constant is larger than what would correspond solely to nuclear polarization.⁵² Following earlier suggestions, we thus also designed ECC force fields with a milder 0.8 scaling factor, both for the cations²⁴ ("Mg_s2" and "Ca_s2" in Table 1) and for the DMP ("GAFF_s2").

Simulations with scaled-charge force fields were performed using Gromacs, following the same protocol as with standard force fields.

Pair-specific Lennard-Jones parameters.

Prompted by the observation that none of the existing non polarizable force fields for Mg^{2+} captures at the same time its hydration thermodynamics, its hydration water exchange kinetics, and its binding free energy to biomolecular phosphate groups, Schwierz and collaborators recently proposed²⁰ a new Mg^{2+} force field (*microMg*) where they selected the best combination of 12-6 Lennard-Jones parameters to capture simultaneously the thermodynamics and kinetics of Mg^{2+} solvation. In addition, to ensure proper binding properties with RNA, they went beyond the standard Lorentz-Berthelot combination rule and optimized specific Lennard-Jones parameters for the pairwise interaction between Mg^{2+} and the phosphate non bridging oxygens, targeting the overall binding free energy. A similar strategy was later adopted for Ca^{2+} and led to the so-called *Ca2* force field.³⁵ *microMg* and *Ca2* parameters were initially optimized for TIP3P water,⁵³ and a GAFF parametrization of dimethylphosphate,²⁰ slightly different from ours, that we hence adopted in simulations with *microMg*. Note that this force field was very recently

extended to be compatible with other water models.⁵⁴ Simulations with these force fields followed the same protocol as with standard force fields.

Table 1 summarizes the combination of force fields used in each simulation with non explicitly polarizable force fields.

Table 1: Combination of force fields used in each simulation with non explicitly polarizable force fields, using labels for each force fields as defined in the text, where we provide the corresponding references.

Label in Fig. 2	Cation	water	DMP
Mg²⁺			
full charges	Mg_m	SPC/E	GAFF
ECC 0.75	Mg_s	SPC/E	GAFF
ECC 0.8	Mg_s2	SPC/E	GAFF
pair LJ	microMg	TIP3P	GAFF
Ca²⁺			
full charges	Ca_m	SPC/E	GAFF
ECC 0.75	Ca_s	SPC/E	GAFF
ECC 0.8	Ca_s2	SPC/E	GAFF
pair LJ	Ca2	TIP3P	GAFF

Drude polarizable force field. The Drude polarizable force field²⁶ takes into account electronic polarization through the Drude oscillator model, with negatively charged particles attached to each heavy atom via harmonic springs. The Drude particles fluctuate around their parent atom during the course of the simulation, depending on the local electric field, thus modeling the dipole response. Together with the Drude polarizable water model SWM4-NDP,⁵⁵ we used Drude parameters for DMP and Mg²⁺ specifically refined⁵⁶ to reproduce Mg²⁺ solvation properties and its binding to nucleic acid moieties. In addition to Drude particles, this force field includes pair-specific Lennard-Jones interactions (NBFIX) and dipole screening (NBTHOLE) terms both for Mg²⁺-water and Mg²⁺-phosphate (non bridging oxygen) interactions. Simulations were performed with the NAMD2.14 program,⁵⁷ with a timestep of

1 fs, in the NpT ensemble at room temperature and pressure, using Langevin dynamics with a 1 ps⁻¹ damping period and a Langevin Piston Nosé-Hoover barostat.⁵⁸ Periodic boundary conditions were used with a PME treatment.⁴⁷ The temperature of the Drude oscillators was maintained at 1 K with a separate Langevin thermostat with a 20 ps⁻¹ damping coefficient.

AMOEBA. The AMOEBA^{37,59} force field models the electrostatic potential with atomic permanent electrostatic multipole moments up to the quadrupole, and includes explicit self-consistent dipole polarization. It has been extended to describe phosphates (incl. DMP)⁶⁰ and nucleic acids.³⁸ Parametrization of the nucleic acid backbone in the most recent version of the force field (amoebanuc17.prm) is based on the initial DMP parametrization,⁶⁰ which we thus used here in combination with the water and Mg²⁺ parameters currently recommended for nucleic acids.³⁸ Simulations with the AMOEBA force field were performed with Tinker⁶¹ on GPU using the Tinker-OpenMM implementation,⁶² in the NVT ensemble at 300 K. The RESPA integrator⁶³ was used with a 2 fs time step. Long-range electrostatic interactions were treated with PME, using a real-space cutoff of 8 Å.

To examine the sensitivity of the binding free energy $\Delta G_{\text{bind}}^{\circ}$ to the set of multipoles, we refined new multipoles using clusters with explicit water molecules, that should thus be suited for simulation of DMP in solution. The geometry of a cluster of 26-water molecules solvating DMP was optimized at the DFT level using the ω B97X-D3 functional and the def2-SVP basis set, with the Orca program package.^{64,65} Following the AMOEBA procedure,⁶⁶ the atomic multipolar distribution of the optimised DMP-water cluster was derived from the MP2/cc-pVDZ density using the Gaussian program⁶⁷ and DMA analysis.⁶⁸ The multipoles were then refined using an adjustment of atomic dipoles and quadrupoles on the electrostatic potential while keeping the charges fixed.

Binding free energy calculations

Binding free energy calculations were performed in each binding mode using the broadly adopted double decoupling theoretical framework.^{69,70} The corresponding thermodynamic cycle and notations are presented in details in Refs. 71 and 72. Two alchemical transformations are performed, where the cation is progressively decoupled from its environment, either in the bulk (ΔG_{bulk}^*) or in the ion pair (ΔG_{site}^*). The free energy contribution due to the change in net charge during the alchemical transformation (which induces a change in neutralizing "gellium" in our simulations with periodic boundary conditions and PME summation) is taken into account with an analytic correction ΔG_{PBC} (eq(17) from Ref. 73). Ion pair geometries during the computation of ΔG_{site}^* are maintained with flat-well harmonic restraints on the distance between the cation and the DMP phosphorus atom, which allows to distinguish the different binding modes (see Figure 1). The free energy of applying the restraints in the uncoupled state with respect to the standard state concentration, $\Delta G_{\text{decoupled}}^{\text{V}^\circ \rightarrow \text{rest}}$, is estimated numerically as $\Delta G_{\text{decoupled}}^{\text{V}^\circ \rightarrow \text{rest}} = -RT \ln \left(\int_0^\infty 4\pi r^2 e^{-\beta U_{\text{rest}}(r)} dr \right) \simeq -RT \ln \left(\int_0^L 4\pi r^2 e^{-\beta U_{\text{rest}}(r)} dr \right)$, where L is the box size and U_{rest} the flat-well potential. In the coupled state, since the employed flat-well potential is very close to the infinite well which is taken as the definition of the bound state, the contribution of the restraints $\Delta G_{\text{coupled}}^{\text{rest} \rightarrow \text{site}}$ (*i.e.* the free energy associated with removing the restraints in the coupled state, to have an unrestrained ion in the defined binding site) is almost exactly 0. Additional details on the employed thermodynamic site and meaning of the different free energy terms can be found in Ref. 72. The standard binding free energy for a given binding mode $\Delta G_{\text{bind}}^\circ$ is thus obtained by adding those contributions:

$$\Delta G_{\text{bind}}^\circ = \Delta G_{\text{bulk}}^* + \Delta G_{\text{PBC}} + \Delta G_{\text{decoupled}}^{\text{V}^\circ \rightarrow \text{rest}} - \Delta G_{\text{site}}^* + \Delta G_{\text{coupled}}^{\text{rest} \rightarrow \text{site}}$$

The overall binding free energy was obtained by combining the binding free energies obtained

in each ion pair geometry, as follows:

$$\Delta G_{\text{bind,tot}}^\circ = -RT \ln \left(e^{-\Delta G_{\text{bind,mono}}^\circ/RT} + e^{-\Delta G_{\text{bind,bi}}^\circ/RT} + e^{-\Delta G_{\text{bind,tri}}^\circ/RT} \right) \quad (1)$$

This alchemical approach for the calculation of binding free energies has been shown previously^{74,75} to be formerly equivalent to PMF-based determinations, if, as we do here, the free energy contributions of the restraints are reliably evaluated. In the present case, the high free energy barriers separating the different ion pairs could make obtaining a converged PMF challenging, especially with Mg^{2+} . In contrast, the separate estimation of affinities for each of the binding modes reduces correlation times and accelerates convergence, hence our method choice here.

Prior to the alchemical transformations, short unrestrained equilibrations were started from different ion pairing geometries (Fig. 1) with each force field to estimate the range of P-cation distances corresponding to each ion pair. The detailed restraint potentials used in each case during the alchemical transformations are provided in the Supporting Information.

The computational details for the alchemical transformations depend on the simulation software, and thus the force field, as follows.

Non polarizable force fields. For all non explicitly polarizable force fields, alchemical transformations were carried on with Gromacs. Electrostatic interactions between the cation and its environment were first turned off (11 windows, $\lambda_{\text{elec}} = 1, 0.9, 0.8, \dots, 0.1, 0$) before switching off van der Waals interactions (4 windows with $\lambda_{\text{vdw}} = 0.7, 0.5, 0.1, 0$). Starting in the coupled state, each window was equilibrated for 250 ps before starting the next one, and another 5 ns sampling was then added independently in each window. The Bennett Acceptance Ratio⁷⁶ (BAR) method, as implemented in Gromacs, was then used to obtain the free energy. The error bar provided on $\Delta G_{\text{bind}}^\circ$ is the sum of the error for each alchemical transformation, which are estimated from the average variance over 5 blocks as implemented by default in `gmx bar`. On a selected system, we

replicated the alchemical transformations three times. The observed standard deviation of the computed free energies is consistent with the error estimate reported by "gmx bar". We also checked that the computed free energy is independent of the direction of the alchemical transformation (forward / backward), which is a good sign of convergence (see details in Supporting Information).

Drude force field. The alchemy module in NAMD is not yet fully compatible with the Drude force field. Hence, we decomposed the decoupling in two stages, as suggested previously.¹⁹ First, the Mg^{2+} charge and associated Drude particle charge were gradually reduced to zero, using scaling factors of 1, 0.95, 0.9, ..., 0.1, 0.05, 0. A 2 ns-simulation was performed with each of the 21 corresponding PSF files, printing out the structures and energies every 1 ps. In a post-processing step, the energy of each frame of the trajectories was recalculated with the potentials of the neighboring windows. These energies were then collected to yield a BAR⁷⁶ estimate of the free energy. In a second stage, the Lennard-Jones interactions of the uncharged Mg^{2+} were gradually removed ($\lambda_{\text{LJ}} = 1, 0.9, 0.8, 0.6, 0.4, 0.2, 0$) using NAMD FEP alchemy module with Interleaved Double-Wide Sampling, which allows the use of the BAR estimator. For both stages, the BAR free energy estimation was performed in a post-processing step using a python jupyter notebook⁷⁷ and the alchemlyb^{78,79} library. Error estimates were obtained by subsampling the energy data to obtain a decorrelated set, then performing the MBAR error analysis as described by Shirts and Chodera⁷⁸ and implemented in the pymbar library.

AMOEBA. Alchemical transformation with the AMOEBA force field were performed with the Tinker-openMM software.⁶² The interactions of the cation with its environment were progressively decoupled using 24 4-ns long windows, first progressively reducing the electrostatic and polarization interactions ($\lambda_{\text{elec}} = 1, 0.9, \dots, 0.1, 0$) before decoupling the van der waals interactions ($\lambda_{\text{LJ}} =$

1, 0.9, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55, 0.5, 0.4, 0.2, 0.1, 0). The final free energy estimate and error was obtained with BAR as implemented in the bar_omm.x Tinker tool.

Computation of free energy landscapes with ABF

Free energy landscapes were sampled and computed using a variant of the Adaptive Biasing Force (ABF) method,⁸⁰ extended-system ABF (eABF).^{81,82}

Consider a vector collective variable (CV) $\mathbf{z} = \xi(\mathbf{x})$, where \mathbf{x} is the vector of atomic Cartesian coordinates. We define the free energy landscape, up to an additive constant, as the quantity

$$A(\mathbf{z}) \equiv -k_B T \ln \rho(\mathbf{z}),$$

where $\rho(\mathbf{z})$ is the marginal probability density of the collective variable. In ABF, the gradient ∇A of the free energy surface $A(\mathbf{z})$ is estimated on-the-fly as the conditional average of a projected force,^{80,83,84} following a Thermodynamic Integration formalism. This running estimate is applied as a biasing force, which tends to level the local slope of the free energy landscape. At long times, as the estimate converges to the free energy gradient, the effective free energy landscape becomes flat and the biased distribution of \mathbf{z} becomes uniform. In ABF, estimating the projected force imposes technical requirements on the CVs.⁸³ These requirements are relaxed in eABF, by not applying the ABF algorithm directly to the CVs of interest, but to fictitious degrees of freedom λ that evolve according to Langevin dynamics and are harmonically coupled to the CVs through a potential $V^{\text{ext}}(\mathbf{x}, \lambda) = \frac{k}{2} |\xi(\mathbf{x}) - \lambda|^2$ with force constant k .^{81,82} The implementation used was the Collective Variables Module (Colvars),⁸⁵ which was recently interfaced with Gromacs. Gromacs patched with Colvars was downloaded from <https://github.com/Colvars/gromacs/tags>. Three collective variables were defined, corresponding to the three Cartesian coordinates of the divalent cation in a moving frame of reference tied to the phosphate group. An

ideal, symmetric geometry of the PO_4 group was created and used to define the reference position to which those atoms were aligned transparently throughout the simulations, by invoking the `fittingGroup` option of Colvars. This allowed for the resolution of well-defined free energy landscapes reflecting the positional distribution of divalent cations with respect to the PO_4 group. Unbiased free energy gradients were obtained by the CZAR estimator,⁸² and integrated on the fly using a Poisson equation formalism,⁸⁶ following the default settings of Colvars. For 2d visualization, free energy surfaces were read into Python notebooks using the `colvars_grid` Python module,⁸⁶ then integrated over relevant slices of the collective variables using `numpy` and plotted using `matplotlib`. For 3d visualization, the DX files written by Colvars were read directly into VMD.¹⁶

The collective variables were discretized on a grid of step size 0.2 Å. For eABF, the extended variables were coupled with an RMS deviation of 0.2 Å and a time constant of 200 fs. ABF forces were fully applied when 50 samples were collected in a given bin.

Quantum mechanical (QM) calculations

We selected from the molecular dynamics simulations 20 snapshots in the monodentate and SShIP geometries. We then extracted from each snapshot a cluster composed of one DMP anion, one Mg^{2+} , and 20 water molecules, selected based on the closest distance d between their H atoms and the cation or DMP oxygen atoms. All water molecules with $d < 3.5$ Å were selected, complemented to 20 by selecting additional water molecules by increasing value of d . The number of 20 water molecules ensured that the entire first hydration shell of both Mg^{2+} and the DMP oxygen atoms (the water molecules directly donating a hydrogen bond) were selected, including in the SShIP. The geometry of these clusters was optimized with Density Functional Theory (DFT) with the ORCA software,^{64,65} in the gas phase, using the $\omega\text{B97X-D3(BJ)}$ functional with Becke-Johnson dispersion correction and the `def2-SVP` basis set. Single point en-

ergies were then obtained with a larger basis set (as in Ref. 60), which proved key to obtain correct interaction energies, both for the full cluster and for the DMP and ion+water fragments. From there, the interaction energy was calculated as $E_{\text{int}} = E_{\text{tot}} - E_{\text{DMP}} - E_{\text{cation+water}}$. Single point energies and the corresponding interaction energies were then calculated on the same optimized fragments with different versions of the AMOEBA force field using the *analyze* program of Tinker. Detailed data are provided in Supporting Information.

Results

Phosphate–cation binding without explicitly polarizable force fields.

Cations can interact with phosphate groups in distinct binding modes. The interaction can take place either through a water molecule, thus forming a solvent-shared ion pair, or with direct contact between the cation and the DMP oxygen atoms. Such a contact ion pair can be either monodentate (interaction with only one of the two non-bridging phosphate oxygen atoms), or bidentate (Fig. 1). To investigate how different force fields capture the interactions of DMP with divalent cations, Mg^{2+} and Ca^{2+} , and how they predict the balance between these binding modes, we computed the standard binding free energy $\Delta G_{\text{bind}}^{\circ}$ between Mg^{2+} or Ca^{2+} and DMP in different ion pair geometries with different non polarizable force fields (Fig 2 and Table S2 in the Supporting Information). The overall binding free energy was compared with the experimental reference (available only for DHP, presumed very similar to DMP), -7.1 kJ/mol for Mg^{2+} ¹² and -5.9 kJ/mol for Ca^{2+} .¹³ As discussed already in the literature^{72,87,88} the definition of the "bound" state depends on the type of experiment performed, since different techniques are sensitive to different ion pairs. Here, following previous works,¹⁸ we assume that the chosen reference free energies report on the formation of both contact and solvent-shared ion pairs. As we discuss later, this is not the case for spectroscopy-derived estimates.

We also note that some works³⁵ chose different, weaker, reference values reported by Sigel and Sigel.⁸⁹ Those values result from an extrapolation of various data obtained on related, but different, systems. We, as others before,¹⁸ thus preferred to use experimental estimates directly obtained on DHP.

$\Delta G_{\text{bind}}^{\circ}$ for Mg^{2+} is strongly overestimated, by more than 20 kJ/mol, with standard non polarizable force fields, which is consistent with the overbinding artefacts reported for Mg^{2+} binding to biomolecules in the literature.^{20,23} This excessive binding is observed independently of the details of the chosen force field and is robust with respect to a change of water model (see Supporting Information). The overbinding is also observed for Ca^{2+} , albeit to a much smaller extent, with the 12-6 force field by Merz and coworkers.³⁹ Other non polarizable force fields exhibit an overbinding for Ca^{2+} almost as pronounced as for Mg^{2+} (see Supporting Information). Additionally, with such force fields, both cations are predicted to interact preferentially in direct contact with the DMP non bridging oxygen atoms (rather than through their hydration shell) which, for Mg^{2+} , goes against the usual picture of such interactions.

The artefacts exhibited by standard force fields are believed to originate from the lack of electronic polarization, which is particularly important to properly capture binding properties of highly charged cations. One possibility to overcome this limitation is to use a mean-field approach and take implicitly into account electronic polarization by scaling the ionic charges.^{23,33,34} This approach is expected to be transferable to all non-transition metal cations, interacting with any charged anion. It has already been successfully used for instance for Na^+ , Li^+ , K^+ , Ca^{2+} , Mg^{2+} , and Zn^{2+} in interaction with simple anions (e.g. Cl^- , Br^-) and typical biomolecular charged groups such as carboxylates or phospholipids.²³ Following the prescriptions of the Electronic Continuum Correction (ECC) theory, we thus developed a scaled-charge force field both for DMP and the divalent cations,^{50,51} using the theoretically-derived $\frac{1}{\sqrt{\epsilon_{\text{el}}}} \simeq 0.75$ scaling fac-

tor. As expected, this scaled-charge force field yields much weaker binding free energies (Fig 2 and Table S2). However, as observed previously with other solutes (e.g. acetate), the overall binding free energy is now underestimated. This led us to prefer a 0.8 scaling factor, which can be justified by the fact that some of the electronic polarization effects are already implicitly incorporated in the water model.^{34,52} With this second version of the ECC force field, the binding free energy for Mg^{2+} is now within less than 3 kJ/mol from the experimental value, probably within experimental (and computational) error bars, or slightly underestimated still. The overall binding free energy for Ca^{2+} is slightly smaller than that of Mg^{2+} , by about 2 kJ/mol, consistently with experimental data. Strikingly, compared to standard non polarizable force fields, in addition to a much weaker binding strength, the favoured binding mode is now the SShIP, for both cations. The contact binding mode is found slightly less stable, by 2–4 kJ/mol, the monodentate geometry being favored over bidentate for both cations. The bidentate binding mode is found unstable with Mg^{2+} (it is not associated with a local free energy minimum and cannot be stabilized even with restraints), while it is metastable with Ca^{2+} . The free energy difference between the monodentate and solvent-shared binding modes is very modest, so that the relative ordering of these two binding modes can be easily switched by changing, for instance, the water model (see discussion in Supporting Information).

Another approach to eliminate overbinding artefacts is to parametrize pair-specific Lennard Jones parameters between the studied cations and the solute they interact with.^{20,28–31} Recently, parameters were developed specifically to capture Mg^{2+} and Ca^{2+} binding free energies with phosphate groups.^{20,35} Since $\Delta G_{\text{bind}}^{\circ}$ was used as a target in the parametrization process—albeit with a different reference than ours for Ca^{2+} ⁸⁹—these parameters give by construction a good overall binding free energy. Comparing panels a and b in Fig 2, Ca^{2+} seems to bind slightly stronger than Mg^{2+} with this family of force fields. However, we should note that a different set of parameters was

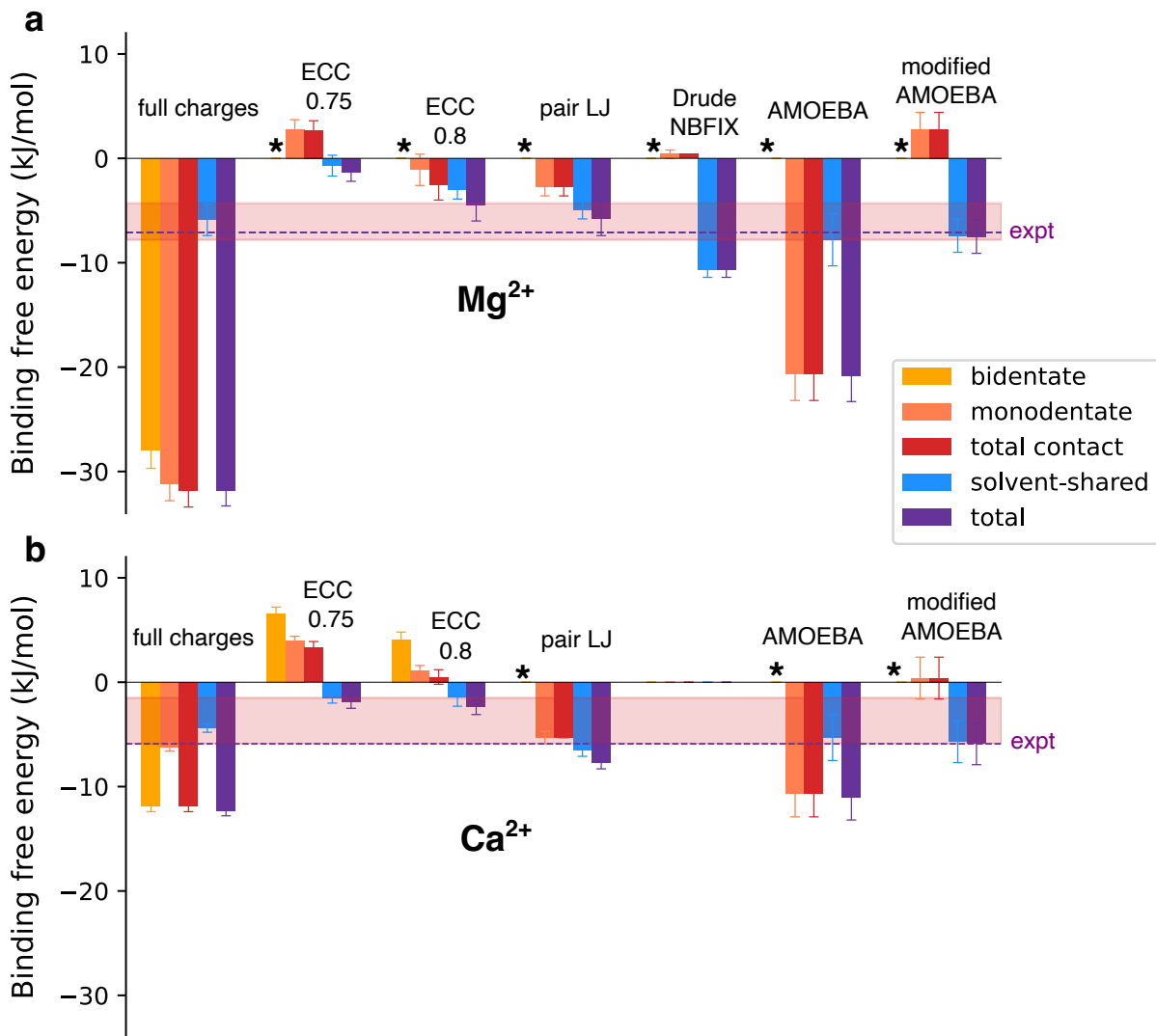


Figure 2: Cation-DMP binding free energy $\Delta G_{\text{bind}}^{\circ}$ (kJ mol⁻¹) for different ion pair conformations (contact monodentate, contact bidentate and SShIP), together with the overall binding free energy $\Delta G_{\text{bind,tot}}^{\circ}$ (Eq. ??), computed either over only the contact binding modes, "total contact", or also including the solvent-shared ion pair, "total". The binding free energies were computed with different force fields, as described in the Methodology section. The specific combination of force fields used in each case is specified in Table 1. The reference experimental binding free energies $\Delta G_{\text{bind,expt}}^{\circ}$ are -7.1 kJ/mol for Mg^{2+} ¹² and -5.9 kJ/mol for Ca^{2+} .¹³ The light red zone indicates the range of binding free energy for the contact ion pairs as estimated from Ref. 15. Unstable ion pair geometries, which are not associated with a local free energy minimum and cannot be stabilized even with restraints, are marked with a * symbol.

recently derived for Mg^{2+} ,³⁵ that binds more strongly than Ca^{2+} (more details in Supporting Information). The relative binding of the two cations thus sensitively depends on details in the parametrization. Here, we decomposed the binding free energy into the contribution of the different ion pairs (Fig 2 and Table S2). As with the ECC force field, the most stable bind-

ing mode is the SShIP, the contact monodentate ion pair being less stable by 1–2 kJ/mol. The contact bidentate ion pair, which, for Ca^{2+} , was only a few kJ/mol less stable than the contact monodentate with ECC, is fully unstable (*i.e.* is not associated with a local free energy minimum) with this description for both cations. This probably comes from the repulsion caused

by the pair-specific Lennard-Jones term.

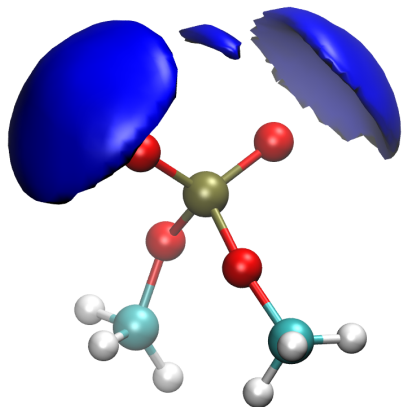


Figure 3: Illustration of the monodentate and bidentate binding modes for the $\text{Mg}^{2+}:\text{DMP}$ ion pair, represented as an iso-free-energy surface computed using eABF for the full charge model. Prepared with VMD.¹⁶

In order to get a better insight into the geometry and relative stability of the different binding modes with the different force fields, we computed, using the eABF enhanced sampling technique,^{81,82} the 3D free energy landscape around the DMP anion. 3D visualization of free energy isosurfaces provides a picture of the spatial extent of each of the binding modes (Figure 3), with symmetric lobes around each oxygen for the contact monodentate ion pair, and a thin lobe, orthogonal to the OPO plane, for the contact bidentate (when stable).

Projecting the 3D free energy data on the plane formed by the phosphorus central atom and the two non bridging oxygens, we obtained 2D free energy maps that enable easy comparison of the binding strength of each mode with the different force fields (Figure 4). While the overall shape of the binding modes is similar with all force fields, clear differences are visible in the binding strength. Full-charges non polarizable force fields yield much deeper free energy basins than the two implicit descriptions;

and ECC, even with 0.8 scaling, yields the most shallow free energy minima, in line with the alchemical evaluation of $\Delta G_{\text{bind}}^{\circ}$. This representation also confirms the absence of a metastable contact bidentate state ($x = 0$) with the use of pair-specific Lennard-Jones parameters, which contrasts with both ECC and full charge descriptions, where the contact bidentate state, albeit the least stable with ECC, indeed corresponds to a metastable state with Ca^{2+} . In addition, even when the contact and SShIP basins are close in free energy, they can be separated—*e.g.* with Mg^{2+} and the ECC force field—by a high (> 15 kJ/mol) free energy barrier, which means that the kinetics of exchange remains slow between the states and won’t be spontaneously sampled in standard non accelerated molecular dynamics simulations.

While scaled ECC force fields and the use of pair-specific LJ parameters are able to reproduce the overall binding free energy and both predict the SShIP to be the lowest free energy minimum, they exhibit differences regarding, for instance, the existence of a metastable bidentate binding mode. Since these force fields only implicitly account for electronic polarization, we now set out to compare their behavior with that of two different explicitly polarizable force fields, the Drude polarizable force field and AMOEBA.

Explicitly polarizable force fields

Drude polarizable force field. The Drude polarizable force field²⁶ takes into account electronic polarization through the Drude oscillator model, with negatively charged Drude particles fluctuating around their parent atom during the simulation. Lemkul and MacKerell⁵⁶ parametrized a Drude polarizable force field for DMP and Mg^{2+} specifically designed to capture Mg^{2+} hydration thermodynamics and structure, as well as its interactions with chloride anions and nucleic acid moieties. The inclusion of pair-specific LJ interactions (NBFIX) between Mg^{2+} and non-bridging phosphate oxygen atoms, in addition to the Drude particles, was found necessary to reproduce the structure of Mg^{2+} -phosphate (DMP) complexes. In addi-

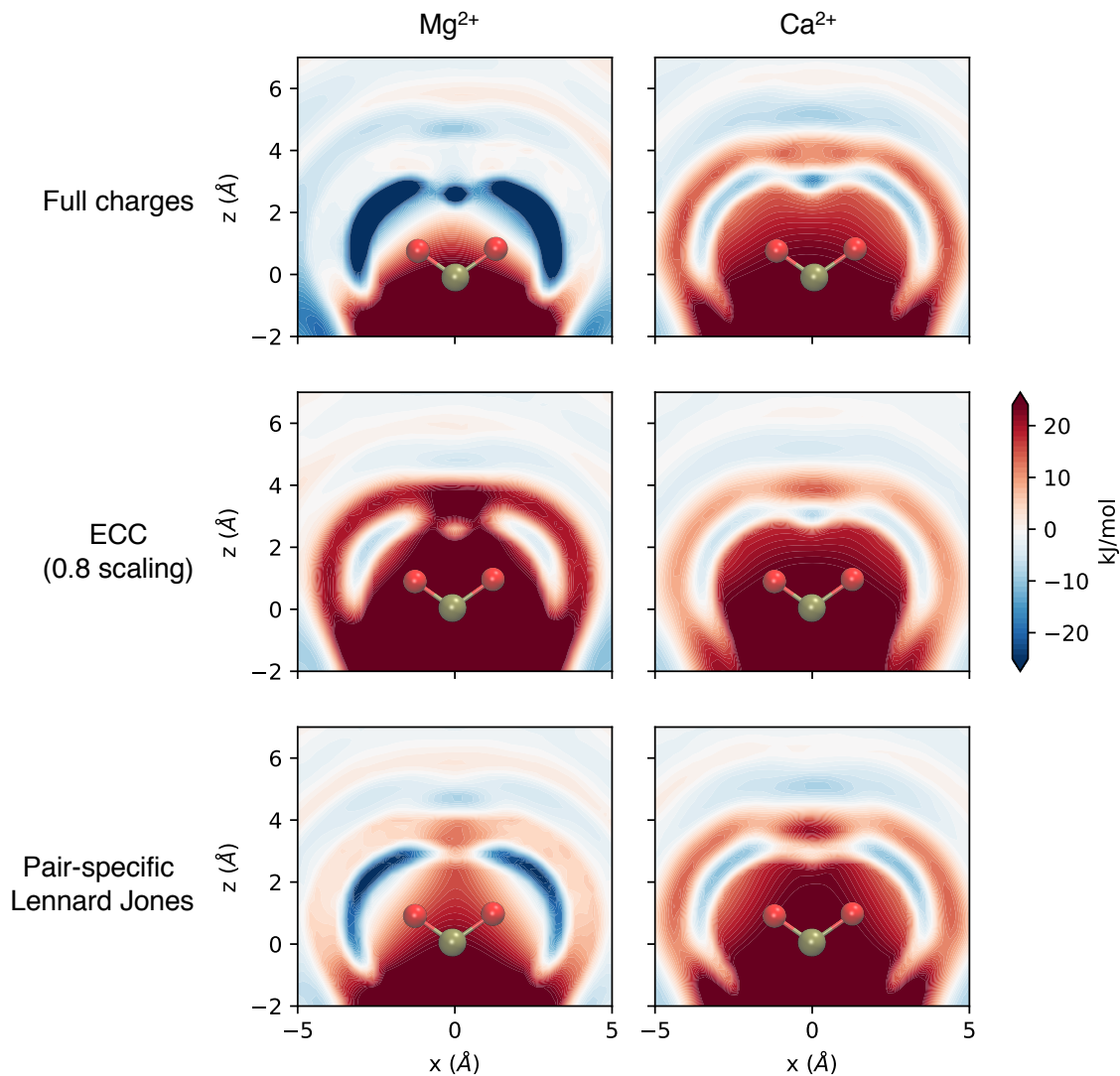


Figure 4: 2D free energy maps for the interaction of a Mg^{2+} or Ca^{2+} cation with dimethylphosphate, computed using eABF, integrated over a 2 \AA -thick slice around the $y = 0$ OPO plane, and capped at $\pm 20 \text{ kJ/mol}$. Rendered with matplotlib.⁹⁰

tion, Villa et al.¹⁹ refined in a similar way (using NBFIX terms), a Drude force field for the interactions of Mg^{2+} cations with methylphosphate. They showed that it quantitatively reproduced the overall binding free energy and predicted the SShIP to be the most stable, the contact monodentate ion pair being significantly less stable by about 9.5 kJ/mol . Using the specifically parametrized force field for DMP and Mg^{2+} ,⁵⁶ we computed the DMP: Mg^{2+} binding free energy for each binding mode (Fig 2

and Table S3). The overall binding free energy (-10.7 kJ/mol) is in good agreement with the experimental reference, and mainly stems from the solvent-shared binding mode, the contact monodentate mode being about 11 kJ/mol high in free energy. These results are fully in line with previous data obtained on methylphosphate.¹⁹ Note that the contact bidentate ion pair is not stable at all with this force field, presumably due to the NBFIX repulsion term, as is also the case with the non polarizable force

field with pair-specific LJ terms (Fig 2 and Table S3).

The Drude polarizable force field for Ca^{2+} -phosphate interactions has not yet been refined with NBFIX terms, as for Mg^{2+} . For the sake of comparison, we still computed the binding free energy in the different modes using the default Ca^{2+} force field. While the SShIP is associated with a reasonable binding free energy (-6.5 kJ/mol), slightly higher than that of Mg^{2+} , which would be in line with experimental data, the contact ion pair is clearly misbehaving, both in terms of structure and energetics (see Supporting Information, Table S3). The monodentate binding mode is unstable, immediately falling in a pathological bidentate conformation, with an extremely small 2.5–2.9 Å Ca^{2+} -phosphorus distance, to be compared with a reference 2.95–3.05 Å distance estimated on 10 clusters (see Methods section) optimized with DFT at the def2-SVP/wB97x-D3(BJ)/PCM level. This kind of deformation is in line with what was observed with Mg^{2+} before the inclusion of the NBFIX term.⁵⁶ Unsurprisingly, the binding free energy corresponding to this bidentate mode is also totally unphysical, leading to an overall $\Delta G_{\text{bind}}^{\circ} = -73.2$ kJ/mol. These results stress that without addition of pair-specific Lennard Jones repulsion terms, the Drude polarizable force field does not properly capture divalent ion binding with biologically relevant moieties, hence the need for careful testing before application to a new family of interaction groups.

AMOEBA. AMOEBA is another family of polarizable force field which uses permanent electrostatic multipole moments up to the quadrupole and includes explicit dipole polarization.³⁷ It has been parametrized for proteins⁹¹ and, more recently, for nucleic acids,³⁸ which now enables its use for simulation of complex biochemical systems. Kumar et al.¹⁸ studied the interactions of Mg^{2+} ions with a model phosphate compound (dihydrogen phosphate, DHP), and showed that it was able to quantitatively reproduce the binding free energy. The most stable ion pair conformation was found to be the SShIP, with both contact ion pairs

(monodentate and bidentate) very close in free energy, about 7 kJ/mol less stable than the solvent shared. More recently, Zhang et al.⁶⁰ derived AMOEBA parameters for dimethylphosphate, which later served as a basis for the nucleic acid backbone in the AMOEBA nucleic acid force field.³⁸ This DMP force field was validated for its interactions with Mg^{2+} against the structure and interaction energy of cation:DMP dimer, in a bidentate geometry optimized at the QM (DFT) level.

With this recent AMOEBA force field for DMP, which should be a good proxy for the RNA/DNA interactions with divalent cations with the AMOEBA force field for nucleic acids, we examined the binding free energy associated with each type of ion pair (Fig 2 and Table S3). The overall Mg^{2+} -DMP binding ($\Delta G_{\text{bind}}^{\circ} = -20.8$ kJ/mol) appears significantly too strong compared to experimental data, especially for the contact monodentate ion pair, which is the most stable. Meanwhile, the SShIP exhibiting a binding free energy much more similar to that obtained with other force fields ($\Delta G_{\text{bind,sship}}^{\circ} = -7.8$ kJ/mol). This contrasts with the results previously obtained on DHP, both quantitatively and qualitatively.

We thus endeavor to better understand the origin of this difference, analyzing how sensitive the binding free energy was to different force field parameters. First, we derived new multipoles for DMP (see Methods), taking into account explicit water molecules in the derivation. However, the binding free energy $\Delta G_{\text{bind}}^{\circ}$ proved very little sensitive to the change of multipoles (see Supporting Information). We then noticed that a significant difference between the early DHP force field¹⁸ and the current DMP force field⁶⁰ was in the value of the atomic polarizability of the phosphorus and non bridging oxygen atoms. For DMP, they were tuned to reproduce the overall DMP polarisability.⁶⁰ Computing the binding free energy for different polarizability values (see Supporting Information) evidenced that the binding free energy of the monodentate binding mode was extremely sensitive to the non-bridging oxygen polarizability and very little to that of the phosphorus. In contrast, the SShIP binding mode was very lit-

tle affected by changes in both polarizabilities.

We then went back to the spirit of the initial development of the AMOEBA force field with the idea of benchmarking some parameters against QM calculations on the structures and energies of small clusters. The polarizability values were thus adjusted on clusters in the monodentate and SShIP binding modes (details in the Supporting Information). We found out that the set of polarizabilities that minimize the error on relative interaction energies ($\alpha_O = 0.837 \text{ \AA}^3$ and $\alpha_P = 3.3 \text{ \AA}^3$, close to those used in the early DHP force field¹⁸) yield binding free energies for both Mg^{2+} and Ca^{2+} in very good agreement with experimental data (see Fig 2 and Table S3), slightly more negative with Mg^{2+} than with Ca^{2+} . In both cases, the SShIP is the most stable ion pair, and the contact monodentate ion pair about 6–10 kJ/mol higher in energy. The contact bidentate ion pair is not stable at all, which is different from what was earlier observed with DHP, but in agreement with the picture provided by most of the other tested force fields. Refinement of the DMP polarizabilities thus had a major impact on binding free energies and led to an ion binding behavior in line with experimental data as well as with simulations with other force fields. Whether a similar change in large scale simulations of nucleic acids would improve their cation binding behavior without deteriorating the properties of the nucleic acid remains to be investigated.

Discussion

The interactions of DMP with Mg^{2+} and Ca^{2+} were examined in details with different force fields that use distinct, implicit or explicit, approaches to take into account electronic polarization. We show that all approaches, after careful refinement, are able to capture the overall binding free energy. From a structural point of view, the obtained binding distances (first peak of the radial distribution function between the DMP non bridging oxygen and the cations, see Table S3 and S4) are all compatible with experimental estimates⁹² (1.92–2.14 Å for Mg^{2+} and 2.2–2.5 Å for Ca^{2+}). From there, a common

picture emerges for divalent cation–phosphate interactions: the binding is overall quite moderate, slightly weaker for Ca^{2+} than Mg^{2+} , and the (slightly) most stable binding mode, for both cations, seems to be the SShIP (outer-sphere). As for the relative stability of the two contact binding modes, monodentate and bidentate, all force fields favor the monodentate binding mode with both cations. However, whether the contact bidentate binding mode can still form a minority population is still an open question. For Mg^{2+} , the contact bidentate binding mode is associated with a local free energy minimum with none of the tested force fields (except for the standard non polarizable full charge, where the strong contact binding is clearly artefactual). This tends to suggest that this binding mode may not exist, which contrasts with the conclusion of an earlier simulation study.¹⁸ For Ca^{2+} , our results are less conclusive, as a (meta)stable state is found with ECC force fields, but does not exist with the tested pair-specific LJ parameters, nor with the AMOEBA polarizable force field. Hence, we cannot fully rule out the existence of this binding mode with Ca^{2+} , which is also difficult to probe experimentally and to distinguish from the monodentate contact ion pair.

In addition, the different tested force fields differ in the balance between SShIP and contact ion pairs. With both implicit approaches (ECC and pair-specific LJ) and for both cations, the contact monodentate ion pair is only slightly (2–4 kJ/mol) less stable than the SShIP. In contrast, with the Drude polarizable force field and AMOEBA, the contact monodentate ion pair is more destabilized with respect to the SShIP (by about 6–10 kJ/mol, more for Mg^{2+} than Ca^{2+}), which means that contact ion pairs would be much less abundant.

Very recently, using 2D-infrared spectroscopy experiments on DMP Mg^{2+} or Ca^{2+} solutions complemented by QM simulations, Schauss et al.¹⁵ suggested the existence of contact ion pairs between DMP and these divalent cations, making use of the fact that the blueshift observed in IR measurements are only caused by the formation of contact ion pairs, and not with solvent-shared. They estimated the con-

centration of DMP:cation contact ion pairs in $C_0 = 0.2$ M solutions between $0.23C_0$ and $0.37C_0$ for Mg^{2+} , and between $0.12C_0$ and $0.3C_0$ for Ca^{2+} . These estimates can be compared with that provided by simulations with different force fields. Direct counting in simulations run at the same concentration as in the experiment is not practical because of the very long exchange times between the different ion pairs, which prevents proper convergence of the contact ion pair concentration. Instead, we can convert the estimated contact ion pair concentrations into the binding free energy associated with contact pairs, $\Delta G_{\text{bind,contact}}$. This thus provides an estimated $\Delta G_{\text{bind,contact}} = -7.8$ to -4.3 kJ/mol for Mg^{2+} and $\Delta G_{\text{bind,contact}} = -5.9$ to -1.5 kJ/mol for Ca^{2+} . $\Delta G_{\text{bind,contact}}$ can be directly compared with that computed with different force fields, from the exponential average of the ΔG_{bind} associated with the two contact pairs (see Fig. 2). All force fields (among those giving a correct overall binding free energy) seem to underestimate the fraction of contact ion pairs. The contact ion pair binding free energy is slightly underestimated by about 2 to 4 kJ/mol with both ECC and pair-specific LJ descriptions, except for Ca^{2+} , where the pair specific LJ description falls within experimental error bars. Both polarizable force fields (Drude and our modified AMOEBA) underestimate even more the contact ion pair free energy, by at least 4.5 kJ/mol for Mg^{2+} .

Overall, implicit and explicit descriptions of electronic polarization yield quite similar pictures of the ion pairing behavior of phosphate groups with Mg^{2+} and Ca^{2+} . However, with the Drude polarizable force field, correct ion binding behavior crucially depends on added pair-specific interaction; with AMOEBA, it is highly sensitive to the DMP atomic polarizabilities of the non-bridging oxygen atoms. Both implicit strategies (ECC and pair-specific LJ) can be directly used with standard simulation software at no additional cost, which makes them ideally suited for large-scale simulations of biological assemblies. However, they also come with their drawbacks. Pair-specific Lennard Jones parameters lack transferability, as they need to

be parametrized specifically for each pair of interacting ions and the relative behavior the two cations proved quite sensitive to details in the parametrization. This approach compensates the short-range overbinding artefacts by an *ad hoc* repulsive term, thus recovering a correct binding behavior, but does not fix the missing electronic term in long-range electrostatics. In contrast, the scaled-charge ECC force fields introduce the missing electronic polarizability in a physically well-grounded mean-field approach. However, this strategy requires scaling the charges of both the cations and the anionic solute, which means that in biosimulations of nucleic acids, the charges of the RNA or DNA moiety should also in principle be scaled. This will affect the nucleic acid–water interactions and could alter the biomolecule conformation. Additional developments and tests are thus needed to be able to use such force fields for more complex systems. Also, for non polarizable force fields, we focused here on the cation force fields and did not explore how modifying the DMP force field (beyond a uniform charge scaling factor) would impact the binding free energy, which should be explored in future works.

Finally, even if we don't explicitly examine the molecular details of the conversion between inner-sphere (contact) and outer-sphere (SShIP) binding modes, as has been done elsewhere,⁹³ we can see that even with the force fields that properly capture the binding free energy, exchange kinetics between the different ion pairs remains slow (compared to typical simulation timescales). This is evidenced, for instance, by the large barriers separating the contact and solvent-shared binding modes in Fig.4. Hence, special care should always be taken to properly sample the distinct binding modes.

Conclusion

While standard non-polarizable force fields cannot properly capture the binding structure and free energy of divalent cations to phosphate moieties, we have shown that force fields

that include electronic polarization in an implicit mean-field way—either through the use of scaled charges (ECC theory) or pair-specific Lennard-Jones parameters—successfully reproduce the modest binding free energy of Mg^{2+} and Ca^{2+} to a model isolated phosphate compound, DMP. Similar results are obtained with explicitly polarizable force fields, provided extra care is taken during the parametrization process, with the addition of pair-specific repulsion for the Drude polarizable model, and fine tuning of the atomic polarizabilities for AMOEBA. From the comparison of these different approaches, a common picture of cation-phosphate binding emerges: with both Mg^{2+} and Ca^{2+} cations, the most stable binding mode is the SShIP, and among the two contact geometries, the monodentate binding mode seems the most stable, even though the various force fields differ in the relative stability of the monodentate and solvent shared binding modes. Comparison with recent experimental estimates of the fraction of contact ion pairs however suggests that all the force fields may underestimate the stability of the contact binding mode. Given that force fields which only implicitly take into account electronic polarization yield very similar descriptions of the cation-phosphate binding modes and binding free energies as explicitly polarizable force fields, these approaches seem very promising for simulation of large scale biochemical assemblies with divalent cations. The two strategies however pose their own challenges, the lack of transferability for the pair-specific LJ approach, and the scaling of the biomolecular charges for ECC. It also remains to be examined whether those different approaches are able to describe the formation of strong binding sites, as found in RNA for instance, that involve multiple phosphate groups.

Acknowledgement This work was supported by the "Initiative d'Excellence" program from the French State (Grants "DYNAMO", ANR-11-LABX-0011, and "CACSIce", ANR-11-EQPX-0008). Computational work was performed using local HPC resources managed by G. Letessier, as well as HPC resources from GENCI (project A0090711021). JP ac-

knowledges a PhD grant from the "DYNAMO" Labex (ANR-11-LABX-0011). We thank Dr. J. Lemkul (Virginia Tech) and Prof. Thomas Simonson (Ecole Polytechnique) for their help to set up simulations and alchemical transformations with the Drude polarizable force field, and Hubert Santuz (CNRS, LBT) for help with interfacing Colvars with Gromacs.

Supporting Information Available

Supporting Information is available, with additional computational details, and complementary analyses. Input and parameter files are shared on a Zenodo public folder (<https://zenodo.org/record/5971002#.YkrBDDw6-YA>).

References

- (1) Westheimer, F. H. Why nature chose phosphates. *Science* **1987**, *235*, 1173–1178.
- (2) Draper, D. E.; Grilley, D.; Soto, A. M. Ions and RNA Folding. *Annu. Rev. Biophys. Biochem. Struct.* **2005**, *34*, 221–243.
- (3) Bowman, J. C.; Lenz, T. K.; Hud, N. V.; Williams, L. D. Cations in charge: Magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.* **2012**, *22*, 262–272.
- (4) Woodson, S. A. Metal ions and RNA folding: A highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* **2005**, *9*, 104–109.
- (5) Hanna, R.; Doudna, J. A. Metal ions in ribozyme folding and catalysis. *Curr. Opin. Chem. Biol.* **2000**, *4*, 166–170.
- (6) Brini, M.; Ottolini, D.; Cali, T.; Carafoli, E. In *Calcium in health and disease*; Sigel, A., Sigel, H., Sigel, R., Eds.; Springer Science+Business Media, 2013; Vol. 13; pp 81–137.
- (7) Egli, M. DNA-Cation Interactions: Quo Vadis? *Chemistry & Biology* **2002**, *9*, 277–286.

- (8) Auffinger, P.; Grover, N.; Westhof, E. *Metal ion binding to RNA.*; 2011; Vol. 9; pp 1–35.
- (9) Klein, D. J.; Moore, P. B.; Steitz, T. A. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **2004**, *10*, 1366–1379.
- (10) Freisinger, E.; Sigel, R. K. From nucleotides to ribozymes—A comparison of their metal ion binding properties. *Coord. Chem. Rev.* **2007**, *251*, 1834–1851.
- (11) Auffinger, P.; Ennifar, E.; D’Ascenzo, L. Deflating the RNA Mg²⁺ bubble. Stereochemistry to the rescue! *RNA* **2021**, *27*, 243–252.
- (12) Verbeeck, R. M.; Bruyne, P. A. D.; Driessens, F. C.; Verbeek, F. Solubility of magnesium hydrogen phosphate trihydrate and ion-pair formation in the system Mg(OH)₂-H₃PO₄-H₂O at 25°C. *Inorg. Chem.* **1984**, *23*, 1922–1926.
- (13) McDowell, H.; Brown, W. E.; Sutter, J. R. Solubility Study of Calcium Hydrogen Phosphate. Ion-Pair Formation. *Inorg. Chem.* **1971**, *10*, 1638–1643.
- (14) Smith, R. M.; Martell, A. E.; Chen, Y. Critical Evaluation of stability constants for nucleotide complexes with protons and the accompanying enthalpy changes. *Pure Appl. Chem.* **1991**, *63*, 1015–1080.
- (15) Schauss, J.; Kundu, A.; Fingerhut, B. P.; Elsaesser, T. Contact Ion Pairs of Phosphate Groups in Water: Two-Dimensional Infrared Spectroscopy of Dimethyl Phosphate and ab Initio Simulations. *J. Phys. Chem. Lett.* **2019**, *10*, 6281–6286.
- (16) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33–38.
- (17) Huston, S. E.; Rossky, P. J. Free energies of association for the sodium-dimethyl phosphate ion pair in aqueous solution. *J. Phys. Chem.* **1989**, *93*, 7888–7895.
- (18) Kumar, M.; Simonson, T.; Ohanessian, G.; Clavaguera, C. Structure and thermodynamics of Mg:phosphate interactions in water: A simulation study. *ChemPhysChem* **2015**, *16*, 658–665.
- (19) Villa, F.; Mackerell, A. D.; Roux, B.; Simonson, T. Classical Drude Polarizable Force Field Model for Methyl Phosphate and Its Interactions with Mg²⁺. *J. Phys. Chem. A* **2018**, *122*, 6147–6155.
- (20) Grotz, K. K.; Cruz-León, S.; Schwierz, N. Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties. *J. Chem. Theory Comput.* **2021**, *17*, 2530–2540.
- (21) Panteva, M. T.; Giambasu, G. M.; York, D. M. Force Field for Mg²⁺, Mn²⁺, Zn²⁺, and Cd²⁺ Ions That Have Balanced Interactions with Nucleic Acids. *J. Phys. Chem. B* **2015**, *119*, 15460–15470.
- (22) Melcr, J.; Piquemal, J. P. Accurate Biomolecular Simulations Account for Electronic Polarization. *Frontiers Mol. Biosci.* **2019**, *6*, 1–8.
- (23) Duboué-Dijon, E.; Javanainen, M.; Delcroix, P.; Jungwirth, P.; Martinez-Seara, H. A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization. *J. Chem. Phys.* **2020**, *153*, 050901.
- (24) Oliveira, D. M. D.; Zukowski, S. R.; Palivec, V.; Hénin, J.; Martinez-Seara, H.; Ben-Amotz, D.; Jungwirth, P.; Duboué-Dijon, E. Binding of divalent cations to acetate: Molecular simulations guided by Raman spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24014–24027.
- (25) Halgren, T. A.; Damm, W. Polarizable force fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236–42.
- (26) Lemkul, J. A.; Huang, J.; Roux, B.; Mackerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983–5013.
- (27) Bedrov, D.; Piquemal, J. P.; Borodin, O.; Mackerell, A. D.; Roux, B.; Schröder, C. Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chem. Rev.* **2019**, *119*, 7940–7995.

- (28) Fyta, M.; Netz, R. R. Ionic force field optimization based on single-ion and ion-pair solvation properties: Going beyond standard mixing rules. *J. Chem. Phys.* **2012**, *136*, 1–11.
- (29) Yoo, J.; Aksimentiev, A. New tricks for old dogs: Improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8432–8449.
- (30) Kashfolgheta, S.; Verde, A. V. Developing force fields when experimental data is sparse: AMBER/GAFF-compatible parameters for inorganic and alkyl oxoanions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 20593–20607.
- (31) Hess, B.; Vegt, N. F. V. D. Cation specific binding with protein surface charges. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13296–13300.
- (32) Leontyev, I. V.; Stuchebrukhov, A. A. Electronic continuum model for molecular dynamics simulations. *J. Chem. Phys.* **2009**, *130*, 085102.
- (33) Leontyev, I.; Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2613–2626.
- (34) Kirby, B. J.; Jungwirth, P. Charge Scaling Manifesto: A Way of Reconciling the Inherently Macroscopic and Microscopic Natures of Molecular Simulations. *J. Phys. Chem. Lett.* **2019**, *10*, 7531–7536.
- (35) Cruz-León, S.; Grotz, K. K.; Schwierz, N. Extended magnesium and calcium force field parameters for accurate ion - nucleic acid interactions in biomolecular simulations. *J. Chem. Phys.* **2021**, *154*, 171102.
- (36) Bruce, E. E.; van der Vegt, N. F. A. Does an electronic continuum correction improve effective short-range ion-ion interactions in aqueous solution? *J. Chem. Phys.* **2018**, *148*, 222816.
- (37) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A. et al. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- (38) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084–2108.
- (39) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational design of particle mesh ewald compatible Lennard-Jones parameters for +2 metal cations in explicit solvent. *J. Chem. Theory Comput.* **2013**, *9*, 2733–2748.
- (40) Callahan, K. M.; Casillas-Ituarte, N. N.; Roeselová, M.; Allen, H. C.; Tobias, D. J. Solvation of magnesium dication: molecular dynamics simulation and vibrational spectroscopic study of magnesium chloride in aqueous solutions. *J. Phys. Chem. A* **2010**, *114*, 5141–5148.
- (41) Mamatkulov, S.; Fyta, M.; Netz, R. R. Force fields for divalent cations based on single-ion and ion-pair properties. *J. Chem. Phys.* **2013**, *138*, 024505.
- (42) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (43) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Comm.* **1995**, *91*, 43–56.
- (44) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.
- (45) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (46) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

- (47) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (48) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (49) Miyamoto, S.; Kollman, P. A. SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (50) Duboué-Dijon, E.; Mason, P. E.; Fischer, H. E.; Jungwirth, P. Hydration and Ion Pairing in Aqueous Mg²⁺ and Zn²⁺ Solutions: Force Field Description Aided by Neutron Scattering Experiments and Ab Initio Molecular Dynamics Simulations. *J. Phys. Chem. B* **2017**, *122*, 3296–3306.
- (51) Martinek, T.; Duboué-Dijon, E.; Štěpán Timr.; Mason, P. E.; Baxová, K.; Fischer, H. E.; Schmidt, B.; Pluharová, E.; Jungwirth, P. Calcium ions in aqueous solutions: Accurate force field description aided by ab initio molecular dynamics and neutron scattering. *J. Chem. Phys.* **2018**, *148*, 222813.
- (52) Vega, C.; Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663–19688.
- (53) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.
- (54) Grotz, K. K.; Schwierz, N. Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties in SPC/E, TIP3P-fb, TIP4P/2005, TIP4P-Ew, and TIP4P-D. *J. Chem. Theory Comput.* **2021**, *18*, 526–537.
- (55) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245–249.
- (56) Lemkul, J. A.; MacKerell, A. D. Balancing the Interactions of Mg²⁺ in Aqueous Solution and with Nucleic Acid Moieties For a Polarizable Force Field Based on the Classical Drude Oscillator Model. *J. Phys. Chem. B* **2016**, *120*, 11436–11448.
- (57) Phillips, J.; Hardy, D.; Maia, J.; Stone, J.; Ribeiro, J.; Bernardi, R.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W. et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 44130.
- (58) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (59) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.
- (60) Zhang, C.; Lu, C.; Wang, Q.; Ponder, J. W.; Ren, P. Polarizable multipole-based force field for dimethyl and trimethyl phosphate. *J. Chem. Theory Comput.* **2015**, *11*, 5326–5339.
- (61) Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J. P.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J. Chem. Theory* **2018**, *14*, 5273–5289.
- (62) Harger, M.; Li, D.; Wang, Z.; Dalby, K.; Lagardère, L.; Piquemal, J. P.; Ponder, J.; Ren, P. Tinker-OpenMM: Absolute and relative alchemical free energies using AMOEBA on GPUs. *J. Comput. Chem.* **2017**, *38*, 2047–2055.
- (63) Tuckerman, M. E.; Berne, B. J.; Martyna, G. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1990–2001.
- (64) Neese, F. The ORCA program system. *Wiley Interdisciplinary Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (65) Neese, F. Software update: the ORCA program system, version 4.0. *Wiley Interdisciplinary Rev.: Comput. Mol. Sci.* **2017**, *8*, e1327.

- (66) Ren, P.; Wu, C.; Ponder, J. W. Polarizable atomic multipole-based molecular mechanics for organic molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143–3161.
- (67) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian09 - Revision D01. 2013.
- (68) Stone, A. J.; Alderton, M. Distributed multipole analysis Methods and applications. *Mol. Phys.* **1985**, *56*.
- (69) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J. Chem. Phys.* **1988**, *89*, 3742–3746.
- (70) Hermans, J.; Shankar, S. The Free Energy of Xenon Binding to Myoglobin from Molecular Dynamics Simulation. *Israel J. Chem.* **1986**, *27*, 225–227.
- (71) Salari, R.; Joseph, T.; Lohia, R.; Hémin, J.; Brannigan, G. A Streamlined, General Approach for Computing Ligand Binding Free Energies and Its Application to GPCR-Bound Cholesterol. *J. Chem. Theory Comput.* **2018**, *14*, 6560–6573.
- (72) Duboué-Dijon, E.; Hémin, J. Building intuition for binding free energy calculations: bound state definition, restraints, and symmetry. *J. Chem. Phys.* **2021**, *154*, 204101.
- (73) Simonson, T.; Roux, B. Concepts and protocols for electrostatic free energies. *Mol. Sim.* **2016**, *42*, 1090–1101.
- (74) Deng, Y.; Roux, B. Computations of standard binding free energies with molecular dynamics simulations. *J. Phys. Chem. B* **2009**, *113*, 2234–2246.
- (75) Gumbart, J. C.; Roux, B.; Chipot, C. Standard binding free energies from computer simulations: What is the best strategy? *J. Chem. Theory Comput.* **2013**, *9*, 794–802.
- (76) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (77) Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S. et al. Jupyter Notebooks – a publishing format for reproducible computational workflows. 2016; pp 87–90.
- (78) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 1–10.
- (79) Chodera, J. D. A Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 1799–1805.
- (80) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169–9183.
- (81) Lelièvre, T.; Rousset, M.; Stoltz, G. Computation of free energy profiles with parallel adaptive dynamics. *J. Chem. Phys.* **2007**, *126*, 134111.
- (82) Lesage, A.; Lelièvre, T.; Stoltz, G.; Hémin, J. Smoothed Biasing Forces Yield Unbiased Free Energies with the Extended-System Adaptive Biasing Force Method. *J. Phys. Chem. B* **2017**, *121*, 3676–3685.
- (83) Hémin, J.; Fiorin, G.; Chipot, C.; Klein, M. L. Exploring multidimensional free energy landscapes using time-dependent biases on collective variables. *J. Chem. Theory Comput.* **2010**, *6*, 35–47.
- (84) Comer, J.; Gumbart, J. C.; Hémin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* **2015**, *119*, 1129–1151.
- (85) Fiorin, G.; Klein, M. L.; Hémin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **2013**, *111*, 3345–3362.
- (86) Hémin, J. Fast and accurate multidimensional free energy integration. *J. Chem. Theory Comput.* **2021**, *17*, 6789–6798.

- (87) Luo, H.; Sharp, K. On the calculation of absolute macromolecular binding free energies. *Proc. Natl. Acad. Sci.* **2002**, *99*, 10399–10404.
- (88) Mey, A. S. J. S.; Allen, B.; Macdonald, H. E. B.; Chodera, J. D.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A. et al. Best Practices for Alchemical Free Energy Calculations. *Living J. Comp. Mol. Sci.* **2020**, 1–51.
- (89) Sigel, R. K.; Sigel, H. A stability concept for metal ion coordination to single-stranded nucleic acids and affinities of individual sites. *Acc. Chem. Res.* **2010**, *43*, 974–984.
- (90) Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.
- (91) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046–4063.
- (92) Kolev, S. K.; Petkov, P. S.; Rangelov, M. A.; Trifonov, D. V.; Milenov, T. I.; Vayssilov, G. N. Interaction of Na⁺, K⁺, Mg²⁺ and Ca²⁺ counter cations with RNA. 2018.
- (93) Neumann, J.; Schwierz, N. Artificial Intelligence Resolves Kinetic Pathways of Magnesium Binding to RNA. *J. Chem. Theory Comput.* **2022**, *18*, 1202–1212.

TOC Graphic

