THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

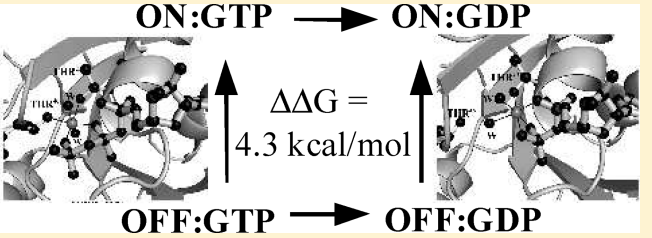
Free Energy Simulations of a GTPase: GTP and GDP Binding to Archaeal Initiation Factor 2

Priyadarshi Satpati,[†] Carine Clavaguéra,[‡] Gilles Ohanessian,[‡] and Thomas Simonson*[,†]

[†]Laboratoire de Biochimie (CNRS UMR7654), Department of Biology, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex (France).

[‡]Laboratoire des Mécanismes Réactionnels (CNRS), Department of Chemistry, Ecole Polytechnique, CNRS, 91128 Palaiseau Cedex (France).

ABSTRACT: Archaeal initiation factor 2 (aIF2) is a protein involved in the initiation of protein biosynthesis. In its GTP-bound, "ON" conformation, aIF2 binds an initiator tRNA and carries it to the ribosome. In its GDP-bound, "OFF" conformation, it dissociates from tRNA. To understand the specific binding of GTP and GDP and its dependence on the ON or OFF conformational state of aIF2, molecular dynamics free energy simulations (MDFE) are a tool of choice. However, the validity of the computed free energies depends on the simulation model, including the force field and the boundary conditions, and on the extent of conformational sampling in the simulations. aIF2 and other GTPases present specific difficulties; in particular, the nucleotide ligand coordinates a divalent Mg^{2+} ion, which can polarize the electronic distribution of its environment. Thus, a force field with an explicit treatment of electronic polarizability could be necessary, rather than a simpler, fixed charge force field. Here, we begin by comparing a fixed charge force field to quantum chemical calculations and experiment for Mg^{2+}:phosphate binding in solution, with the force field giving large errors. Next, we consider GTP and GDP bound to aIF2 and we compare two fixed charge force fields to the recent, polarizable, AMOEBA force field, extended here in a simple, approximate manner to include GTP. We focus on a quantity that approximates the free energy to change GTP into GDP. Despite the errors seen for Mg^{2+}:phosphate binding in solution, we observe a substantial cancellation of errors when we compare the free energy change in the protein to that in solution, or when we compare the protein ON and OFF states. Finally, we have used the fixed charge force field to perform MDFE simulations and alchemically transform GTP into GDP in the protein and in solution. With a total of about 200 ns of molecular dynamics, we obtain good convergence and a reasonable statistical uncertainty, comparable to the force field uncertainty, and somewhat lower than the predicted GTP/GDP binding free energy differences. The sign and magnitudes of the differences can thus be interpreted at a semiquantitative level, and are found to be consistent with the experimental binding preferences of ON- and OFF-aIF2.



1. INTRODUCTION

Proteins can often occupy different conformations whose population is controlled by the binding of specific ligands. This allows ligand-induced conformational switching, which can underlie important biophysical and biochemical processes, such as allostery, energy transduction, signaling, gene regulation, and catalysis.[1−3] ATPases and GTPases are two important classes of such proteins, which play essential roles in the transduction of energy and information in biological cells.[4,5] They are capable of hydrolyzing ATP (or GTP) into ADP (GDP), and they typically occupy different conformations depending on which ligand (ATP/GTP or ADP/GDP) is bound. A well-known example is ATP synthase,[6,7] which produces most of the ATP in cells by cycling between several conformations, some of which are capable of converting ADP to ATP, and whose populations depend on the motion of protons between a series of specific sites. Other examples are given by several GTPases that act as molecular switches to help control and coordinate protein biosynthesis.[8−10] In this work, as a paradigm for this class of proteins, we have chosen the archaeal initiation factor 2, or aIF2.[9] Here, "initiation" refers to the first steps in decoding an mRNA in order to synthesize the corresponding protein. Indeed, aIF2 is largely responsible for recruiting the first, initiator tRNA to the ribosome and positioning it correctly, in register with the start codon of the ribosome-bound mRNA.[10] Initiation is the most highly regulated step in protein biosynthesis, and correct positioning of the initiator tRNA by aIF2 is essential for reading mRNA in the correct frame, or register. Like other GTPases, aIF2 cycles between an ON state, stabilized by GTP and competent to bind its partners (the ribosome and tRNA) and an incompetent, OFF state, stabilized by GDP. Analyzing the molecular interactions that control the selective, conformation-dependent binding of GTP and GDP is thus essential to understand the function of aIF2, and the same is true for all GTPases and ATPases.

Molecular dynamics (MD) and Monte Carlo simulations are powerful tools for studying both protein:ligand interactions and

Received: February 28, 2011
Revised: April 13, 2011
Published: May 02, 2011

conformational dynamics.[11] They have been extensively applied to complexes between proteins and nucleoside di- or triphosphates, such as GTP. Thus, MD has been used to study the GTPases Ras, RhoC, and Rab5a and their complexes with GTP or GDP. Classical force fields with fixed atomic charges were used to characterize structural transitions and identify highly populated conformations,[12−16] and to study the orientation of a catalytic water in Ras.[17] MD simulations were also applied to the protein kinase PAK1 to identify key functional interactions with bound GTP[18] and to study ADP release from protein kinase A.[19,20] Two systems that have been extensively studied are the molecular motors myosin and ATP synthase. Conformational transitions in both systems have been studied,[21−25] using a wide range of techniques for conformational exploration, all involving classical mechanical models with fixed atomic charges. In contrast, hybrid, quantum mechanical/molecular mechanics models (QM/MM) were used to study the catalysis of ATP hydrolysis by myosin and GTP hydrolysis by Ras.[26−28]

Another class of applications has focused on binding free energies and free energy differences between ligands.[29−32] Thus, alchemical MD free energy simulations (or MDFE) were used to compare the binding of several GTP analogues to the cell division protein FTSZ,[33] and to compare ATP and ADP binding to an aminoacyl-tRNA synthetase.[34,35] ATP synthase was studied in depth by Karplus and co-workers.[36,37] Their MDFE simulations compared ATP and ADP binding at different sites of the ATPase, and were validated by comparison to an extensive body of experimental data. All of these MDFE studies used classical mechanical, fixed charge force field models.

In this work, we examine whether MDFE with a classical mechanical, fixed charge force field is an appropriate tool to understand the interactions that control selective nucleotide binding to aIF2. aIF2 is composed of three subunits, α, β, and γ. The γ subunit forms the core of the heterotrimer. It contains two loops of 20 amino acids each; the conformational changes that define the ON and OFF states are mainly localized in these loops, which are called switch 1 and switch 2 (sw1, sw2). GTP and GDP bind in a pocket at the surface of the γ subunit, making contact with both sw1 and sw2. A Mg^{2+} ion coordinates the phosphate groups—the β and γ groups in the GTP complex and the β group with GDP. Thus, replacing GTP by GDP involves the removal of one Mg^{2+}:phosphate interaction. This is one of the potential difficulties we have to address: can a fixed charge force field accurately capture the balance between GTP and GDP binding when electrostatic interactions with a divalent ion are involved? The divalent ion is expected to polarize its electronic environment.[38−41] Such electronic polarization is not represented explicitly with a fixed charge force field; it is represented implicitly, and this may not be sufficiently accurate. A second difficulty arises from the long relaxation times associated with long-range electrostatic interactions in a heterogeneous system of this complexity:[42] we expect that, to obtain converged and precise MDFE results, a great deal of conformational sampling will be needed. This, in turn, makes it attractive to use a simplified model for distant portions of the protein, specifically, a dielectric continuum model.[30,43−46] Such hybrid models, with a fully atomistic region and another, simplified region have a long history.[42,47−49] Our preferred variant,[46] though generally applicable, has not been widely used and will have to be specifically tested for the present system. One last difficulty arises whenever one compares a highly populated binding state (such as ON-aIF2 with bound GTP) to a weakly populated state (such as ON-aIF2 with bound GDP): the minor complex does not have a known crystal structure and must be modeled.

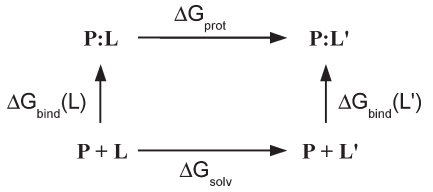


Figure 1. Thermodynamic cycle for protein:ligand binding. Vertical legs correspond to binding; horizontal legs correspond to the alchemical transformation of the ligand from L into L′, either in the solvated protein (above) or in solution (below). The L/L′ binding free energy difference is $\Delta\Delta G = \Delta G_{prot} - \Delta G_{solv} = \Delta G_{bind}(L') - \Delta G_{bind}(L)$.

We address nucleotide binding to aIF2 using the thermodynamic cycle in Figure 1. For each conformational state, ON or OFF, we compare the binding free energies of GTP and GDP. In practice, we use MD simulations to follow the horizontal legs of the cycle, transforming GTP into GDP alchemically and reversibly.[50,51] Our simulation model considers a spherical subset of the protein, centered on the ligand, and treats it in atomic detail. This spherical subset is solvated by a cubic water box, and simulated with periodic boundary conditions using particle mesh Ewald summation for long-range electrostatic interactions. More distant parts of the protein (beyond 26 Å) are reintroduced in a second step, treated as a dielectric continuum;[46] see Methods. Most of the simulations are done with the Charmm27 force field, which employs fixed atomic charges, modeling electronic polarization in an implicit way.[52] With this force field, we assume that the effect of electronic polarization by the divalent Mg^{2+} will approximately cancel when the horizontal legs of the cycle are subtracted. We refer to this assumption as the "additivity" hypothesis. With the fixed charge force field, we can run long simulations, and investigate the convergence of the MDFE results in some detail. By collecting a total of about 200 ns of MD, the statistical uncertainty for the double free energy difference, $\Delta G_{GTP} - \Delta G_{GDP}$, is reduced to an acceptable level, about 2−3 kcal/mol. The contribution of distant protein regions, treated here as a dielectric continuum, is shown to be much less than this uncertainty.

To test the additivity hypothesis, selected calculations are done with a recent, high quality force field that includes electronic polarizability explicitly: the AMOEBA force field, implemented in the Tinker program.[53,54] This force field uses not only atomic charges but also dipoles and quadrupoles on each atom, along with atomic dipolar polarizability. The force field was extended here in a simple, *ad hoc* way to include GTP (see Methods). Calculations are also done with a second fixed charge force field: a recent variant of the Amber force field.[55] Before considering the Mg^{2+}:phosphate interactions in the protein, we study small model systems in solution, using both the fixed charge force field and quantum mechanical calculations to compute Mg^{2+}:phosphate binding free energies. Although many authors have applied quantum calculations to Mg^{2+}:phosphate interactions, it appears that none of them have computed binding free energies and compared them to experiment. The quantum mechanical treatment is accurate for a singly charged phosphate but gives significant errors for dianionic phosphate. The fixed charge force field gives even larger errors, overestimating the binding free energy of Mg^{2+} to inorganic phosphate in solution by an order of magnitude, compared to experiment (see Results).

However, when we compare GTP and GDP binding to aIF2, by subtracting the horizontal legs of our thermodynamic cycle, the double free energy difference $\Delta G_{GTP} - \Delta G_{GDP}$ depends only moderately on the force field employed. When the protein ON and OFF states are compared, in particular, the fixed charge Charmm27 and Amber force fields and the polarizable AMOEBA force field all give results that agree within statistical uncertainty, about 2−3 kcal/mol. This suggests that the additivity hypothesis is reasonable.

All of these results indicate that MDFE with a classical, fixed charge force field, a hybrid atomistic/continuum MD setup, and long simulations lead to results that are reasonably accurate and precise. In particular, the overall uncertainty is estimated to be smaller than the binding free energy difference between GTP and GDP, for both the ON and OFF states. The difficulties considered here—force field dependency of MDFE when Mg^{2+}:phosphate interactions are involved, effect of long-range interactions and boundary conditions, conformational sampling, and convergence—are relevant for many GTPases and ATPases. We expect that the present conclusions will be valid for these proteins, so that the simulation strategy presented here (and similar strategies used earlier[33,34,36,37]) will be of general interest for this class of proteins.

This Article is organized as follows. The Methods section describes the MD setup, the MDFE protocols, the quantum mechanical treatment, the strategy used to compare the fixed charge and polarizable force fields, and the continuum electrostatic treatment of distant protein groups. In Results, we first describe the small model systems: the quantum mechanical study of Mg^{2+}:phosphate binding, the calculation of the binding free energy using MDFE and a fixed charge force field, and an MDFE study of Mg^{2+} binding to GTP and GDP. Next, we describe our analysis of the force field dependency of ligand binding. Then, we present the calculation of the contribution of the distant protein region to the GTP/GDP binding free energy difference. Finally, we present MDFE results for GTP/GDP binding to aIF2 and describe the convergence of the simulations. The last section is a discussion.

2. METHODS

2.1. Protein MD Simulations. Structures of aIF2 from *Sulfolobus solfataricus* in its ON and OFF conformations were taken from the Protein Data Bank (entries 2AHO and 2QN6; crystallographic resolution of 3.0 and 2.15 Å).[9,56] Our simulation model included protein residues having at least one non-hydrogen atom within a 26 Å radius sphere, centered at the center of the nucleotide, GTP or GDP. In addition to crystal waters within the binding pocket (2 and 9 water molecules for the ON and OFF states, respectively), a cubic water box with a 74 Å edge was overlaid and waters that overlapped protein were removed. This box size corresponds to a low protein concentration of about 4 mM. The total number of water molecules in our MD models are 12007 and 12067 for the ON and OFF states, respectively. The numbers of non-hydrogen protein atoms in our truncated protein models are 1515 and 1432 for the ON and OFF states, respectively (see Figure 2). The total number of atoms in both simulation models is about 39080. Protein atoms between 22 and 26 Å from the sphere's center were harmonically restrained to their experimentally determined positions, with force constants increasing from 0.1 to 1.5 kcal/mol/Å^2 as one moves closer to the outer boundary; a few ionized side chains at the boundary have a

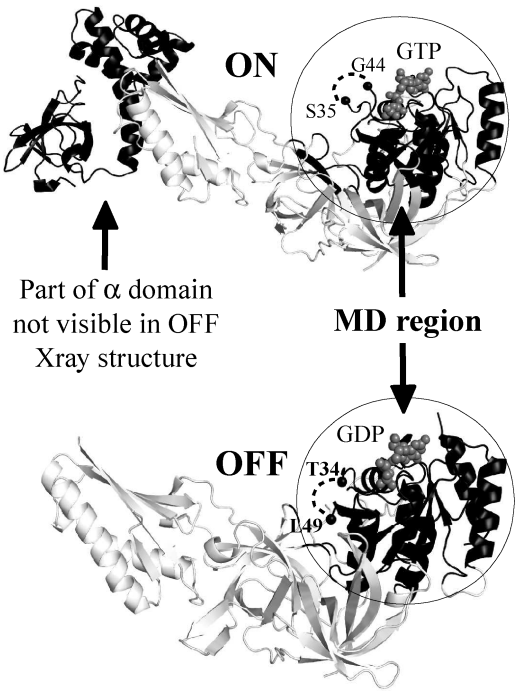


Figure 2. Crystal structures of ON and OFF aIF2. The spherical region included explicitly in the MD model is indicated and colored black; a part of the α domain not seen in the OFF X-ray structure is also indicated. The GTP and GDP ligands are shown (CPK view, colored gray) and labeled. Missing switch 1 residues are indicated by a dashed line; the adjacent residues are shown as spheres and labeled: Ser35 and Gly44 (ON state); Thr34 and Leu49 (OFF state).

stronger force constant of 10 kcal/mol/Å^2. Simulations were performed with periodic boundary conditions, using the particle mesh Ewald method[57] for long-range electrostatics, with tinfoil boundary conditions.[58,59] The van der Waals interaction was spatially truncated at a 16 Å cutoff distance. Five sodium counterions were added to reduce the overall charge of our MD models (for both the ON and OFF states), giving overall charges of −2, −1, 0, 1, or 2, depending on the system. The temperature and pressure of the system were maintained at 295 K and 1 bar. The temperature was controlled by using Langevin dynamics for the protein and solvent atoms other than hydrogens, with a coupling coefficient of 5 ps^{-1}; pressure was controlled by a Langevin piston Nose−Hoover method. Covalent bond lengths were held fixed, and the MD time step was 2 fs. The CHARMM27 force field was used for the protein and the GTP/GDP ligands[52] and the TIP3P model for water.[60] Calculations were done with the CHARMM[61,62] and NAMD[63] programs.

In both the ON and OFF crystal structures, the switch 1 is partly disordered, so that several amino acids are missing from the crystal structure:[9,56] 8 in the ON state and 14 in the OFF state. The 8 amino acids missing in the ON structure are LYS36, HIS37, SER38, GLU39, GLU40, LEU41, LYS42, and ARG43 (for a net missing charge of +1). The 14 amino acids missing from the OFF structure are SER35, LYS36, HIS37, SER38, GLU39, GLU40, LEU41, LYS42, ARG43, GLY44, MET45, THR46, ILE47, and LYS48 (for a net missing charge of +2). The amino acids adjacent to this missing segment, SER35 and GLY44 (ON state) or THR34 and LEU49 (OFF state), are well within the 26 Å limit of our model but rather distant from the

ligand, especially its phosphate moieties; e.g., SER35 is 10 Å from the β phosphate of GTP in the ON structure, while MET45 is 9 Å away. These distances are somewhat greater than the physiological Debye−Hückel screening length. On the basis of the overall orientation of the visible portion of sw1, the missing amino acids are expected to be even further away. Therefore, rather than attempting to model these weakly ordered regions explicity, we have simply restrained the adjacent amino acids to stay close to their crystal positions, using harmonic restraints and a 4 kcal/mol/Å^2 force constant. As a result, their response to the GTP/GDP exchange is effectively modeled through the water dielectric response.

The ON and OFF X-ray structures contain GTP and GDP, respectively, with a single coordinating Mg^{2+} ion. The terminal phosphate group was assumed to be fully deprotonated in each case, consistent with biochemical data and crystal structures.[64−68] To compare GTP and GDP binding (Figure 1), we need to also consider a GTP-bound OFF state and a GDP-bound ON state. To model GTP (GDP) in the OFF (ON) structure, we have simply overlaid GTP (GDP) onto the GDP (GTP) of the crystal structure, taking the new ligand from the other crystal structure. Introducing GTP required the removal of 1−2 water molecules but did not lead to any steric clashes with protein groups.

2.2. MDFE Protocol for the Protein. To calculate the GTP/GDP binding free energy difference $\Delta\Delta G$, we follow the horizontal legs of the thermodynamic cycle in Figure 1. We use a hybrid energy function U that represents a mixture of the two end point states for the particular horizontal leg and depends on two coupling coordinates, λ_{elec} and λ_{vdw}, which are used to scale, respectively, electrostatic and van der Waals terms in the energy function:[51]

$$\begin{aligned} U &= U(\lambda_{elec}, \lambda_{vdw}) \\ &= \lambda_{elec} U_{elec}^{GTP} + (1-\lambda_{elec}) U_{elec}^{GDP} + \lambda_{vdw} U_{vdw}^{GTP} + (1-\lambda_{vdw}) U_{vdw}^{GDP} \end{aligned} \quad (1)$$

Here, U_{elec}^{GTP} (respectively, U_{elec}^{GDP}) represents the Coulomb interactions involving the GTP (GDP). Notice that these terms include small contributions due to interactions between GTP (GDP) molecules in different boxes within the periodic lattice. With a 74 Å box length, such interactions are small.[30,58] They are included, nevertheless, in our calculation of the free energy derivative (see below). By first varying λ_{elec} from 1 to 0 over a series of MD simulations with decreasing λ_{elec} values, we effectively remove the atomic charges of the GTP γ-phosphate; at the same time, we switch the charges of the β-phosphate to their values in GDP. This fictitious, intermediate ligand will be referred to below as GDP-P^0. In a second series of simulations, we vary λ_{vdw} from 1 to 0, removing the van der Waals interactions of the GTP γ-phosphate, leaving GDP in its place. From Boltzmann statistics, the free energy derivative with respect to either of the coupling constants can be written as $\partial G/\partial\lambda = \langle \partial U/\partial\lambda \rangle_\lambda$, where $\lambda = \lambda_{elec}$ or λ_{vdw} and the brackets represent an average over an MD trajectory performed with a particular value of λ. For the alchemical transformation of the electrostatic term, the successive λ_{elec} values were 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, and 0.0. Each simulation, or "window", lasted 2−4 ns; details for the individual runs are given below. Only the last half of each window was used to estimate $\langle \partial U/\partial\lambda \rangle_\lambda$. For the van der Waals transformation, the successive λ_{vdw} values were 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1, 0.05, 0.01, and 0.001. Each simulation lasted 1−4 ns; the last half was used for averaging. The small, final λ_{vdw} values reflect a very gradual removal of the γ-phosphate, to accurately account for the singularity of the free energy of particle removal.[69] The free energy derivative at each λ value (where $\lambda = \lambda_{elec}$ or λ_{vdw}) was computed from a finite-difference estimate. For the van der Waals windows, we used

$$\frac{\partial G}{\partial \lambda} \approx \frac{1}{2\delta} \langle U(\lambda+\delta) - U(\lambda-\delta) \rangle_\lambda \quad (2)$$

where $\delta = 0.002$. For the electrostatic windows, we used the difference $\langle U(\lambda = 1) - U(\lambda = 0) \rangle_\lambda$. We then calculated the free energy change using numerical integration. We integrated the derivatives using a standard trapezoidal method, except for the small-λ_{vdw} region of the van der Waals transformation. For the region between $\lambda_{vdw} = 0.05$ and 0, the free energy derivative was fitted to the function $A_0 \lambda_{vdw}^{-A_1}$, where A_0 and A_1 are adjustable parameters. Uncertainties of the free energy derivative at each λ value were estimated by dividing the trajectory segment used for averaging into two equal batches and taking the deviation between the batch averages. Two runs were performed for each state; total run lengths ranged from 29 to 71 ns; details (including window lengths) are listed in Results.

The GTP → GDP transformation was also performed in solution, using the same protocol, except that the cubic water box had an edge length of 60 Å and the windows had lengths of 600 ps (electrostatic and van der Waals steps). Experimentally, ATP is known to exist in solution mainly in the $[Mg{:}ATP]^{2-}$ form, with trace amounts of $[Mg{:}HATP]^{-}$ and $Mg_2{:}ATP$.[64,65] We may assume this situation also holds for GTP. Here, the transformation was done both in the presence and absence of bound Mg^{2+}. The nucleotide phosphates were fully deprotonated in all cases. Two runs in the GTP → GDP direction were done for the Mg^{2+} complex; one run in each direction was done for GTP/GDP alone in solution. A harmonic spring is present at the "N9" position of GTP/GDP; the spring contribution to the free energy cancels when the systems with and without Mg^{2+} are compared.

2.3. MDFE Protocol for Mg^{2+}:HPO_4^{2-} Binding. To compute the Mg^{2+}:HPO_4^{2-} binding free energy, we have reversibly deleted the solvated phosphate in the presence/absence of Mg^{2+}, using a multistep protocol illustrated in Figure 3. A harmonic spring (force constant = 2 kcal/mol/Å^2) was applied throughout the free energy simulations to the P and Mg atoms (legs 2), maintaining them close to a predefined position, with a 2.4 Å separation. This was the separation observed after several nanoseconds of free dynamics for the solvated pair. When the horizontal legs are subtracted to obtain ΔG_{bind}, the P spring contribution cancels exactly (legs 1a, 1a′, Figure 3). The free energy to introduce the Mg spring when Mg^{2+} is bound to the phosphate (leg 1b) is expected to be small, since the ion is held in position by the phosphate, even without the spring. In this situation, the effect of the spring is to remove the rotational freedom of the Mg^{2+}:HPO_4^{2-} complex. The rotational contribution to the free energy has a simple relation to the moments of inertia of the complex,[70] and can be estimated to be −5.0 kcal/mol. Finally, the free energy to introduce the spring when Mg is unbound (leg 1b′) can also be computed analytically; it has the form $\Delta G(\text{spring insertion}) = k_b T \ln V/(2\pi k_b T/\kappa)^{3/2}$, where k_b is the Boltzmann constant, T is the temperature, κ is the force constant for the spring, and V is the volume per Mg, which is just the inverse of the Mg concentration: $V = 1/[Mg]$.[70] The deletion of HPO_4^{2-} was done in two steps (legs 2), as above, with the charges deleted first and the van der Waals interactions second. We used a cubic box with an edge length of 50 Å. For the coupling

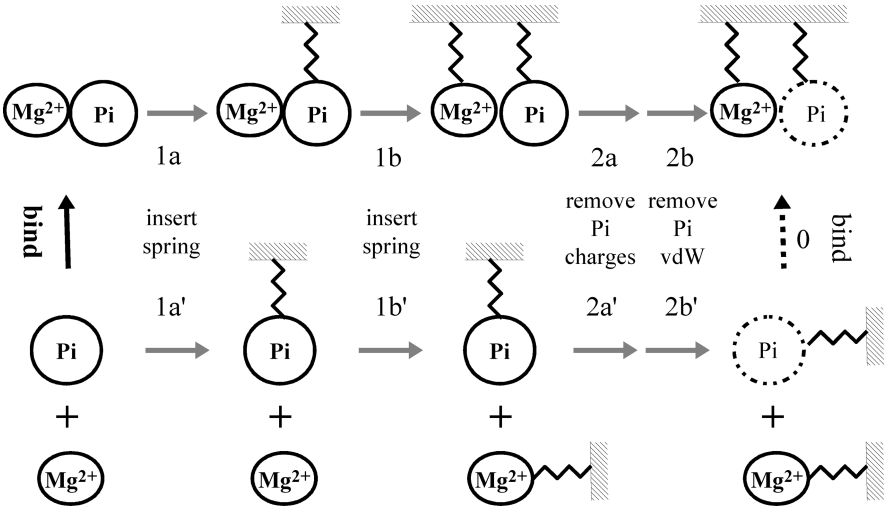


Figure 3. Thermodynamic cycle for Mg^{2+}:Pi binding, where Pi designates inorganic phosphate. Vertical legs correspond to binding. Horizontal legs correspond to the alchemical transformation of Pi into a ghost particle in solution, in the presence or absence of Mg^{2+}. Free energy changes for the different legs are computed by MDFE (2a, 2b, 2a′, 2b′), analytically (1a, 1b, 1a′, 1b′), or known to be zero (leg 0).

constants, we used the same values as before, with 600 ps of MD per window. For the phosphate alone, two runs were performed, respectively deleting and reintroducing the ion. For the Mg^{2+}:HPO_4^{2-} complex, a single run was performed, deleting the phosphate.

2.4. Quantum Chemical Calculations. Quantum chemical calculations were done for Mg^{2+} complexes with dihydrogen phosphate (DHP) and monohydrogen phosphate (MHP), surrounded by up to nine explicit water molecules, embedded in a dielectric continuum representing bulk aqueous solvent. Structures were prepared with a variety of interaction modes, including zero, one, or two direct interactions between Mg and phosphate oxygens. When six explicit water molecules were used, zero direct interactions amounts to bare phosphate interacting with $Mg^{2+}(H_2O)_6$, while two direct interactions means that two phosphate oxygens were bound to $Mg^{2+}(H_2O)_4$ and were capped with one water each. Analogous conformations were considered with nine explicit waters, in which the three additional molecules were placed so as to interact with the other phosphate oxygens. Several starting geometries were used in each case. They were energy-minimized at the density functional theory (DFT) level,[71] using the B3LYP-D functional, which includes an empirical correction for dispersion interactions.[72,73] This quantum solute was placed in a dielectric continuum, and the self-consistent reaction field was computed with the conductor-like screening model (COSMO) algorithm.[74] This model assumes nearly ideal conductor behavior for the solvent, which is particularly adequate for solvents with high dielectric constants such as water. The part of the electron density that extends into the solvent region is treated with the outlying charge correction.[75] The aug-cc-pVDZ atomic basis set was used for energy minimization as well as for normal mode calculations. Final energetics were computed using the aug-cc-pVTZ basis and the structures optimized with the aug-cc-pVDZ basis. For $(H_2PO_4^-)(Mg^{2+})(H_2O)_9$, this basis set involves a total of 1158 contracted Gaussian functions. Basis set superposition errors (BSSE), evaluated with the counterpoise correction method, were computed for both basis sets. The isolated solutes, $Mg^{2+}(H_2O)_6$, $H_2PO_4^-(H_2O)_3$, and $HPO_4^{2-}(H_2O)_3$, were treated in the same way. Calculations were done with the Turbomole package.[76] The free energy of a given solute included a vibrational entropy term obtained from a normal mode calculation. Rotational entropy was computed with a rigid rotor approximation. Translational entropy was computed assuming a 1 M standard state concentration.

2.5. Survey of Protein:Nucleotide Complexes in the Protein Data Bank. For comparison with the computed structures, we did a survey of protein structures in the Protein Data Bank (PDB) that have bound ATP, ADP, GTP, GDP, or close chemical analogues. 1019 structures that contained the words ATP, ADP, GTP, or GDP in the "structure description" field were identified using the PDB search engine and downloaded. It may be that additional structures exist that could be detected with a more sophisticated search procedure. Out of the 1019 structures, 285 included at least one magnesium ion. The atoms coordinating Mg directly in the structure (those within 3 Å) were identified using a perl script, and the coordination distances were recorded. Structures where a magnesium was coordinated by two oxygens from the same phosphate group were inspected visually using the Pymol program.[77]

2.6. Energy Calculations with the Fixed Charge and Polarizable Force Fields. To compare the different force fields, Charmm, Amber99sb,[78] and AMOEBA, we focus on a physical quantity that approximates the free energy to convert GTP into GDP. For a given conformational state, say ON, we consider the solvated protein with bound GTP, as above, and two different charge states for the GTP: the normal charges of GTP (with the particular force field) and a modified set, which mimics the charge distribution in GDP. Specifically, we add a charge of +0.25 to the four terminal atoms of GTP: three oxygens and the phosphorus. We refer to the ligands with the two charge states as L = GTP and L′ = GTP^0 (not to be confused with GDP-P^0). The superscript indicates that the terminal phosphate has a reduced charge. The net charge of L is −4; that of L′ is −3. For a given structure, taken from an MD simulation, the energy to change the ligand charge distribution from the L to the L′ values will be denoted ΔE. It is analogous to the energy gap for removing an electron instantaneously from a protein structure, which is a standard quantity in electron transfer theory.[42] We will consider an average of ΔE over an ensemble of protein structures, drawn

from an MD simulation. We consider two MD simulations, both performed with the Charmm force field. The first corresponds to the protein:GTP complex; the second corresponds to the end point of the electrostatic transformation in the MDFE simulations above, the protein:GDP-P^0 complex. In this end point, the γ-phosphate of GTP has its charges set to zero, while the β-phosphate charges are those of GDP. These two conformational ensembles, referred to as 1 and 0, are approximately representative of the L and L′ charge states, respectively. Finally, we consider the overall average, $(\langle \Delta E \rangle_1 + \langle \Delta E \rangle_0)/2$, which we denote $\delta g(1 \rightarrow 0)$. Indeed, if the ensembles 1 and 0 really do approximate the equilibrium states of the protein with bound L and L′, then δg has a simple interpretation: it represents a linear response approximation to the free energy to transform L into L′, and a rough approximation to the free energy ΔG to perform the first, electrostatic step in the GTP → GDP transformation. Our goal here is not to make a highly accurate estimate of ΔG. Rather, we view $\delta g(1 \rightarrow 0)$ as a good indicator of the force field dependency of the MDFE results. In fact, we will see that, when we estimate $\delta g(1 \rightarrow 0)$ with the Charmm force field, the result really does approximate ΔG rather closely.

To compute $\delta g(1 \rightarrow 0)$ with the polarizable, AMOEBA force field, we use the Tinker program.[79] The force field does not include optimized parameters for the triphosphate moiety of GTP, but approximate parameters can be inferred from the other nucleic acid parameters and from the Charmm force field. Highly accurate parameters are not needed, since our purpose is to check the sensitivity of $\delta g(1 \rightarrow 0)$ to the use of a polarizable force field. The most important point is to adopt reasonable values for the charges and polarizabilities, and we do this by combining the Charmm phosphate charges with the atomic polarizabilities of a phosphate group in AMOEBA DNA, described in the Amoebabio09 force field file in the Tinker distribution.[79] Atomic dipoles and quadrupoles were simply set to zero for the GTP phosphate groups. This simplified AMOEBA parametrization, combined with the full AMOEBA protein, guanosine, magnesium, and water parameters, should give a reasonable picture of the magnitude of polarization effects and deviations from the additivity hypothesis. The inserted charge is concentrated on the terminal phosphate group. For this group, it may not be fully consistent to combine the Charmm charges (which include polarization implicitly) with nonzero atomic polarizabilities. However, the inconsistency only affects a few atoms, so its effect should be small.

Calculations were done for both the ON and OFF states, using three different force fields: the polarizable Amoebabio09 force field,[53] modified to include GTP, and the fixed charge Amber99sb force field,[78] both implemented in the software package Tinker 4.2,[79] and the fixed charge Charmm27 force field,[52] implemented in the NAMD package.[63] The spherically truncated protein is solvated by a cubic water box, as above, for a total of about 39080 atoms. Setting up the truncated, solvated protein with Tinker was technically difficult, at least in our hands, and is described below. Energies were calculated with the two charge states, L and L′, with periodic boundary conditions and "plain" Ewald summation for the atomic multipoles and smooth particle mesh Ewald (PME) for charge−charge interactions. Energies were averaged over the last 1 ns of a 2 ns MD trajectory for either the 1 or 0 end point states, with structures sampled every picosecond. Calculations with the fixed charge Amber99sb force field were done with Tinker in the same manner.

To truncate the protein with Tinker, a rather complex procedure was necessary. First, the complete protein, the GTP,

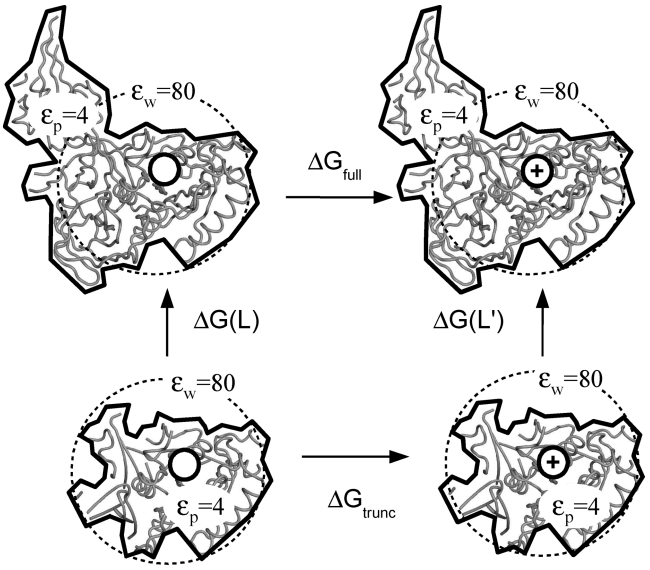


Figure 4. Scheme to compute the contribution of distant protein regions to the GTP/GDP binding free energy difference. Poisson−Boltzmann calculations are used to compare the spherically truncated model, limited to the MD region (below), and a larger model, spherically truncated with a 55 Å radius ("full" system, above). The 26 Å radius "MD" sphere is indicated as a dashed line for clarity. Each system is divided into an inner, protein medium, with a dielectric of four, and an outer, solvent medium, with a dielectric of 80. The dielectric boundary is schematized by a thick line. Horizontal legs represent the transformation of GTP by insertion of a positive charge onto its γ phosphate, mimicking the electrostatic step of the GTP → GDP transformation (see main text), and schematized here as the insertion of a "+" sign.

and the solvent were each set up separately, in the form of Tinker "xyz" files, and then merged into a single xyz file using Tinker. Manual corrections were needed for the water and the GTP xyz files. This led to an initial, reference file, "xyz_0". The coordinates in this reference file were those of the first MD snapshot, along with the crystal coordinates for the outer protein region. At the same time, the Tinker atom numbers of the atoms to be deleted (those outside the MD region) were recorded. Next, the other MD snapshots were modified by adding back the protein region not present in the MD model (using the crystal coordinates) and then converted to Tinker xyz files. These files contained atom connectivity and name errors, especially for the GTP, and could not be used directly. Therefore, a perl procedure was required, which cross-checked the coordinates in each (malformed) xyz file against those in the corresponding MD snapshot to identify corresponding atoms, and then edited the reference file "xyz_0", updating the atomic coordinates to be those of the current MD snapshot. Finally, the outer atoms were deleted from each xyz snaphot using the xyzedit module of Tinker, slightly modified for our purposes.

2.7. Long Range Correction: Poisson−Boltzmann Setup. To reduce the computational cost of the long MDFE runs, we have considered a spherically truncated protein model (radius = 26 Å; see above), solvated by a 74 Å water box. The portion of protein outside the 26 Å radius is neglected during the MD simulations and replaced by explicit waters (see Figure 2). In a second step, the distant protein is reintroduced and the corresponding free energy change is computed, using a dielectric continuum model (Figure 4).[46] The protein is reintroduced for

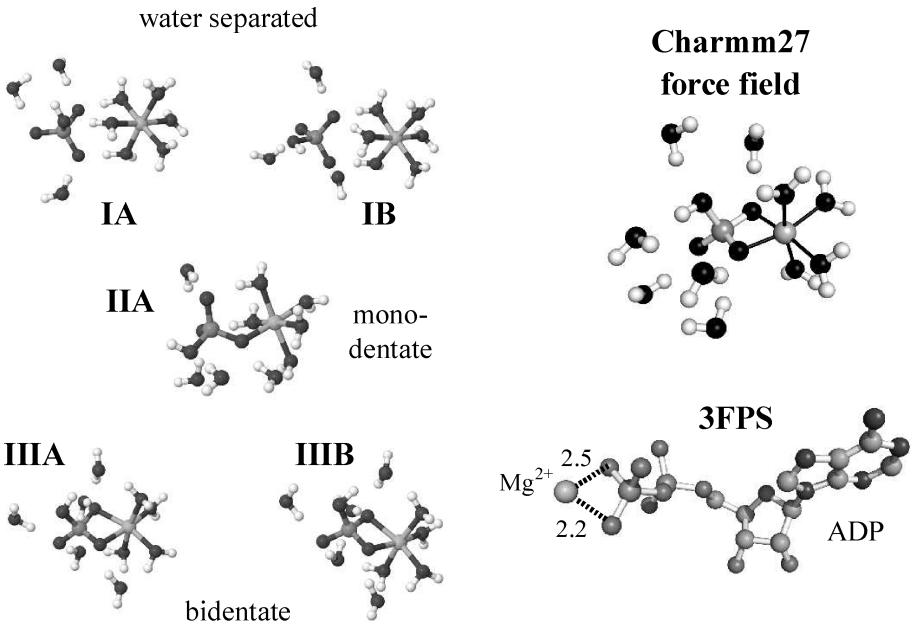


Figure 5. Mg^{2+}:phosphate complexes from the quantum chemical model (left), the Charmm27 force field model (upper right), and the Protein Data Bank (lower right). The five quantum structures include Mg^{2+} (on the right, surrounded by 4−6 waters), HPO_4^{2-} (on the left), and nine water molecules; the structures were energy-minimized at the B3LYP-D/aug-cc-pVDZ level in the presence of a dielectric continuum solvent (see text). Structures IA and IB have a water-mediated Mg^{2+}:phosphate interaction; IIA, IIIA, and IIIB have a direct interaction, either mono- or bidentate. The Charmm27 force field structure includes 10 waters that coordinate Mg^{2+} or hydrogen bond to HPO_4^{2-}. The Mg^{2+}:ADP complex (lower right) is from the 3FPS Protein Data Bank entry. It has a bidentate Mg^{2+}:phosphate interaction, with distances indicated (in Å).

both end points of the first, electrostatic step of the GTP → GDP transformation ($\lambda_{elec} = 1 \rightarrow \lambda_{elec} = 0$). Presumably, the main effect of the distant protein is to act as a low dielectric and modify (slightly) the screening of electrostatic interactions of the protein and solvent with the GTP and GDP ligands. This long-range effect should be accurately modeled by continuum electrostatics. For the second, van der Waals step of the GTP → GDP transformation, we assume the contribution of the distant protein is negligible; this assumption is reasonable, since the effect on the electrostatic step is found to be small, and the effect on the van der Waals step will almost certainly be even smaller.

To compute the free energies of the system with and without the distant protein, we considered the crystal structure, either intact or spherically truncated, with all the waters removed. Protein and ligand (GTP or GDP-P^0) were treated as a single dielectric medium with a dielectric constant of 4; solvent was treated as another medium with a dielectric constant of 80. The boundary between the two media was defined as the protein/ligand molecular surface, computed with a probe sphere having a 3 Å radius. The system was discretized using a cubic grid with 241 regularly spaced planes and a spacing of 1 Å. The ionic strength in the aqueous medium was set to zero. The Poisson equation was solved numerically, with Coulombic boundary conditions, using the Charmm program (PBEQ module).[80] A second, "focussing" calculation was then performed using a smaller grid with 241 planes but a 0.5 Å spacing. The potential from the first step was used to define the boundary potential on the surface of the second, smaller grid.[81] To better understand the long-range effects, we did two separate calculations. In the first, for the complete system, protein charges outside the 26 Å sphere were set to zero; i.e., the outer region acts only by desolvating the inner, spherical region. In the second, the outer protein charges were present, and their direct contribution to the GTP → GDP-P^0 transformation was evaluated. The net charge of the protein for the outer amino acids is +4 for the ON structure; the charge of the region between 26 and 50 Å is just +2.

3. RESULTS

3.1. Magnesium:Phosphate Binding: Model Systems.

3.1.1. Quantum Chemical Treatment. We used a quantum chemical treatment to study Mg^{2+} binding to either dihydrogen phosphate (DHP) or monohydrogen phosphate (MHP). Our strategy is based on a combination of explicit, microsolvation and implicit, continuum modeling of bulk water. It was first deemed essential to "cover" the magnesium dication highly localized charge with a complete coordination shell, involving six ligands. Several arrangements were considered, illustrated in Figure 5. Structures of type I start with $Mg^{2+}(H_2O)_6$, to which a bare phosphate (either DHP or MHP) is added, establishing several hydrogen bonds between phosphate oxygens and water OHs. For structures of types II and III, one or two direct contacts between Mg^{2+} and a phosphate oxygen are also established, and one or two water molecules are displaced to interact with the phosphate oxygens. Three additional waters were then added to each of these structures, positioned to make hydrogen bonds with several of the phosphate oxygens and to provide explicit electrostatic screening, in between these charged oxygens and the polarizable continuum outer region. Several starting geometries were tried for each structural type; the most stable of each type are shown in Figure 5 for the MHP complex. The binding enthalpies and free energies are listed in Table 1 for both MHP and DHP. To maintain consistency between the hydrated Mg^{2+}:phosphate complexes and the separated phosphate and Mg^{2+}, all structures having explicit water molecules in the second coordination sphere of the phosphate were eliminated. Although we did

dx.doi.org/10.1021/jp201934p |J. Phys. Chem. B 2011, 115, 6749–6763

Table 1. Quantum Chemical Mg^{2+}:Phosphate Binding Free Energies (kcal/mol)[a]

structure	ΔH	ΔG	BSSE-corrected ΔG
	HPO_4^{2-}:Mg^{2+}:$(H_2O)_9$		
IA	−26.27 (−27.13)	−11.58 (−12.46)	−10.91 (−10.22)
IB	−26.56 (−27.68)	−11.05 (−12.15)	−10.33 (−9.55)
IIA	−26.96 (−28.04)	−9.98 (−11.05)	−8.90 (−7.56)
IIIA	−29.45 (−31.46)	−11.89 (−13.90)	−10.62 (−10.17)
IIIB	−30.48 (−32.56)	−11.10 (−13.18)	−9.86 (−9.11)
	$H_2PO_4^-$:Mg^{2+}:$(H_2O)_9$		
IC	−22.46 (−23.21)	−4.02 (−4.78)	−3.44 (−2.75)
ID	−21.99 (−23.37)	−3.01 (−4.43)	−2.37 (−2.01)
IIB	−19.69 (−20.91)	−0.84 (−2.08)	0.17 (−1.24)
IIC	−24.64 (−25.74)	−3.97 (−5.07)	−2.87 (−1.05)
IIIC	−25.00 (−26.89)	−2.27 (−4.04)	−0.86 (0.84)

[a] Enthalpies, free energies, and BSSE-corrected free energies of binding of Mg^{2+} to HPO_4^{2-} (MHP) and $H_2PO_4^-$ (DHP) in water at 298 K. All values in kcal/mol. Computations were done at the B3LYP-D/aug-cc-pVTZ level; B3LYP-D/aug-cc-pVDZ results are also given in parentheses.

not try to explore the potential energy surfaces exhaustively, it seems unlikely that other microsolvation structures exist that would be significantly more stable than the ones described here.

Before discussing the relative energies of the various structures, it is worth describing the convergence of the computational strategy. Since negatively charged species are involved, all atomic basis sets used are rich in diffuse functions. The aug-cc-pVDZ basis used for geometry optimizations and vibrational frequency calculations is, however, not expected to be sufficient for accurate energetics. In particular, the very strong electrostatic interactions lead to short interatomic distances and thus to potentially large basis set superposition errors (BSSE) on binding energies. The results in Table 1 indicate that BSSE values with the aug-cc-pVDZ basis are typically 2.5−5 kcal/mol, which is annoyingly large. Energy calculations for the same geometries, using the aug-cc-pVTZ basis, lead to reduced BSSEs, now lying in the 0.5−1.2 kcal/mol range. The BSSE calculation shows that the binding energies are slightly underestimated with the smaller basis set, underlying the need for extended basis set calculations. Overall, these results indicate that basis set convergence is close to being attained at the aug-cc-pVTZ level with the B3LYP-D functional.

The Mg^{2+}:phosphate binding enthalpies and free energies reported in Table 1 are computed relative to hydrated Mg^{2+}-$(H_2O)_6$ and $HPO_4^{2-}(H_2O)_3$ or $H_2PO_4^-(H_2O)_3$. While this is the intuitively obvious dissociation of structures of type I, it is less so for structures II and III. The general principle should be to microhydrate both the complex and the dissociated fragments in the same way. However, there is not a single solution with a given number of explicit waters. To settle this issue, calculations were also carried out using hydrated $Mg^{2+}(H_2O)_4$ and HPO_4^{2-}-$(H_2O)_5$ as the dissociation products. The results indicated this combination to be higher in energy than $Mg^{2+}(H_2O)_6$ and $HPO_4^{2-}(H_2O)_3$, so this latter dissociation product was used throughout.

The results in Table 1 indicate that the various types of interaction between Mg^{2+} and MHP are associated with very similar energetics. Structures of type III are favored by enthalpic effects, as expected from the very strong electrostatic interaction between a dianion and a dication with two direct contacts. However, this leads to more rigid structures and an associated entropic penalty. Overall, the binding free energies for the five structures lie within a 2 kcal/mol window. Especially striking is the close proximity of the binding free energies obtained for structures of types I and III, which are within 0.3 kcal/mol. Binding of Mg^{2+} to DHP involves weaker electrostatic interactions and binding enthalpies about 5 kcal/mol weaker than for MHP. This difference is increased further by entropic effects, leading to binding free energies that are smaller by about 7 kcal/mol. As with MHP, this effect is strongest for structures of type III, so that, in the DHP case, structure IIIC is less favorable than the type I and II structures. This result is in line with the infrared signatures of diethyl phosphate, which suggest that water positions itself between it and Mg^{2+}.[82]

The most stable structure of $H_2PO_4^-$:$Mg^{2+}(H_2O)_9$ is associated with a binding free energy of −3.3 kcal/mol, a value in good agreement with that derived from experimental association constants, −1.7 kcal/mol.[64] While Boltzmann averaging over all the structures reported would lead to a smaller computed value, it is also conceivable that another, slightly more stable structure exists, so that −3.3 kcal/mol is close to our best estimate. This comparison to experiment provides support for the reliability of modeling approaches for the interaction of singly charged phosphates with divalent cations in water.[83,84]

The situation is less favorable for HPO_4^{2-}:$Mg^{2+}(H_2O)_9$, for which the calculations give an estimated binding free energy of −10.8 kcal/mol, significantly greater than the experimental value of −3.7 kcal/mol. The error is thus about 7.1 kcal/mol. There is essentially no quantum chemical literature for such difficult cases of multiply charged cations and anions interacting in a condensed phase, so our results may be considered as a benchmark, from which further progress should be made. Notice that nonideal corrections to the binding free energy[70] are considered in the next section and shown to be negligible.

3.1.2. Mg^{2+}−HPO_4^{2-} Binding in Solution: Fixed Charge Force Field. The experimentally reported[66] dissociation constant for the Mg^{2+}:HPO_4^{2-} complex is K_d = 1.95 mM at room temperature (298.15 K) and pressure (1 atm), so that the standard binding free energy is ΔG^0_{bind} = −3.68 kcal/mol. To compute this binding free energy with the Charmm27 fixed charge force field, we use molecular dynamics and a thermodynamic cycle shown in Figure 3. For the lower leg, two runs were averaged, removing and reinserting HPO_4^{2-} into solution, respectively. The free energy is 399.7 kcal/mol, including 397.1 kcal/mol from step 2a′ (electrostatic term) and 2.6 kcal/mol from 2b′ (van der Waals term). For the upper leg, phosphate deletion from the complex, steps 2a and 2b yield free energy changes of 443.0 and −3.5 kcal/mol, for a total of 439.5 kcal/mol, estimated from a single MDFE run. A "backward" run was also attempted, reinserting HPO_4^{2-} into the complex; however, the phosphate remained trapped in a high free energy geometry, separated from the Mg^{2+} ion by a water molecule. Rather than putting additional effort into the convergence of the backward run (by reversibly deleting the bridging water in additional steps, for example), we simply adopted the estimate from the forward run. The free energy for Mg spring insertion in step 1b′ is 3.8 kcal/mol when the Mg^{2+} concentration is 1 M. The free energy for Mg spring insertion in the upper leg (1b) is estimated to be about 5.0 kcal/mol.

Overall, subtracting the upper and lower legs of the cycle, we obtain a standard binding free energy of −41.0 kcal/mol, an

order of magnitude larger than the experimental value. This indicates a strong systematic error, presumably due at least partly to the lack of explicit electronic polarizability in the fixed charge force field. Notice that the quantum chemical treatment gave a much smaller but still significant error, indicating that even that model has difficulty describing the Mg^{2+}:MHP interactions. The most favorable MD conformation for Mg^{2+}:HPO_4^{2-} involves a direct, bidentate coordination of Mg^{2+} by two of the phosphate oxygens, shown in Figure 5. In the quantum chemical structures, above, a water-separated interaction (structure IA, Figure 5) was slightly preferred (by 0.3 kcal/mol) over the bidentate conformation (structure IIIA). Water-separated interactions are seen in some Mg:phosphate crystal structures (e.g., Cambridge Crystallographic Database entries KIMNID and GATLUI).[85] In a diethylphosphate:Mg^{2+} crystal structure, a monodentate coordination is seen, similar to IIA,B.[86]

For the Mg^{2+} + HPO_4^{2-} ionic solution, we should also consider the possibility of nonideal corrections to the free energy, due to long-range interactions between the divalent Mg^{2+} and HPO_4^{2-} ions. In the upper leg, after phosphate deletion, direct interactions between Mg^{2+} ions in neighboring boxes contribute a term $(z^2e^2\pi/2L^3\kappa^2)$ to the free energy,[58,87] where z is the valency ($z = 2$), e is the charge of a proton, L is the box length, and $\kappa = 0.34$ Å is the charge width used in the Ewald setup, giving a small contribution of 0.14 kcal/mol (opposing phosphate deletion). In the upper leg, the same effect favors phosphate deletion, so the two legs cancel. The second long-range effect corresponds to stabilizing correlations between ions in the same or different boxes. The earliest and simplest theory to estimate these effects is the Debye−Hückel theory.[70] In solution, an ion of a given type will be preferentially surrounded by ions of the other type, leading to correlations between ions and stabilizing interactions, in contrast to the ideal dilute situation, where ions of a given type do not interact with each other. In the simulation model, we have a periodic array of cubic boxes, with the constraint that different boxes undergo the exact same motions. This situation is obviously very different from the ideal dilute state, where the ions all move independently. Nevertheless, it is usually taken as an acceptable approximation to the ideal dilute case, as long as the box size is fairly large. An important point is that in the periodic situation, a given ion (say, Mg^{2+}) has no way to preferentially surround itself with ions of the other type (HPO_4^{2-}), except for its partner within the same box. Thus, while correlations between ions do exist due to the periodicity, they do not have the systematic, stabilizing character of a real, nonperiodic solution at finite dilution. We will therefore make the usual assumption that the MD setup represents an ideal dilute situation, and add to the free energy the Debye−Hückel correction, reflecting the stabilizing effect of interion correlations at finite dilution. This correction has the form[70]

$$G_{\mathrm{Debye}} = -\frac{2k_bT}{3}\sqrt{\rho\pi}\left(\frac{z^2e^2}{\varepsilon_W k_b T}\right)^{3/2} \qquad (3)$$

where ρ is the ionic concentration, z and e were defined above, ε_W is the dielectric constant of water, and T is the temperature. Our box size (50 Å) corresponds to an ionic concentration of 13.3 mM, so the free energy contribution is $G_{\mathrm{Debye}} = -0.24$ kcal/mol (opposing binding), which is negligible compared to the MDFE binding free energy term.

3.1.3. Mg^{2+} Binding to GTP and GDP in Solution. We next consider Mg^{2+} binding to GTP and GDP in solution, using MDFE and the Charmm27 force field. We compute the Mg^{2+} binding free energy difference between GTP and GDP by alchemically transforming one into the other, either in complex with Mg^{2+} or alone in solution. In the MD structures, GTP and GDP have an extended conformation and Mg^{2+} coordinates all the phosphate groups. A free energy study of ATP with the present force field showed that extended structures with coordination of either three phosphates or just the β and γ phosphates were equistable and correspond to the global free energy minimum.[88] We recall that the nucleotide protonation state and the Mg^{2+} binding stoichiometry are supported by biochemical data.[64−66] In the presence of Mg^{2+}, we obtain $\Delta G(\mathrm{GTP} \rightarrow \mathrm{GDP}) = 432.3$ kcal/mol; the electrostatic step contributes 443.1 kcal/mol; the van der Waals step contributes −10.9 kcal/mol. In the absence of Mg^{2+}, we obtain $\Delta G(\mathrm{GTP} \rightarrow \mathrm{GDP}) = 419.6$ kcal/mol; the electrostatic step contributes 431.7 kcal/mol; the van der Waals step contributes −12.1 kcal/mol. Thus, the computed binding free energy difference is −12.8 kcal/mol, favoring GTP. Experimentally, the dissociation constants for ATP and ADP binding are available,[66] giving a standard binding free energy difference of −2.1 kcal/mol, which is expected to be very close to the GTP/GDP difference. Thus, in a different context, the force field gives again a large error (almost 11 kcal/mol) for Mg^{2+}:phosphate binding.

In two ATP:Mg^{2+} crystal structures, Mg^{2+} coordinates all three phosphates, as in our simulations; see Cambridge Structural Database entries DECDIY and CICPOT.[67,85,89] An X-ray structure also exists for ATP:Na_2:$(H_20)_3$,[68] and quantum calculations have been used to substitute a Mg^{2+} for one of the sodium ions; in the quantum structure, Mg^{2+} coordinates all three phosphates of an ATP (monodentate coordination with no bridging waters). To test the MD structures further, we also did a survey of protein structures in the Protein Data Bank containing ATP, ADP, GTP, GDP, or close chemical analogues, along with at least one Mg^{2+} ion. We only considered structures with direct Mg^{2+}:phosphate interactions (type II and III structures), not water-separated interactions (type I), since the latter are harder to identify in the PDB structures; in particular, a Mg^{2+} ion that does not directly coordinate the nucleotide may be marked as a water in the PDB file. Within our data set, nearly all the nucleotides have an extended conformation (see ref 90 for a more detailed survey). Most of the ions coordinate two distinct phosphate groups. For these biphosphate cases, the coordination always involves a single oxygen from each phosphate, never two oxygens from a single phosphate. About one-third of the structures (94 structures) have a Mg^{2+} ion coordinated by a single phosphate group, usually the terminal phosphate of the nucleotide. Structures where the Mg^{2+} is coordinated by a negatively charged Asp or Glu side chain, in addition to the phosphate, were not included in this set. Of the 94 structures, 92 have a monodentate coordination, while two have a bidentate coordination (they bind two oxygens from the same phosphate): 2J1L and 3FPS. 2J1L is the human Rho-related protein HP1, solved at a 2.5 Å resolution (unpublished structure; see PDB), while 3FPS is a Ca^{2+}-transporting ATPase from rabbit, solved at a medium, 3.2 Å resolution.[91] The coordinating phosphate in these two cases is the terminal phosphate of GDP (2J1L) or ADP (3FPS). The two shortest Mg^{2+}:oxygen distances in these structures are 2.2 and 2.6 Å (2J1L) and 2.2 and 2.5 Å (3FPS). The 3FPS structure is included in Figure 5. Thus, we observe direct, bidentate Mg^{2+}:phosphate interactions in about 2% of the relevant PDB structures and monodentate interactions in the

Table 2. Force Field Sensitivity of the $\Delta g(1 \rightarrow 0)$ Free Energy Estimator[a]

	Amoeba	Amber99sb	Charmm	Charmm (rigorous MDFE)
solution	260.4 (0.9)	227.8 (1.4)	228.8 (1.3)	
ON-aIF2	260.4 (0.2)	231.0 (1.0)	234.6 (3.1)	
OFF-aIF2	256.5 (2.0)	223.7 (2.3)	227.9 (0.2)	
ON/solution difference	0.1 (0.9)	3.2 (1.7)	5.8 (3.4)	2.2
OFF/solution difference	−3.8 (2.2)	−4.0 (2.7)	−0.9 (1.3)	−2.2
ON/OFF difference	3.9 (2.0)	7.3 (2.5)	2.7 (3.1)	

[a] Free energies in kcal/mol. Values in parentheses are the uncertainty, estimated for the upper values (solution, ON, OFF) as the difference between the results from the first and second halves of each ensemble of structures. For the lower values (differences between states), the uncertainty is estimated by propagation of the individual uncertainties of each state.

Table 3. Contribution of Protein outside the MD Sphere to $\Delta G(\text{GTP} \rightarrow \text{GDP} - \text{P}^0)$[a]

	outer charges set to zero				
	ΔG_{full}	ΔG_{trunc}	difference	contribution of outer charges	total contribution of outer protein
ON	−436.5	−436.0	−0.5	1.2	0.7
OFF	−404.1	−403.8	−0.3	1.1	0.8
difference			−0.2[b]	+0.1[c]	−0.1[d]

[a] Free energies in kcal/mol. Full and truncated models are illustrated in Figure 2. [b] This term reflects the contribution to $\Delta G(\text{GTP} \rightarrow \text{GDP} - \text{P}^0)$ arising from desolvation of the inner region by the outer protein. [c] This is the direct contribution of charges in the outer protein region. [d] Total contribution of the outer protein to the ON/OFF difference in $\Delta G(\text{GTP} \rightarrow \text{GDP} - \text{P}^0)$.

Table 4. Free Energy Results for GTP → GDP Transformation[a]

			free energy for GTP → GDP transformation		
system	run direction	simulation length	ΔG^{elec}	ΔG^{vdw}	$\Delta G^{\text{elec}} + \Delta G^{\text{vdw}}$
ON-aIF2	forward	24 + 7.6 ns	446.1 (1.8)	−13.4 (1.0)	432.7 (2.1)
ON-aIF2	forward	15.2 + 13.8 ns	444.7 (1.4)	−12.2 (1.0)	432.5 (1.7)
OFF-aIF2	forward	24 + 13.8 ns	440.2 (2.9)	−11.1 (0.1)	429.1 (2.9)
OFF-aIF2	forward	15.2 + 7.8 ns	441.7 (0.6)	−14.2 (0.2)	427.5 (0.6)
solution	forward	6.6 + 11.4 ns	443.1 (0.6)	−10.7 (0.9)	432.4 (1.1)
solution	forward	6.6 + 11.4 ns	443.3 (0.6)	−11.1 (0.3)	432.2 (0.7)

[a] Free energies in kcal/mol. Apparent statistical uncertainty in parentheses. Simulation lengths correspond to the sum of all windows for the electrostatic + van der Waals steps. Forward runs transform GTP into GDP. Solution runs include an Mg^{2+} ion coordinating the nucleotide.

other 98%. This suggests that, for the phosphate case above, both the force field and the quantum chemical model may overestimate the occupancy of the bidentate conformation. Conversely, the observation of two bidentate structures (and not zero) also suggests that the true free energy difference between mono/bidentate conformations is not much greater than a few kcal/mol.

3.2. Testing the Additivity Hypothesis: Comparing Polarizable and Fixed Charge Force Fields. Our goal is to understand the specific aIF2:nucleotide binding using MDFE simulations. Both GTP and GDP bind aIF2 with a cobound Mg^{2+} ion, which coordinates the β and the γ phosphate (when present). Thus, when GTP is replaced by GDP as the ligand, a Mg^{2+}:phosphate interaction is removed. We saw above that the Charmm27 fixed charge force field gives large errors for the corresponding binding free energy. However, we expect that a large part of the error will cancel when we compare the GTP → GDP transformation in solution and/or in two different protein states (ON, OFF). To test this "additivity" hypothesis, we focus on a simplified transformation where a positive charge is added to the GTP γ phosphate (yielding GTP^0; see Methods) and a simplified energy difference: $\Delta g(1 \rightarrow 0)$, which approximates the corresponding free energy change (see Methods). We estimate $\Delta g(1 \rightarrow 0)$ using three different force fields, and two ensembles of structures, drawn from MD simulations of aIF2 bound to either the plain GTP or the modified, GDP-P^0 ligand. Obviously, these ensembles of structures, produced with the Charmm27 force field, do not represent true equilibrium ensembles for the polarizable force field model or the Amber99sb model, nor for the GTP^0 complexes. Nevertheless, we expect that the variations of $\Delta g(1 \rightarrow 0)$ with different force fields should be a reasonable indicator of the MDFE force field sensitivity.

Calculations were done for GTP/GTP^0 in solution and for both the ON and OFF states of the protein. Both in the protein and in solution, a bound Mg^{2+} ion is present. Results are summarized in Table 2. We focus on the differences between the protein and solution, or between the two protein states (bottom part of Table 2). For these quantities, the differences between the force fields remain significant. The largest

dx.doi.org/10.1021/jp201934p | J. Phys. Chem. B 2011, 115, 6749–6763

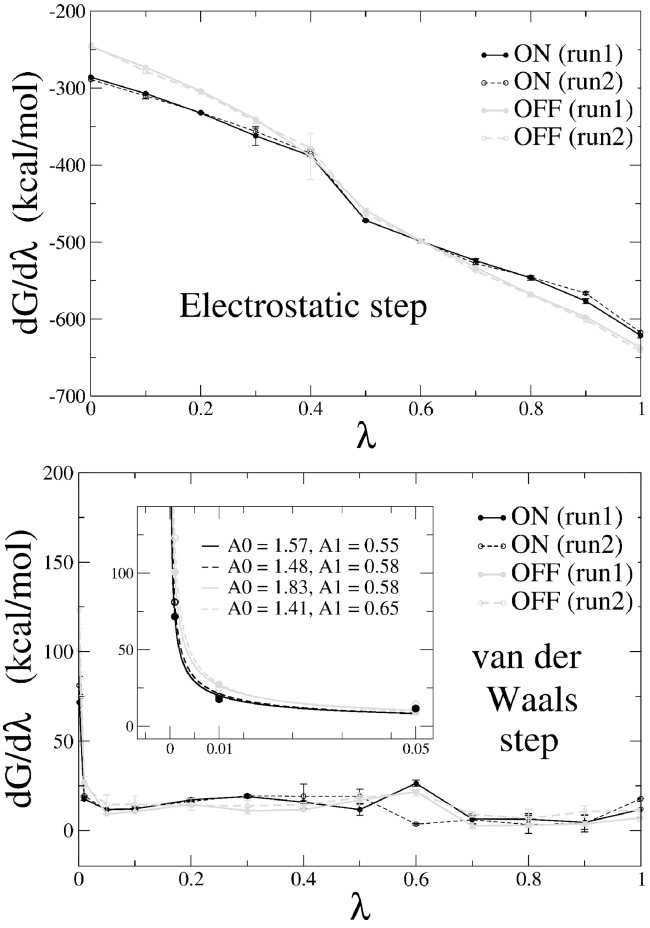


Figure 6. Above: Free energy derivative $\partial G/\partial\lambda$ for the electrostatic step of the GTP → GDP transformation in the ON and OFF states, as a function of the coupling coordinate, λ. Data for two runs are shown. Uncertainties are shown as vertical bars (often too small to be seen). Below: $\partial G/\partial\lambda$ for the van der Waals step of the transformation. The inset enlarges the region near $\lambda = 0$; the smooth curves are fits to the numerical derivatives with the function $A_1\lambda^{A_2}$ (see text).

discrepancy is the difference between AMOEBA and Charmm for the ON/solution difference: $\Delta\Delta g(1 \rightarrow 0)$ = 0.1 kcal/mol with AMOEBA and 5.8 kcal/mol with Charmm. This is much smaller than the differences between Charmm27 and experiment for Mg^{2+}:phosphate binding in solution but still quite significant. The uncertainty for the Charmm result is rather large, ±3.4 kcal/mol. The MDFE estimate of the GTP → GDP-P^0 free energy change, which involves considerably more MD sampling, is 2.2 kcal/mol. This value falls within the Charmm uncertainty range for $\Delta g(1 \rightarrow 0)$ but is much closer to the AMOEBA result. The OFF/solution differences estimated with the different force fields are also similar, with AMOEBA and Charmm27 differing by just 2.7 kcal/mol and AMOEBA and Amber99sb almost identical. Finally, when the two protein states are compared, AMOEBA and Charmm27 are in good agreement, with a 1.2 kcal/mol difference. This last difference may be fortuitously small, given the uncertainties. Overall, though, it is clear that the additivity hypothesis is at least roughly valid when two protein conformations are compared; at the same time, caution will have to be used when interpreting the MDFE results, below.

3.3. Long-Range Contributions from Distant Protein. The MD simulations use a spherically truncated protein model,

Table 5. Free Energy Derivatives for the GTP → GDP Transformations in the Protein[a]

ON state	run 1		run 2	
λ	elec	vdw	elec	vdw
1.0	−621.3 (3.3)	11.8 (4.5)	−617.50 (2.1)	17.82 (0.2)
0.9	−576.6 (3.3)	4.5 (3.9)	−566.47 (1.6)	3.94 (4.8)
0.8	−546.5 (3.0)	6.3 (3.1)	−545.42 (1.2)	3.32[d] (5.0)
0.7	−524.1 (2.5)	6.4 (2.8)	−527.90 (1.2)	6.00 (2.3)
0.6	−499.0 (0.3)	26.4[c] (1.7)	−498.88 (2.3)	3.52[d] (0.4)
0.5	−471.9 (0.2)	11.6 (3.1)	−472.12 (0.9)	19.1[d] (4.0)
0.4	−388.1[b] (0.9)	15.8 (0.8)	−384.73[c] (4.5)	19.12 (6.8)
0.3	−362.1 (12.3)	19.0 (0.8)	−356.26 (4.3)	19.46 (0.5)
0.2	−332.4 (0.1)	17.2 (0.8)	−331.68 (0.6)	16.29 (0.5)
0.1	−307.1 (1.1)	12.1 (0.5)	−310.26 (3.7)	12.48 (1.1)
0.05		11.7 (0.2)		11.37 (0.0)
0.01		17.6 (0.7)		19.35 (0.1)
0.001		71.7 (2.0)		81.12 (5.1)
0.0	−285.8 (0.5)		−288.92 (1.9)	

OFF state	run 1		run 2	
λ	elec	vdw	elec	vdw
1.0	−637.3 (4.2)	7.1 (2.0)	−641.4 (1.3)	11.4 (0.9)
0.9	−597.3 (0.2)	3.7 (1.3)	−600.8 (2.9)	10.5 (3.3)
0.8	−567.6 (0.9)	3.0 (1.0)	−568.7 (2.1)	6.7 (5.1)
0.7	−533.0 (0.8)	2.6[d] (1.6)	−537.3 (0.6)	9.1 (1.6)
0.6	−498.4 (2.3)	21.6[d] (0.9)	−499.2 (0.1)	21.2 (1.8)
0.5	−458.7 (2.7)	17.0[d] (2.8)	−463.0 (4.3)	18.4 (2.6)
0.4	−388.7[b] (30.1)	11.5 (0.5)	−378.2[c] (0.3)	14.3 (1.3)
0.3	−340.3 (0.1)	11.0 (1.8)	−343.4 (2.6)	13.8 (2.9)
0.2	−303.7 (0.2)	14.6 (3.2)	−305.4 (2.6)	14.1 (0.6)
0.1	−273.0 (1.8)	10.5 (0.5)	−278.3 (2.5)	14.5 (4.8)
0.05		9.3 (0.2)		14.4 (5.5)
0.01		27.2 (3.0)		24.9 (1.3)
0.001		100.8 (12.1)		123.1 (24.8)
0.0	−245.9 (1.8)		−244.5 (3.1)	

[a] In kcal/mol. Uncertainty in parentheses (difference between the results from the first and second halves of each window). Windows for the electrostatic steps lasted 2 ns (run 1) or 1.2 ns (run 2) and those for the van der Waals step lasted 600 ps, except for selected windows[b,c,d] run for a longer period. [b] 4 ns. [c] 3.2 ns. [d] 2.6 ns.

surrounded by aqueous solvent, which is very efficient and allows long simulations.[46] To estimate the effect of this approximation, we consider a process where the protein outside the sphere is reintroduced reversibly, and the effect on the free energy is computed with a dielectric continuum model (Figure 4). The protein and ligands are treated as a single dielectric medium with a dielectric constant of 4. This value has been found to give reasonable results for a wide range of ligand binding studies.[30,92] The free energy is computed by solving the Poisson equation. Atomic charges in the external protein region are initially set to zero (see Methods). Thus, introducing the external protein region only acts to change the solvent shielding of charges of the inner, 26 Å sphere. Results are given in Table 3. The free energy changes when GTP is changed to GDP-P^0 are very large, over 400 kcal/mol. A large part of this comes from the

dx.doi.org/10.1021/jp201934p |J. Phys. Chem. B 2011, 115, 6749–6763

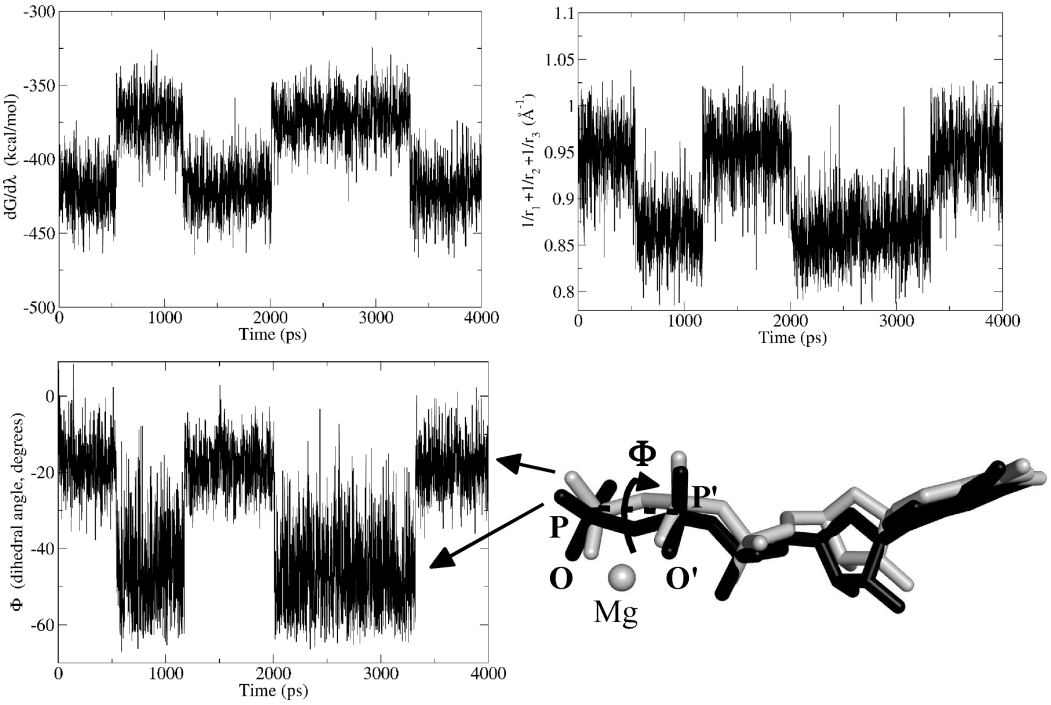


Figure 7. Sampling during the $\lambda_{elec} = 0.4$ window of run 1, OFF state. Upper left: free energy derivative. Upper right: inverse O:Mg distance, summed over the three terminal GTP oxygens. Bottom left: pseudodihedral angle Φ, defined by the atoms O−P−P′−O′ and highlighted by a dashed line in the bottom right panel. Bottom right: GTP structures representative of the two isomers seen in the dihedral plot.

interactions of the new charge with the rest of the GTP moiety. This contribution can be viewed as an artifact of the point charge model used here; however, it cancels exactly when the complete and truncated systems are compared, and also largely when the ON and OFF states are compared. Thus, the differences between the full and truncated systems are quite moderate: −0.5 kcal/mol for the ON state and −0.3 kcal/mol for the OFF state. The ON/OFF difference is just −0.2 kcal/mol. Next, we introduce the protein charges in the outer region, either in the presence of GTP or GDP-P^0. The contribution of the outer charges to the GTP → GDP-P^0 transformation is +1.2 kcal/mol in the ON state and +1.1 kcal/mol in the OFF state, for a difference of +0.1 kcal/mol. Taking the two calculations together, the outer protein region contributes just −0.1 kcal/mol, favoring GDP binding to the ON state. This small value is not very sensitive to ionic strength. Thus, computing the direct effect of the outer protein charges with ionic concentrations of 0.1 or 0.2 M, we obtain smaller values of 0.3−0.5 kcal/mol; however, the ON/OFF difference is unchanged, remaining equal to 0.1 kcal/mol. The total contribution of the outer protein region is essentially negligible.

3.4. aIF2:GTP/GDP Binding: MDFE Results and Convergence. Having analyzed the uncertainty associated with our choice of force field and our use of a truncated protein model, we have used this efficient setup to perform multiple MDFE runs comparing GTP and GDP binding to aIF2. We are especially interested in the convergence of the simulations and the reproducibility of the results. The structures of the various complexes are quite stable during the simulations, with rms deviations from the relevant crystal structure (either ON- or OFF-aIF2) of just 0.6−0.9 Å for backbone atoms in the MD region (values averaged over the last nanosecond of a 5 ns simulation of each complex, excluding portions of the backbone that are harmonically restrained). The individual runs are detailed in Table 4. The electrostatic and van der Waals steps are shown separately. For each one, the uncertainty was estimated by comparing results obtained with the first and second halves of each window. For the total free energy change, the electrostatic and van der Waals uncertainties were propagated. The free energy derivatives are shown as a function of the coupling constants in Figure 6 and Table 5. The agreement between runs is excellent, well within the statistical uncertainty, not only for the overall free energy changes but also for the free energy derivatives at each λ value. The main exception, where the derivatives are significantly different in different runs, is in the $\lambda = 0.4$ region of the electrostatic step for the OFF state. Here, the uncertainty for the derivative is very large in run 1, ±30.1 kcal/mol. To understand the origin of this uncertainty, the original MDFE window in run 1 was extended, from 2000 to 4000 ps. Examining the conformations sampled during the MD trajectory, the uncertainty can be attributed to rotations of the terminal PO_3 of the GTP ligand, which oscillates back and forth between two states. This has a large effect on the Mg^{2+}:phosphate interactions. The oscillations are illustrated in Figure 7, where the sum of the inverse distances between Mg^{2+} and the three terminal oxygens of GTP is plotted with respect to simulation time. The transitions between states are frequent enough to allow identification of the states but not to provide a highly converged Boltzmann average for the free energy derivative. Despite this uncertainty, we can see that, when the two runs are averaged, the resulting free energy derivative curve is very smooth, suggesting that the average over the two runs is probably accurate in the $\lambda = 0.4$ region. Ten other windows in the four different runs led initially to a high statistical uncertainty, and were systematically extended by adding 2 ns to their production length. In all cases except the window discussed above, the uncertainty was reduced to a much more acceptable level, as indicated in Table 5.

Comparing the GTP → GDP free energy differences in the ON and OFF states, we obtain a binding free energy difference of 4.2 kcal/mol, favoring GTP binding to the ON state. This value is an average over two forward runs in both the ON and OFF states (Table 4), and also contains a −0.1 kcal/mol contribution from the distant protein regions (treated as a dielectric continuum; Table 3). The experimental binding free energy difference is not available, but the sign obtained here is correct, since the ON state is known to preferentially bind GTP, and the magnitude appears biochemically plausible. It is slightly larger than the force field uncertainty and the statistical uncertainty taken separately (2−3 kcal/mol each) but approximately equal to the overall, combined uncertainty. The force field sensitivity analysis suggests that Charmm has a tendency to overstabilize GDP in the protein (Table 2), but there should be substantial cancellation of the error for the ON/OFF difference. A detailed analyis of the structural basis for the binding free energy differences is outside the scope of this paper, and will be reported elsewhere.

4. CONCLUDING DISCUSSION

ATPases and GTPases are central actors in bioenergetics, signaling, and many other cellular processes. The archaeal initiation factor aIF2 serves here as a paradigm for these proteins, as well as for other classes of protein:nucleotide interactions. To elucidate these interactions with MD free energy simulations, three main difficulties must be overcome. The treatment of strong electrostatic interactions, such as those between divalent Mg^{2+} and phosphate groups, must be sufficiently accurate. Sufficient conformational sampling must be achieved. Finally, the treatment of long-range interactions and boundary conditions should be efficient but also accurate.

The most efficient simulation approach is to use a fixed charge, classical mechanical force field and a hybrid simulation model, with a detailed, atomistic treatment of regions close to the nucleotide ligand but a simplified treatment of more distant regions. To determine the suitability of the fixed charge force field, we first studied several small model systems involving Mg^{2+}:phosphate interactions in water, using both a fixed charge force field and quantum chemical methods, and comparing to experimental binding free energies. The fixed charge force field gave very large errors for Mg^{2+}:phosphate interactions; for example, the standard Mg^{2+}:monohydrogen-phosphate binding free energy was overestimated by a factor of 10 (−41 kcal/mol vs −3.7 kcal/mol from experiment). This large error can be presumably attributed in large part to the lack of explicit electronic polarizability in the fixed charge model.[38−40] On the quantum chemical side, smaller but still significant errors were obtained for the dianionic monohydrogen phosphate using DFT with a B3LYP-D functional, a large, aug-cc-pVTZ basis set, and a dielectric continuum solvent model. The computed binding free energy was −10.8 kcal/mol for Mg^{2+}:monohydrogen-phosphate, which corresponds to a 7.1 kcal/mol error. This error level was achieved in part through an *ad hoc* procedure, where nine explicit waters were added to the model. While the water placement was plausible, as shown by the good accuracy obtained for Mg^{2+}:DHP binding, its predictive value for other cases is not completely clear. Part of the remaining error undoubtedly comes from the simplified solvent treatment of the nine explicit waters, as well as the remaining, dielectric continuum water. As usual with this methodology, the rotational degrees of freedom of the explicit waters could not be rigorously sampled, for reasons of cost. Rather, the binding free energy estimate only considered a few minimum energy basins (structures of types I, II, and III) in conformational space, explored by normal mode calculations, whereas other microsolvation configurations were neglected. For the other solvent molecules, a dielectric continuum treatment was used, which is known to be qualitatively but not quantitatively accurate. In particular, nonlinear effects are known to occur in the vicinity of divalent ions.[93−95] Also, the atomic radii used to define the solute/solvent boundary should in principle have a dependency on the local charge density,[96,97] although this effect is certainly attenuated here by the use of nine explicit waters. Overall, more work is needed to fully explore the potential and limitations of the quantum mechanical model for these systems.

To compare GTP and GDP binding to aIF2, we computed the free energies for the horizontal, alchemical legs in the thermodynamic cycle of Figure 1, which transform GTP into GDP. This transformation removes a phosphate group and thus a Mg^{2+}:phosphate interaction. Even though the fixed charge force field gave large errors for the absolute, Mg^{2+}:phosphate binding free energies, much of the error cancels out when we compare the GTP → GDP transformation in protein and in solution. The cancellation is even better when we compare the two protein conformational states: ON and OFF. Taking the ON/OFF difference for the free energy estimator Δg, the Charmm and Amoeba results differ by 1.2 ± 3.7 kcal/mol (Table 2). This is smaller than the statistical uncertainty associated with the finite simulation length, 2−3 kcal/mol (Table 4). This situation can be compared to other biochemical transformations where a charge is inserted either into a protein or into solution, such as redox potential shift or pK_a shift calculations: fixed charge force fields often give good agreement with experiment,[30,98] although an explicit treatment of polarazability can obviously be necessary for some cases or for specific free energy components (such as the reorganization energy).[99]

To reduce the statistical error to such an acceptable level, we collected more than 200 ns of MD for the complete set of MDFE windows. To perform such long MD simulations, an efficient simulation setup was needed. We considered explicitly a spherical subset of the protein, with a 26 Å radius, solvated by a cubic water box with a 74 Å edge length. With this box size, interactions between GTP/GDP ligands in neighboring boxes are very small, and the ionic concentration is small enough so that nonideal, Debye−Hückel corrections to the free energies can be neglected. Notice also that, although the GTP → GDP transformation changes the total charge of the molecular system, the use of tinfoil boundary conditions ensures that, as the charge changes, a compensating charge density arises automatically, which is spread uniformly throughout the simulation box and does not contribute to the atomic forces.[58,59] In a separate stage, the missing protein region is reintroduced with a dielectric continuum model; its contribution to the GTP → GDP free energy change is negligible.

Overall, when we compare the GTP/GDP binding free energy differences in the ON state, we obtain a +0.3 kcal/mol difference, favoring GTP binding. For the OFF state, we obtain −4.0 kcal/mol, disfavoring GTP and favoring GDP. As expected, the ON state is predicted to prefer GTP while the OFF state prefers GDP. The difference between these two effects is presumably more reliable than either one separately, because the force field errors cancel to a larger extent when ON- and OFF-aIF2 are compared. For this difference, we obtain 4.2 kcal/mol. The precise magnitude of all these preferences is not known

experimentally. Here, despite the significant uncertainty associated with the force field and the finite sampling, we obtain a somewhat more precise picture. A more detailed interpretation of the free energy differences can be made, in principle, by analyzing the contributions of individual interactions, by means of a free energy component analysis, for example.[30,100] Such a detailed, biochemical analysis will be reported in another paper. Here, our main focus was to establish the uncertainty and reliability of the MDFE approach for comparing GTP and GDP binding to a GTPase. We expect that the observations reported here will be relevant for many other proteins of this class.

AUTHOR INFORMATION

Corresponding Author
*E-mail: thomas.simonson@polytechnique.fr.

ABBREVIATIONS

aIF2, archaeal initiation factor 2; PDB, Protein Data Bank; MDFE, molecular dynamics free energy; PB, Poisson−Boltzmann; FF, force field

REFERENCES

(1) Perutz, M. *Mechanisms of cooperativity and allosteric regulation in proteins*; Cambridge University Press: Cambridge, U.K., 1990.
(2) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
(3) Grant, B. J.; Gorfe, A. A.; McCammon, J. A. *Curr. Opin. Struct. Biol.* **2010**, *20*, 142–147.
(4) Sprang, S. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 849–856.
(5) Vetter, I. R.; Wittinghofer, A. *Science* **2001**, *294*, 1299–1304.
(6) Boyer, P. D. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2296–2307.
(7) Walker, J. E. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2309–2319.
(8) Hauryliuk, V. V. *Mol. Biol.* **2006**, *40*, 688–701.
(9) Yatime, L.; Méchulam, Y.; Blanquet, S.; Schmitt, E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18445–18450.
(10) Myasnikov, A. G.; Simonetti, A.; Marzi, S.; Klaholz, B. P. *Curr. Opin. Struct. Biol.* **2009**, *19*, 300–309.
(11) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646–651.
(12) Gorfe, A. A.; Grant, B. J.; McCammon, J. A. *Structure* **2008**, *16*, 885–896.
(13) Grant, B. J.; Gorfe, A. A.; McCammon, J. A. *PLoS Comput. Biol.* **2009**, *5*, e1000325.
(14) Grant, B. J.; McCammon, J. A.; Gorfe, A. A. *Biophys. J.* **2010**, *99*, 87–89.
(15) Lukman, S.; Grant, B. J.; Gorfe, A. A.; Grant, G. H.; McCammon, J. A. *PLoS Comput. Biol.* **2010**, *6*, e1000922.
(16) Wang, Y.; Sadreyev, R. I.; Grishin, N. V. *Nucleic Acids Res.* **2009**, *37*, 3522–3530.
(17) Friedman, Z. Y.; Devary, Y. *Proteins* **2005**, *59*, 528–533.
(18) Ng, Y. W.; Raghunathan, D.; Chan, P. M.; Baskaran, Y.; Smith, D. J.; Lee, C. H.; Verma, C.; Manser, E. *Structure* **2010**, *18*, 879–890.
(19) Lu, B.; Wong, C. F.; McCammon, J. A. *Protein Sci.* **2005**, *14*, 159–168.
(20) Khavrutskii, I. V.; Grant, B.; Taylor, S. S.; McCammon, J. A. *Biochemistry* **2009**, *48*, 11532–11545.
(21) Fischer, S.; Windshügel, B.; Horak, D.; Holmes, K. C.; Smith, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6873–6878.
(22) Koppole, S.; Smith, J. C.; Fischer, S. *Structure* **2007**, *15*, 825–837.
(23) Yu, H.; Ma, L.; Yang, Y.; Cui, Q. *PLoS Comput. Biol.* **2007**, *3*, 214–230.
(24) Cecchini, M.; Houdusse, A.; Karplus, M. *PLoS Comput. Biol.* **2008**, *4*, e1000129.
(25) Pu, J. Z.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1192–1197.
(26) Yang, Y.; Yu, H. B.; Cui, Q. J. *Mol. Biol.* **2008**, *381*, 1407–1420.
(27) Grigorenko, B. L.; Nemukhin, A. V.; Shadrina, M. S.; Topol, I. A.; Burt, S. K. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 456–466.
(28) Heesen, H. T.; Gerwert, K.; Schlitter, J. *FEBS Lett.* **2007**, *581*, 5677–5684.
(29) McCammon, J. *Curr. Opin. Struct. Biol.* **1998**, *8*, 245–249.
(30) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430–437.
(31) Jorgensen, W. L. *Science* **2003**, *303*, 1813–1818.
(32) Chipot, C.; Mark, A. E.; Pande, V. S.; Simonson, T. In *Free energy calculations: theory and applications in chemistry and biology*; Chipot, C., Pohorille, A., Eds.; Springer Verlag: New York, 2007; Chapter 11.
(33) Hritz, J.; Lappchen, T.; Oostenbrink, C. *Eur. Biophys. J.* **2010**, *39*, 1573–1580.
(34) Thompson, D.; Simonson, T. *J. Biol. Chem.* **2006**, *281*, 23792–23803.
(35) Aleksandrov, A.; Thompson, D.; Simonson, T. *J. Mol. Recognit.* **2010**, *23*, 117–127.
(36) Yang, W.; Gao, Y. Q.; Cui, Q.; Ma, J.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 874–879.
(37) Gao, Y. Q.; Yang, W.; Karplus, M. *Cell* **2005**, *123*, 195–205.
(38) Ponder, J.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27.
(39) Warshel, A.; Kato, M.; Pisliakov, A. V. *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.
(40) Jiao, D.; Golubkov, P. A.; Darden, T. A.; Ren, P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6290–6295.
(41) Piquemal, J. P.; Perera, L.; Cisneros, G. A.; Ren, P. Y.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, No. 054511.
(42) Simonson, T. *Rep. Prog. Phys.* **2003**, *66*, 737–787.
(43) Beglov, D.; Roux, B. *J. Chem. Phys.* **1994**, *100*, 9050–9063.
(44) Roux, B.; Beglov, D.; Im, W. In *Simulation and theory of electrostatic interactions in solution*; Pratt, L., Hummer, G., Eds.; American Institute of Physics: College Park, Maryland, 1999; pp 473−491.
(45) Im, W.; Bernèche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924–2937.
(46) Simonson, T. *J. Phys. Chem. B* **2000**, *104*, 6509–6513.
(47) Warshel, A. *Computer modelling of chemical reactions in enzymes and solutions*; John Wiley: New York, 1991.
(48) Simonson, T. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243–252.
(49) Schaefer, M.; Vlijmen, H. W. T. v.; Karplus, M. *Adv. Protein Chem.* **1998**, *51*, 1–57.
(50) Tembe, B.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281–283.
(51) Simonson, T. In *Computational Biochemistry & Biophysics*; Becker, O., Mackerell, A., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; Chapter 9.
(52) Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(53) Ponder, J. W.; Wu, C. J.; Ren, P. Y.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; I, I. H.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
(54) Ren, P.; Ponder, J. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.
(55) Cornell, W.; Cieplak, P.; Bayly, C.; Gould, I.; Merz, K.; Ferguson, D.; Spellmeyer, D.; Fox, T.; Caldwell, J.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(56) Yatime, L.; Mechulam, Y.; Blanquet, S.; Schmitt, E. *Structure* **2006**, *14*, 119–128.
(57) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(58) Hummer, G.; Pratt, L.; Garcia, A. *J. Phys. Chem.* **1996**, *100*, 1206–1215.
(59) Bogusz, S.; Cheatham, T. E.; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070–7084.
(60) Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.
(61) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(62) Brooks, B.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
(63) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
(64) Verbeeck, R. M. H.; Bruyne, P.; Driessens, F. C. M.; Verbeek, F. *Inorg. Chem.* **1984**, *23*, 1922–1926.
(65) Storer, A. C.; Cornish-Bowden, A. *Biochem. J.* **1976**, *159*, 1–5.
(66) Alberty, R. A.; Goldberg, R. N. *Biochemistry* **1992**, *31*, 10610–10615.
(67) Cini, R.; Burla, M. C.; Nunzi, A.; Polidori, G. P.; Zanazzi, P. F. *J. Chem. Soc., Dalton Trans.* **1984**, *11*, 2467–2476.
(68) Akola, J.; Jones, R. O. *J. Phys. Chem. B* **2006**, *110*, 8110–8120.
(69) Simonson, T. *Mol. Phys.* **1993**, *80*, 441–447.
(70) Fowler, R. H.; Guggenheim, E. A. *Statistical Thermodynamics*; Cambridge University Press: Cambridge, U.K. 1939.
(71) Koch, W.; Holthausen, M. C. *A chemist's guide to density functional theory*; Wiley: New York, 2001.
(72) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.
(73) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
(74) Klamt, A.; Schürmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799–805.
(75) Klamt, A.; Jonas, V. *J. Chem. Phys.* **1996**, *105*, 9972–9981.
(76) TURBOMOLE V6.2 2010, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
(77) The pymol molecular graphics system. DeLano, W. L.; DeLano Scientific: San Carlos, CA, 2002.
(78) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.
(79) Tinker software tools for molecular design. Jay Ponder Lab.; Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Saint Louis, MO, 1987−2011.
(80) Im, W.; Beglov, D.; Roux, B. *Comput. Phys. Commun.* **1998**, *111*, 59–75.
(81) Gilson, M.; Honig, B. *Proteins* **1988**, *4*, 7–18.
(82) Stangret, J.; Savoie, R. *Can. J. Chem.* **1992**, *70*, 2875–2883.
(83) Rulisek, L.; Sponer, J. *J. Phys. Chem. B* **2003**, *107*, 1913–1923.
(84) Petrov, A. S.; Pack, G. R.; Lamm, G. *J. Phys. Chem. B* **2004**, *108*, 6072–6081.
(85) Allen, F.; Kennard, O. *Chem. Des. Autom. News* **1993**, *8*, 31–37.
(86) Ezra, F. S.; Collin, R. L. *Acta Crystallogr., Sect. B* **1973**, *29*, 1398–1403.
(87) Hunenberger, P.; McCammon, J. *Biophys. Chem.* **1999**, *78*, 69–88.
(88) Liao, J.-C.; Sun, S.; Chandler, D.; Oster, G. *Eur. J. Biophys.* **2004**, *33*, 29–37.
(89) Freisinger, E.; Sigel, R. *Coord. Chem. Rev.* **2007**, *251*, 1834–1851.
(90) Thompson, D.; Plateau, P.; Simonson, T. *ChemBioChem* **2006**, *7*, 337–344.
(91) Friedman, Z. Y.; Devary, Y. *J. Biol. Chem.* **2009**, *284*, 13513–13518.
(92) Simonson, T. In *Free energy calculations: theory and applications in chemistry and biology*; Chipot, C., Pohorille, A., Eds.; Springer Verlag: New York, 2007; Chapter 12.
(93) Papazyan, A.; Warshel, A. *J. Chem. Phys.* **1997**, *107*, 7975–7978.
(94) Sandberg, L.; Edholm, O. *J. Chem. Phys.* **2002**, *116*, 2936–2944.
(95) Gong, H.; Hocky, G. M.; Freed, K. F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11146–11150.
(96) Tomasi, J. *Chem. Rev.* **2005**, *105*, 2999.
(97) Zhou, B. J.; Agarwal, M.; Wong, C. F. *J. Chem. Phys.* **2008**, *129*, No. 014509.
(98) Blumberger, J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5651–5667.
(99) Tipmanee, V.; Oberhofer, H.; Park, M.; Kim, K. S.; Blumberger, J. *J. Am. Chem. Soc.* **2010**, *132*, 17032–17040.
(100) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25–33.