Supporting Information 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

Additional computational details

DMP force field

We followed the GAFF procedure to develop a GAFF forcefield for DMP. The employed atom types and charges are listed Table S1. The resulting force field presents charges on the phosphate moieties that are very similar to that of the phosphate backbone in recent Amber nucleic acid force fields. Our DMP model should thus be a reasonable proxy for ion binding to nucleic acids. The corresponding parameter files are provided together with the rest of the inputs in a public Zenodo folder.¹

Table S1: GAFF force field for DMP. Charges of analogous atoms, when relevant, in the recent Amber ff99+bsc0 force field for nucleic acids are provided for comparaison. For the briging oxygens O1/O4, both values for the analogous O3' and O5' atoms are reported.

atom index	atom name	atom type	Charge (e)	Charge DNA
1	C1	CT	0.044479	
2	H1	HC	0.040363	
3	H2	HC	0.040363	
4	H3	HC	0.040363	
5	O1	OSM2P	-0.469008	-0.4954 / -0.5232
6	P1	p5	1.181336	1.16590
7	O2	OM2P	-0.787228	-0.77610
8	O3	OM2P	-0.787228	-0.77610
9	O4	OSM2P	-0.469008	-0.4954 / -0.5232
10	C2	CT	0.044479	
11	H4	HC	0.040363	
12	H5	HC	0.040363	
13	H6	HC	0.040363	

Restraint potentials

During the alchemical binding free energy calculations, the geometry of the ion pair was maintained in a defined binding mode (monodentate, bidentate or solvent-shared) using flat-well harmonic restraints on the distance between the DMP phosphorus atom and the cation: the ion:P distance varied freely in a range [a,b]; outside this range, the restraint energy increased harmonically, with a force constant of k = 500 kcal mol⁻¹ Å⁻¹. The range of distances used to define the flat bottom interval was determined based on short equilibrations without restraints starting from different bound geometries.

For Mg²⁺, with non polarizable force fields, the flat-bottomed interval [a,b] was typically 2.7 Å to 3.2 Å for the contact bidentate binding mode (when metastable at all), 3.2 Å to 3.7 Å for the contact monodentate, and 4.1 Å to 6.1 Å for the solvent-shared binding mode. With AMOEBA, the intervals were slightly shifted to closer distances for the contact binding mode: 2.9 Å to 3.6 Å for the contact monodentate ion pair.

As expected from their relative sizes, the flat-bottomed interval [a,b] for Ca^{2+} was slightly larger than with Mg²⁺, typically 2.6 Å to 3.4 Å for the contact bidentate binding mode (when metastable at all), 3.4 Å to 4.1 Å for the contact monodentate, and 4.3 Å to 6.5 Å for the solvent-shared binding mode.

Binding free energies

The error bars reported in Fig 2 of the main manuescript were obtained by summing the error estimated by "gmx bar" for the two alchemical transformations. In order to verify the accuracy of this error estimate, we performed (using the calcium system with the ECC force field with 0.8 scaling) three independent replicas of the alchemical transformations (both in water and in the monodentate ion pair). The standard deviation observed between the replicas (see Table S2) is consistent with the estimated error from bar (0.2 kJ/mol for the isolated ion, 0.3 kJ/mol for the monodentate ion pair). We also checked that the difference between the free energy estimates obtained from forward / backward transformations (Table S2) also falls within the reported error.

The numerical values for the binding free energies corresponding to Fig 2 of the main manuscript are provided in Table S3 for force fields without explicit polarization and Table S4 for polarizable force fields.

Table S2: Free energy $(kJ \text{ mol}^{-1})$ associated with the alchemical transformations for the decoupling of a calcium cation in water or in a monodentate ion pair with DMP (force field "ECC 0.8", see Table 1 main text), for each of 3 replicas of the same transformation, as well as for a transformation performed in the "backward" direction.

	Ca^{2+} in water	monodentate ion pair
replica 1	806.79	808.44
replica 2	806.76	808.29
replica 3	806.83	808.75
average	806.79	808.49
standard deviation	0.10	0.26
backward - replica 1	-806.69	-808.23

Additional force fields. The free energy associated with Mg^{2+} and Ca^{2+} binding to DMP was computed with different force fields. In addition to those presented in Table 1 of the manuscript, we investigated how the binding free energy depended on the exact cation parametrization and of a change of water model. Table S5 summarizes the force field combinaison employed in all the new simulations.

In addition to the 12-6 Lennard-Jones parameters suggested by the Merz group⁴ used in the manuscript, we also computed $\Delta G^{\circ}_{\text{bind}}$ with the Mg²⁺ force field (labelled "Mg_c" in Table S6) suggested by D. Tobias and coworkers,⁵ and with the Ca²⁺ force field developed in the Netz group⁶ ("Ca_n" in Table S6). Also, an alternative force field using pair-specific Lennard Jones parameters was recently suggested for Mg²⁺ ("Mg2"),⁷ as well as a "fast exchange" ("nMg") version with fast nanosecond exchange of water from the hydration shell,⁸ which can be an advantage to accelerate sampling.

The strong overbinding exhibited by standard full charges non polarizable force field proves independent of the exact set of ionic parameters. For Ca^{2+} , it is even more pronounced (on the same order of magnitude as with Mg^{2+}) with parameters from the Netz group.⁶ As for force fields based on pair-specific Lennard Jones parameters for the magnesium–phosphate interaction, it is worth noting that depending on the exact parametrization (with different targets in mind), the overall binding free energy varies between -4.5 and -8.9 kJ/mol for Mg^{2+} , with key differences in the relative free energies of the different binding modes: the Table S3: 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 $\Delta G_{\text{bind}}^{\circ}$, computed with different force fields, as described in the Methodology section. The reference experimental binding free energies $\Delta G_{\text{bind,expt}}^{\circ}$ are -7.1 kJ/mol for Mg²⁺² and -5.9 kJ/mol for Ca²⁺.³ The reported binding distance (in Å) is the first peak of the radial distribution function between the DMP non bridging oxygen atoms and the cation.

Force field	bidentate	monodentate	SShIP	Total contact	Total	Binding
						distance
${f Mg}^{2+}$						
full charges	-28.0	-31.2	-5.9	-31.8	-31.8	1.93
ECC (0.75 scaling)	unstable	+2.8	-0.7	+2.8	-1.3	1.97
ECC (0.8 scaling)	unstable	-1.1	-3.0	-1.1	-4.5	1.93
pair specific LJ	unstable	-2.7	-4.9	-2.7	-5.8	2.06
\mathbf{Ca}^{2+}						
full charges	-11.9	-6.2	-4.4	-11.9	-12.3	2.34
ECC (0.75 scaling)	+6.6	+4.0	-1.5	+3.3	-1.9	2.34
ECC (0.8 scaling)	+4.1	+1.1	-1.4	+0.5	-2.4	2.36
pair specific LJ	unstable	-5.3	-6.5	-5.3	-7.7	2.28

Table S4: 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 $\Delta G_{\text{bind}}^{\circ}$, computed with different force fields, as described in the Methodology section. The reference experimental binding free energies $\Delta G_{\text{bind,expt}}^{\circ}$ are -7.1 kJ/mol for Mg²⁺² and -5.9 kJ/mol for Ca²⁺.³ The reported binding distance (in Å) is the first peak of the radial distribution function between the DMP non bridging oxygen atoms and the cation.

Force field	bidentate	monodentate	SShIP	Total contact	Total	Binding
						distance
${f Mg}^{2+}$						
Drude–NBFIX	unstable	+0.4	-10.7	+0.4	-10.7	2.05
AMOEBA	unstable	-20.7	-7.8	-20.7	-20.8	1.97
AMOEBA modified	unstable	+2.8	-7.4	+2.8	-7.5	2.06
\mathbf{Ca}^{2+}						
Drude (no NBFIX)	-73.2	unstable	-6.5	-73.2	-73.2	2.16
AMOEBA	unstable	-10.7	-5.3	-10.7	-11.0	2.27
AMOEBA modified	unstable	+0.4	-5.7	+0.4	-5.9	2.29

"Mg2" force field favors contact monodentate ion pairs against SShIP, while the opposite is true for the other parametrizations.

Table S5: Combination of force fields used in each additional simulation with non explicitly polarizable force fields, using labels for each force fields as defined in the text, where we provide the corresponding references.

Cation	water	DMP
Mg^{2+}		
Mg_c	SPC/E	GAFF
Mg_mt	TIP3P	GAFF
Mg_s2	SPC/E	$GAFF_s2$
Mg2	TIP3P	GAFF
nMg	TIP3P	GAFF
Mg_s	SPC/E	GAFF
Mg_s2	SPC/E	GAFF
$\mathbf{C}\mathbf{a}^{2+}$		
Ca_n	SPC/E	GAFF
Ca_mt	TIP3P	GAFF
Ca_s2	TIP3P	$GAFF_s2$
Ca_s	SPC/E	GAFF
Ca_s2	$\mathrm{SPC/E}$	GAFF

Table S6: Cation-DMP binding free energy $\Delta G_{\text{bind}}^{\circ}$ (kJ mol⁻¹) for different ion pair conformations (contact monodentate, contact bidentate and solvent shared), together with the overall $\Delta G_{\text{bind}}^{\circ}$, computed with different force fields, with the specific force field combinaison detailed in Table S5. The reference experimental binding free energies $\Delta G_{\text{bind,expt}}^{\circ}$ are -7.1 kJ/mol for Mg²⁺² and -5.9 kJ/mol for Ca²⁺.³ The error bars on the computed $\Delta G_{\text{bind}}^{\circ}$ are about 1–2 kJ/mol.

Force field	bidentate	monodentate	SShIP	Total
${f Mg^{2+}}$				
Mg_c	-38.5	-30.0	-5.3	-38.6
$Mg_mt + TIP3P$	-38.7	-37.3	-0.99	-39.9
$Mg(ECC \ 0.8 \ scaling) + TIP3P$	+2.57	-3.73	+0.33	-4.3
Mg2	unstable	-8.6	-3.9	-8.9
nMg	unstable	-0.7	-3.9	-4.5
$Mg(ECC \ 0.75 \ scaling) + DMP(full)$	-23.2	-11.7	-3.1	-23.2
$Mg(ECC \ 0.8 \ scaling) + DMP(full)$	-28.3	-17.1	-4.6	-28.3
\mathbf{Ca}^{2+}				
Ca_n	-28.9	-16.9	-4.8	-28.9
$Ca_mt + TIP3P$	-16.3	-6.4	-1.3	-16.3
$Ca(ECC \ 0.8 \ scaling) + TIP3P$	+5.1	-2.1	-0.6	-1.6
Ca(ECC 0.75 scaling)+DMP(full)	-4.3	-3.1	-3.8	-6.5
$Ca(ECC \ 0.8 \ scaling) + DMP(full)$	-8.7	-3.4	-4.1	-9.3

Mixed ECC-full charges tests. Even if the ECC theory prescribes scaling the charges of both the cation and phosphate moieties, it could be tempting, for applications in large scale biochemical systems, to scale only the cation, leaving the biomolecular force field untouched. Such mixed approaches have been shown earlier on cation-acetate interaction to be insufficient to yield correct binding free energies. Here again, scaling only the Mg^{2+} cation still leads to strong overbinding to DMP (Table S6). This is less true with Ca^{2+} , where the overall binding free energy with a mixed strategy would be only little too strong with respect to experimental data, especially with the 0.75 scaling. However note that this mixed approach then predicts the contact monodentate ion pair to be the most stable, which contrasts with all other force fields (among those that correctly capture the total binding free energy).

Water model. On a few selected cases, we examined how a change in the water model (TIP3P, often used with GAFF, rather than SPC/E) impacts the computed binding free energies. For the full charge force fields proposed by the Merz group, we picked the TIP3Pspecific cation parametrizations ("Mg mt" and "Ca mt"), which we now combined with GAFF for DMP and TIP3P water (see Tables S4 and S5). As can be seen when comparing the TIP3P results (Table S5) with those obtained with SPC/E initially presented (Table S2), this change in the water model (and associated cation parameters) changes by a few kJ/mol the binding free energies for each ion pair, but the overall picture remains the same, with the contact ion pairs strongly overbinding. Doing the same with our best ECC force field (scaling 0.8), we see similar changes in the binding free energies. However, if the relative ordering of the ion pairs is not changed with Ca²⁺, the monodentate ion pair becomes more stable than the SShIP with Mg^{2+} . Note however, that the ECC ion parameters were derived specifically for SPC/E, and we did not derive new parameters for TIP3P here. Hence, we tend to rely more on our initial calculations, even if this test points out that the relative stability of the ion pairs can be sensitive to changes in the water force fields, especially given the small free energy differences between states.

Refinement of the AMOEBA force field for phosphate– cation binding

Binding free energies

We examined the sensitivity of cation (starting with Mg^{2+})-DMP binding free energies with respect to different parameters in the AMOEBA force field. First, we refined new multipoles for DMP using clusters with explicit water molecules, that should thus be well suited for simulation of DMP in solution. With these modified multipoles, we computed the associated binding free energy (Table S7) for the monodentate and SShIP ion pairs (the contact bidentate is not stable with any of the parameters). The binding free energy is strikingly similar with both sets of multipoles.

Observing that the non bridging oxygen polarizability was modified in the DMP force field⁹ compared to other oxygen atoms (in particular that of water), and that the employed polarizabilities for the non bridging DMP oxygen and phosphorus atom ($\alpha_O = 1.724$ Å³ and $\alpha_P = 1.788$ Å³ respectively) strongly differed from that used in an early study on DHP ($\alpha_O = 0.837$ Å³ and $\alpha_P = 3.72$ Å³, respectively), we computed the binding free energy for different values and combinations of non bridging-O and P polarizabilities (Table S7). The binding free energy associated with the contact monodentate ion pair is mainly determined by the non bridging oxygen polarizability, with $\Delta G^{\circ}_{\text{bind}}$ changing by more than 20 kJ/mol when modifying $\alpha_O =$ from 1.724 to 0.837 Å³. In contrast, a change in the phosphorus polarizability has very little impact on the binding free energy.

Benchmark against QM calculations on clusters

In the spirit of the development of the AMOEBA force field, we examined whether optimal polarizability values could be rationally chosen from single point energy calculations on small-medium size clusters, rather than from the quite costly systematic calculation of $\Delta G_{\text{bind}}^{\circ}$.

Table S7: Mg^{2+} -DMP binding free energy ΔG_{bind}° (kJ mol⁻¹) for different ion pair conformations (contact monodentate and solvent shared), together with the overall ΔG_{bind}° , computed with different modified versions of the AMOEBA force field. Polarizabilities of the DMP non bridging oxygen atom, α_O , and of the phosphorus atom, α_P are provided in Å³. The reference experimental binding free energy $\Delta G_{bind,expt}^{\circ}$ is -7.1 kJ/mol for Mg²⁺.² The error bars on the computed ΔG_{bind}° are about 1–2 kJ/mol.

Force field	monodentate	SShIP	Total
original ($\alpha_O = 1.724, \alpha_P = 1.788$)	-20.7	-7.8	-20.8
new multipoles	-21.1	XXX	-21.1
$\alpha_O = 0.837, \alpha_P = 3.72$	+2.3	-7.8	-7.9
$\alpha_O = 0.837, \alpha_P = 1.788$	+3.4	-6.5	-6.5
$\alpha_O = 1.724, \alpha_P = 3.72$	-24.1	-8.1	-24.1
$\alpha_O = 0.837, \alpha_P = 3.3$	+2.8	-7.4	-7.5
$\alpha_O = 1.3, \alpha_P = 3.3$	-10.2	-7.6	-10.9

As described in the main text, we optimized with DFT (def2-SVP/ ω B97X-D3(BJ)) the geometry of 40 clusters (20 in the monodentate and 20 in the solvent-shared ion pair geometry) made of one DMP anion, one Mg²⁺, and 20 water molecules. Single point energies were then calculated on these optimized geometries, both with DFT and a larger basis set (6-311++G(2d,2p) as in Ref. 9), which proved key to obtain correct interaction energies, and with different versions of the AMOEBA force field, systematically varying the polarizabilities α_O and α_P of the non-bridging oxygens and phosphorus atoms of DMP. The error with respect to the reference DFT calculations was quantified with the standard deviation on the 40 snapshots between QM and AMOEBA energies.

We systematically quantified the error between the reference DFT energies and AMOEBA with varied polarizabilities, both for the total relative energy of the snaphots (E_{rel} , Table S8), and for the interaction energy, defined as $E_{int} = E_{tot} - E_{DMP} - E_{cation+water}$, both absolute (E_{int} , Table S9) and relative between snapshots ($E_{int,rel}$, Table S10), the energies being aligned on the average interaction energy of the monodentate snaphots.

It is important to notice that the set of polarizabilities that minimize the error between AMOEBA and DFT energies highly depends on the chosen energy. While examination of the total relative energy of the snapshots yields a minimal error for $\alpha_O = 1.5$ Å³ and

Table S8: Standard deviation (kcal/mol) between DFT and AMOEBA relative total single point energies computed with different modified versions of the AMOEBA force field. Polarizabilities of the DMP non bridging oxygen atom, α_O , and of the phosphorus atom, α_P , are provided in Å³.

α_P	1.788	2.3	2.8	3.3	3.72
1.724	4.40	4.32	4.25	4.20	4.17
1.5	4.30	4.22	4.16	4.12	4.10
1.3	4.31	4.24	4.45	4.43	4.15
1.0	4.54	4.48	4.45	4.24	4.44
0.837	4.79	4.74	4.71	4.71	4.72

Table S9: Standard deviation (kcal/mol) between DFT and AMOEBA interaction energies computed with different modified versions of the AMOEBA force field. Polarizabilities of the DMP non bridging oxygen atom, α_O , and of the phosphorus atom, α_P , are provided in Å³.

α_P	1.788	2.3	2.8	3.3	3.72
1.724	13.50	13.07	12.62	12.16	11.76
1.5	13.38	12.93	12.47	11.99	11.76
1.3	13.35	12.88	12.45	11.11	11.46
1.0	13.48	12.97	12.45	11.92	11.46
0.837	13.70	13.15	12.61	12.05	11.58

Table S10: Standard deviation (kcal/mol) between DFT and AMOEBA relative interaction energies computed with different modified versions of the AMOEBA force field. Polarizabilities of the DMP non bridging oxygen atom, α_O , and of the phosphorus atom, α_P , are provided in Å³.

$\boxed{\begin{array}{c} \alpha_P \\ \alpha_O \end{array}}$	1.788	2.3	2.8	3.3	3.72
1.724	3.42	3.34	3.28	3.24	3.23
1.5	3.37	3.29	3.23	3.20	3.19
1.3	3.32	3.24	3.12	3.12	3.16
1.0	3.25	3.17	3.12	3.10	3.11
0.837	3.21	3.13	3.09	3.08	3.09

 $\alpha_P = 3.72$ Å³, examination of the interaction energies would give as optimal parameters $\alpha_O = 1.3$ Å³ and $\alpha_P = 3.3$ Å³. Finally, looking instead at the relative interaction energies yields optimal values of $\alpha_O = 0.837$ Å³ and $\alpha_P = 3.3$ Å³. Not surprisingly given the difference in α_O which is the most sensitive parameter, these 3 optimal sets of polarizabilities yield

highly different binding free energies for the contact monodentate ion pair (Table S7), only the last one ($\alpha_O = 0.837$ Å³ and $\alpha_P = 3.3$ Å³) yielding an overall binding free energy in agreement with experimental estimates.

Hence, it is important to note that the polarizabilities that minimize the error on the relative energy of the snapshots do not yield good binding energies, so minimize the error on the relative energy computed on small clusters does not guarantee a proper binding behavior. From our tests, only examination of the relative interaction energy between snapshots in different ion pair geometries yields accurate binding free energies.

References

- (1) Zenodo dataset DOI:10.5281/zenodo.5971002.
- (2) Verbeeck, R. M.; De Bruyne, P. A.; 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.
- (3) Mcdowell, H.; Brown, W. E.; Sutter, J. R. Solubility Study of Calcium Hydrogen Phosphate. Ion-Parr Formation. *Inorg. Chem.* 1971, 10, 1638–1643.
- (4) 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.
- (5) 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.
- (6) 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.

- (7) 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.
- (8) Grotz, K. K.; Cruz-Leon, S.; Schwierz, N. Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water- Exchange Properties. J. Chem. Theor. Comput. 2021, 17, 2530–2540.
- (9) 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.