This is an open access article published under a Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

ACS AUTHORCHOICE

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

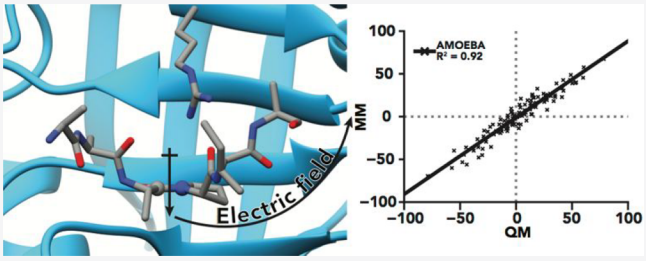
The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields?

Richard T. Bradshaw,*,▽ Jacek Dziedzic,▽ Chris-Kriton Skylaris, and Jonathan W. Essex*

Cite This: J. Chem. Inf. Model. 2020, 60, 3131−3144 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Preorganization of large, directionally oriented, electric fields inside protein active sites has been proposed as a crucial contributor to catalytic mechanism in many enzymes, and it may be efficiently investigated at the atomistic level with molecular dynamics simulations. Here, we evaluate the ability of the AMOEBA polarizable force field, as well as the additive Amber ff14SB and Charmm C36m models, to describe the electric fields present inside the active site of the peptidyl-prolyl isomerase cyclophilin A. We compare the molecular mechanical electric fields to those calculated with a fully first-principles quantum mechanical (QM) representation of the protein, solvent, and ions, and find that AMOEBA consistently shows far greater correlation with the QM electric fields than either of the additive force fields tested. Catalytically relevant fields calculated with AMOEBA were typically smaller than those observed with additive potentials, but were generally consistent with an electrostatically driven mechanism for catalysis. Our results highlight the accuracy and the potential advantages of using polarizable force fields in systems where accurate electrostatics may be crucial for providing mechanistic insights.



■ INTRODUCTION

Over the past 30 years, molecular simulations have become an integral and routine part of understanding biological processes at the molecular level. Advances in both CPU and GPU hardware, and more reliable and sustainable software packages, have led to both wider uptake and greater numbers of applications across multiple fields.[1] Microsecond-length simulations are now routine, allowing exploration of biomolecular dynamics on time scales relevant for a variety of biological processes, whether through conventional simulation, enhanced sampling, or kinetic modeling of the underlying process.

Simultaneously, the major families of biomolecular force fields have undergone multiple rounds of development and improvement.[2−7] For any force field, this reparametrization process requires significant time and resource investment to identify errors, develop corrections, and validate the accuracy of the new parameters. The functional form of each force field, which provides the framework for parameter optimization, has therefore remained relatively unchanged and similar between families for decades.[8,9] With small variations between force field families, each typically features a simple harmonic representation of bonds and angles, a Fourier series for proper dihedrals, a Lennard-Jones model of van der Waals (vdW) interactions, and most importantly, an additive, fixed-point-charge representation of electrostatic interactions.[10−12] The philosophy for parametrization of these fixed-point charges varies between families and has seen substantial modification between force field iterations, but generally entails a significant focus on fitting to either molecular electrostatic potentials (for example, side-chain fragments), interaction energies, or derived thermodynamic properties.[10−14] For a given biomolecule, these electrostatic and vdW parameters are then relied upon to accurately re-create interatomic interactions in all possible environments.

Given the diversity of biomolecular environments and the relative simplicity of common potential functions, it is perhaps surprising that biomolecular force fields have seen such broad and successful applications.[1,15−17] However, the latest additive force fields feature a delicate balance of force-field terms built up over decades, that continue to evolve to be usefully applied in new contexts. Where force fields have failed, parameter improvements have often focused on tweaks to backbone and side-chain torsional potentials,[6,18] but have also included changes to the Lennard-Jones potential,[19] pair-specific corrections to nonbonded interactions in defined environments,[20] adaptations of parametrization protocols to develop implicitly polarized partial charges,[21,22] and even application-specific water models.[23]

Received: March 2, 2020
Published: April 16, 2020

Downloaded via GDANSK UNIVERSITY OF TECHNOLOGY on March 15, 2024 at 10:43:11 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Nevertheless, fixed, atom-centered partial charges suffer from known deficiencies in their representation of intermolecular interactions.[24] First, molecular electrostatic potentials, which are often the main target for fitting electrostatic parameters, are more accurately re-created if higher-order, atom-centered multipoles are included in the fitted potential function.[25,26] Second, additive models do not include an explicit representation of induction, and may therefore poorly represent electrostatics for molecular conformations, functional groups, or environments not well-validated in the parametrization process.[27−30] This lack of transferability may pose challenges in the use of force fields for new, untested applications.

Alternative, more-advanced models of electrostatics have been implemented in new families of biomolecular force fields,[31,32] of which the Charmm Drude[33] and AMOEBA force fields[34] are among the most extensively developed. Both include explicit representation of mutual polarization, and AMOEBA also incorporates a classical multipole representation of fixed electrostatics. These potential functions have been repeatedly shown to improve representation of electrostatic properties around biologically relevant moieties, and they can improve the reliability and convergence properties of hybrid quantum mechanical/molecular mechanical (QM/MM) studies.[35−44] Nevertheless, the accuracy advantages of polarizable force fields in the context of real protein environments are still poorly understood, and should be explored in systems where quantitatively accurate intramolecular electrostatics may be crucial to provide mechanistic insight. Biological catalysis is a particular focus in this regard, since, in both naturally occurring and designed systems, enzyme active sites may preorient substrates and stabilize reaction transition states with focused electric fields or other electrostatic motifs.[45,46] Although studies of the enzyme-catalyzed chemical reactions require QM/MM techniques, the reactant and product states can be studied efficiently with purely MM models, provided the underlying electrostatic model is sufficiently accurate to provide useful information on the active site environment.

Given this interest, here, we test the ability of biomolecular force fields to reproduce the electric fields inside an entire protein and its associated solvent and ionic environment, using the peptidyl-prolyl isomerase cyclophilin A (CypA) as a model system. We choose to evaluate the accuracy of the AMOEBA polarizable force field, as it has recently been promisingly used in proof-of-concept work for enzyme design.[47,48] Although ultimately successful in redesigning active site interactions, the additional computational cost of AMOEBA has prevented lengthy studies of active site dynamics, and there has been little assessment so far of the physical accuracy of AMOEBA electric fields within the folded protein environment over typical simulation time scales.

Assessing this physical accuracy by comparison to experiment has substantial difficulties. Biophysical techniques for measuring intramolecular electric fields, such as vibrational Stark effect spectroscopy, can provide precise and sensitive measurements of electric fields inside biomolecular environments, but remain technically challenging and may require the introduction of unnatural amino acids to act as vibrational probes.[49] Instead, we evaluate the accuracy of AMOEBA electrostatics by comparison to a fully quantum mechanical representation, calculated using the ONETEP linear-scaling density functional theory (DFT) software.[50] ONETEP is based on a reformulation of DFT, which takes advantage of the locality of electronic structure to allow DFT calculations with computational effort that increases linearly with the number of atoms, as compared to conventional DFT approaches where the computational effort increases with the third power. As a result, ONETEP can perform calculations on many thousands of atoms, such as entire proteins in solvent, as we do here. A unique characteristic of ONETEP is that it retains the full near-complete basis set accuracy of conventional DFT via *in situ* optimization of Nonorthogonal Generalized Wannier Functions (NGWFs);[51] these are expressed in terms of a periodic sinc (psinc) basis set, which is equivalent to plane waves. The ONETEP program has been developed to run on parallel computers, using hybrid MPI-OpenMP parallelism.[52]

To test the ability of biomolecular force fields to reproduce the electric fields derived from DFT, we first perform ns-length AMOEBA molecular dynamics (MD) simulations of CypA to generate structural ensembles for wild-type (WT) and mutant systems. We then calculate electric fields at the reaction site using AMOEBA,[53] and the widely used Amber ff14SB[2] and Charmm C36m[3] additive protein force fields. Each set of electric fields is then compared with those calculated at the DFT level with ONETEP. We find that AMOEBA electric fields show far greater correlation to DFT than those of additive potentials, and with smaller systematic error. Our results highlight the utility of polarizable potentials in biomolecular environments, where the accurate reproduction of electrostatics is crucial. If mechanistic insights are desired, recent high-performance software implementations obviate the time scale limitations of standard polarizable simulations, to some extent,[54−56] but cross-validation with experimental data would remain highly desirable to complement the computational predictions.

METHODS

CypA System Preparation. The initial structures for simulated systems were taken from crystal structures of wild-type human CypA bound to HIV capsid protein (PDB entries 1M9C and 1M9Y).[57] In both cases, chain A was taken as the structure of the CypA receptor, and the substrate was trimmed to a six-residue model peptide, of sequence HAGPIA (from PDB 1M9C) (with the G−P peptide bond in the trans configuration), or AAAPIA (from PDB 1M9Y) (with the A−P peptide bond in the cis configuration). Substrate N and C termini were not capped and peptides were modeled in their zwitterionic states. Basic and acidic residues were modeled in their ionized states and histidine residues modeled in their neutral, N_δ-protonated form. Finally, R55A mutant structures with equivalent cis- and trans-proline substrates were created by manually truncating the Arg55 residue in the crystal structures to the C_β atom.

CypA MD Simulations. Simulations of both WT and R55A mutant CypA, bound to cis- and trans-proline substrates, were performed in triplicate using the AMOEBA polarizable force field. A control simulation of WT CypA bound to the cis-proline substrate was also performed with the Amber ff14SB additive force field. See Table S1 in the Supporting Information for a list of all simulations performed. Initial protein crystal structures were protonated and solvated using the tleap module of the Amber14 package.[58] CypA-peptide complexes were neutralized with Cl^- anions, solvated in an ∼63 Å × 65 Å × 62 Å periodic box of water, such that no solute atom was <8.0 Å from the box edge, and Na^+/Cl^- ions were added to create a 150 mM NaCl ionic atmosphere. The

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

simulation of the WT CypA cis-proline system performed with the Amber additive force field used the same starting configuration as the first AMOEBA replicate, except ion positions in bulk solvent were randomized. In total, system sizes varied from 17 727 atoms (R55A trans-proline system) to 19 030 atoms (R55A cis-proline system).

AMOEBA simulations used the AMOEBA 2013 force field for protein and ions, and the AMOEBA 2003 water model.[53,59] Simulations were performed in triplicate, with unique random seeds for the thermostat and barostat. Systems were energy-minimized for 2500 steps with a steepest descent algorithm, then heated to 300 K over 50 ps in the NVT ensemble, followed by equilibration to 1.0 bar over 100 ps in the NPT ensemble. Production simulations were then performed for a total of 25 ns each, with trajectory frames saved at 10 ps intervals for a total of 2500 frames per trajectory. An Andersen thermostat and Monte Carlo barostat were used to maintain a temperature of 300 K and a pressure of 1 bar throughout. A velocity Verlet integrator with a 1 fs time step was used in all simulations. Long-range electrostatic interactions were treated with a Particle Mesh Ewald summation with a 7.0 Å real-space cutoff. vdW interactions were subject to a 9 Å cutoff with an analytical long-range correction. Induced dipole convergence was set to 10^{-5} D/atom. All MD simulations were performed with the Tinker-OpenMM plugin, using Tinker version 7.1, and OpenMM version 6.3.[55,60]

The WT cis-proline Amber simulation used the ff14SB force field for protein, TIP3P model for water, and Joung and Cheatham parameters for ions.[2,61,62] The system was energy-minimized for 5000 steps with a steepest descent algorithm, followed by 5000 steps with a conjugate gradient algorithm. Next, the system was heated to 300 K over 50 ps in the NVT ensemble, equilibrated to 1.0 bar over 100 ps in the NPT ensemble, and simulated for an additional 25 ns under NPT conditions, with trajectory frames saved every 10 ps for a total of 2500 per trajectory. A Langevin thermostat and Berendsen barostat were used to maintain temperature and pressure. A 2 fs time step was used throughout, and the SHAKE algorithm used to constrain all bonds involving hydrogen.[63] Long-range electrostatic interactions were treated with a Particle Mesh Ewald summation with an 8.0 Å real-space cutoff, while vdW interactions were calculated with an 8 Å cutoff and long-range analytical correction. MD simulations were performed with the pmemd.cuda module of the Amber14 software.[58]

MM Field Calculations. Using the 2500 frames extracted from each trajectory, electric fields were calculated using the AMOEBA 2013, Amber ff14SB, and Charmm C36m force fields. The electric fields experienced by the peptide substrate were evaluated by calculating the environmental field, $\vec{E}_{\mathrm{env}}$, following the approach of Fried and Wang:[64]

$$\vec{E}_{\mathrm{env}} = \vec{E}_{\mathrm{complex}} - \vec{E}_{\mathrm{ligand}} \tag{1}$$

Here, $\vec{E}_{\mathrm{complex}}$ and $\vec{E}_{\mathrm{ligand}}$ refer to the electric field at a given site in the fully solvated protein−ligand complex, and in a ligand-only system stripped of receptor and solvent, respectively. Fields were first calculated at the two atoms of the isomerized peptide bond—the proline N atom and the carbonyl C atom of the preceding residue (Figure 1). Atom-centered fields were then linearly interpolated to provide the field experienced at the bond midpoint. After AMOEBA simulations, the electric fields $\vec{E}_{\mathrm{complex}}$ at the C and N atoms were calculated directly for each trajectory frame using the instantaneous induced atomic dipole at each atomic site, $\vec{\mu}_{\mathrm{ind}}$, and the relevant atomic polarizability taken from the AMOEBA force field, α:

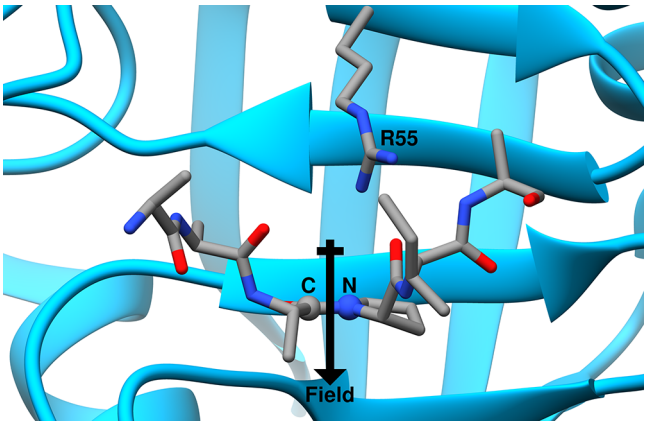


Figure 1. Orientation of CypA (blue cartoon) and AAAPIA truncated peptide substrate (gray sticks) in a cis-proline conformation generated from PDB entry 1M9Y. Residue R55 of CypA is positioned approximately above the peptide proline ring, and, in the ground state, interacts with the proline carbonyl oxygen atom. Substrate C and N atoms for the calculation of environment field are depicted as balls, and the approximate orientation of the proposed environmental electric field is overlaid with a black arrow.

$$\vec{E}_{\mathrm{complex,polarizable}} = \frac{\vec{\mu}_{\mathrm{ind}}}{\alpha} \tag{2}$$

For additive force fields, the electric field was calculated by dividing the electrostatic force, $\vec{F}_{\mathrm{ele}}$, exerted on the desired atom, by its partial charge, q:

$$\vec{E}_{\mathrm{complex,additive}} = \frac{\vec{F}_{\mathrm{ele}}}{q} \tag{3}$$

For all force fields, post-processing of trajectory frames to remove the CypA receptor, solvent, and ion coordinates allowed the equivalent calculation of electric fields in the ligand-only system, $\vec{E}_{\mathrm{ligand}}$. Finally, the electric field arising from the protein/solvent/ionic atmosphere of the system and acting on the rotatable amide bond, $\vec{E}_{\mathrm{env}}$, was calculated via eq 1.

For each trajectory structure, fields were calculated and compared across force fields in two distinct ways. First, x, y, and z field components at the substrate proline N and preceding C atom were evaluated directly in the Cartesian coordinate frame. Second, the field at each atom was also projected along a vector perpendicular to the plane of the substrate proline ring, calculated individually for each frame as the unit-length cross product of the N−C_α and N−C_δ bond vectors. This vector approximates the proposed orientation of the peptidyl carbonyl group at the cis−trans transition state.[65] Therefore, the projection of the substrate N and C field vectors along the same direction estimates the magnitude of the environmental electric field aligned with the peptidyl carbonyl group during the peptide bond rotation. Finally, N and C atom projections were averaged to interpolate the field magnitude at the bond midpoint. These two distinct approaches compare (1) the full electric field vectors and (2) the electric field magnitude oriented along a catalytically relevant direction.

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

For the additive models, electric fields were calculated using an in-house python script making use of the Amber *sander* python API. For the AMOEBA model, fields were calculated using a separate in-house python script, linking to the *analyze* module of Tinker 7.1.

Finally, 10 frames were extracted at 2.5 ns intervals from the first AMOEBA simulation performed for each of the four CypA-substrate systems, for comparison with ONETEP DFT fields.

ONETEP Field Calculations. All DFT calculations in this work were performed with the ONETEP linear-scaling DFT package in the norm-conserving pseudopotential approximation. The PBE exchange-correlation functional was used. A minimal *in-situ*-optimized NGWF basis was employed, with an 8.0 a_0 (∼4.23 Å) localization radius. Density kernel truncation was not applied.

For consistency with MM calculations, periodic boundary conditions were employed throughout. The MM frames were not truncated in any way. A plane-wave kinetic energy (k.e.) cutoff of 800 eV was assumed. Given that the dimensions of the simulation cell varied slightly between MM frames (due to the use of NPT conditions), and that the cell edges must be divisible into an integer number of grid points, the actual k.e. cutoffs were between 798 eV and 832 eV. The resultant subtle difference in basis set quality between frames was neglected.

To calculate electric fields, we first calculated electrostatic potentials on a Cartesian grid with a spacing of ∼0.13 Å (a so-called double grid, with a spacing sufficiently fine to represent densities). Two separate components to the potentials were considered—one due to the valence-electronic pseudodensity (Hartree potential), and one due to the local pseudopotential. Both components were obtained in reciprocal space using fast Fourier transforms (FFTs) using standard ONETEP methodology. Exchange and correlation potentials, representing nonclassical effects, were excluded from analysis. Electric fields ($E(\vec{g})$) were obtained from the potential ($U(\vec{g})$) in reciprocal space, at the reciprocal space vector $\vec{g}$ where the gradient operator is simply $i\vec{g}$:

$$\vec{E}(\vec{g}) = -i\vec{g}U(\vec{g}) \tag{4}$$

and subsequently transformed back to real space via an inverse FFT. Final electric fields at arbitrary points in space were obtained via trilinear interpolation from the nearest grid points.

Quantities that underwent FFT processing on a double grid suffer from a small amount of ringing artifacts with a period of the original (single) grid spacing. While the magnitude of the ringing is small, we are investigating sums of two components to the electric field (Hartree and pseudopotential) that have opposite signs and similar magnitudes, thus canceling out to a large degree. This magnifies relative errors in the total electric field. To improve the accuracy of the calculated fields, we smoothed the field values on the grid using a 27-point three-dimensional (3D) stencil, where the stencil weights were given by $\frac{1}{8}(2^{-d_1})$, where d_1 is the taxicab distance of a stencil point from the middle of the stencil. This procedure was found to efficiently attenuate the ringing artifacts.

The use of Cartesian grids in the DFT calculations with localized orbitals and in subsequent processing is expected to subtly break rotational invariance (by distinguishing certain spatial directions corresponding to grid edges) and translational invariance (the so-called "eggbox effect"). These effects are minimized in ONETEP, where the local orbitals (NGWFs) are optimized *in situ*, but they are not completely eliminated.[66] To alleviate any concerns over whether these have a measurable effect on the obtained electric fields, we performed six additional calculations for one of the MM frames, where the systems were variously rotated and translated relative to the grid. Typical error magnitudes were found to be in the order of 0.1 MV/cm, with a maximum below 0.4 MV/cm, making them negligible in our analysis.

Data Availability. All underlying data used for this study is freely available (DOI: 10.5281/zenodo.3678278), including the simulation trajectories, calculated electric fields, and analysis code for calculating fields from AMOEBA and fixed-charge simulations. The code and underlying data used to create figures is also available in this repository.

RESULTS

CypA as an Exemplar for Electrostatically Driven Enzyme Mechanisms. The peptidyl-prolyl isomerase CypA catalyzes the cis/trans isomerization of the amide preceding proline residues in proteins, inducing a structural change in peptide chains that would otherwise be extremely slow under physiological conditions. This fundamental catalyst for protein conformational change is ubiquitously expressed and plays a role in a wide variety of biomolecular mechanisms, including protein folding, trafficking, and signaling and regulation. Equally, however, CypA has been implicated in a wide variety of disease processes, particularly the facilitation of viral infection and replication.[67]

Because of this broad spectrum of activity and therapeutic interest, the CypA structure, dynamics, and catalytic mechanism have been well-studied for decades using both experimental and computational structural biology approaches.[65,68−74] Many of these studies have identified residue R55, conserved across the cyclophilin family, as mechanistically crucial—either a R55A or R55K mutation resulting in a similar reduction of catalytic efficiency.[71,75] In the case of the lysine mutation, this change in activity occurs despite the minor structural perturbation of the active site.[71,72] A subtle combination of electrostatic, structural, and/or dynamical effects has therefore long been thought to underlie the contribution of R55 to catalysis. Intriguingly, in a combined NMR and molecular dynamics study, Camilloni and coworkers proposed that R55 provides a stabilizing electric field that aligns with the electric dipole of the peptidyl carbonyl group as the pseudopeptide bond rotates through the cis/trans transition state.[65] This "electrostatic handle" was suggested to drive the rotation to occur only via positive peptide ω angles by reducing the activation energy when the carbonyl dipole is aligned with the R55 electric field, and increasing it when nonaligned.

From simulations of WT CypA, Camilloni et al. calculated the mean electric field component aligned perpendicular to the plane of the proline ring (chosen to be roughly normal to the peptide bond in either of the cis/trans ground states) to be stabilizing, but relatively small at −40 MV cm^{-1} to −50 MV cm^{-1}.[65] Solvatochromism experiments suggest this magnitude lies between the electric field strengths exerted on probe molecules by polar, hydrogen-bonding environments such as alcohols or water, and those exerted by apolar solvents such as hexane.[64] Therefore, the CypA active site is likely to be representative of typical intramolecular field strengths encountered in protein environments with both polar and nonpolar moieties. Potential energy surfaces calculated for a model

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

proline dipeptide by Camilloni et al. also suggested that a field of −50 MV cm^{-1} aligned with the cis/trans transition state would lower the isomerization activation energy by ∼30 kJ mol^{-1}, qualitatively consistent with experimentally observed reaction speedups over the uncatalyzed reaction.[65]

Nevertheless, the electrostatic handle hypothesis was proposed using CypA field strengths determined with ONIOM DFT calculations, the relatively high computational cost of which precludes the use of this approach to study electric fields in proteins more generally. However, the hypothesis does suggest that CypA provides a well-suited test case to evaluate the ability of the AMOEBA force field to re-create electric fields in protein active sites, which would provide a much faster route to evaluating ground-state electrostatics, energetics, and dynamics.

AMOEBA Field Magnitudes Projected Perpendicular to the Substrate Proline Ring Are Smaller than Those of Additive Force Fields. To evaluate the ability of AMOEBA to model typical intramolecular fields in CypA, we first calculated electric fields present in structural ensembles of the WT, cis-proline, CypA complex (PDB 1M9Y,[57] prepared and simulated as per the Methods section). For each simulation frame from the triplicate simulations performed with AMOEBA, the Cartesian x, y, and z components of environmental electric field were calculated at the C and N atoms of the substrate peptide bond. Each atom-centered field was projected along the vector perpendicular to the plane of the peptide proline ring, which reflects the magnitude of the field aligned with the carbonyl group of the peptide bond when rotated during the substrate isomerization. Projections were then linearly averaged to estimate the field magnitude at the peptide bond midpoint (Figure 1). After projection, intramolecular electric field magnitudes ranged from −40 MV cm^{-1} to +40 MV cm^{-1} (Figure 2A), even in these short simulations. However, the distributions overlapped substantially and the mean field strengths from each independent simulation spanned a much smaller range, from −6.9 MV cm^{-1} to −16.8 MV cm^{-1}.

Part of the variation in field strength originates from the projection of the Cartesian field components along a vector perpendicular to the proline ring at each time step. The direction of this vector is not identical for every MD snapshot, as torsional motions and out-of-plane bending subtly affect the conformation of the substrate proline ring. Thus, even if the CypA environment provided a constant electric field, a distribution of field strengths would be observed, because of the conformational fluctuation of the peptide. To interrogate the effects of substrate conformational fluctuation, for the first AMOEBA replicate, we visualized the Cartesian components of the vector perpendicular to the proline ring, and compared them with the Cartesian components of the calculated electric fields at the substrate peptide N and C atoms (Movies S1 and S2 in the Supporting Information). Even after structural superimposition of the frames, the electric field vectors were distributed across a wide range of orientations. However, the vectors perpendicular to the proline ring remained in a compact distribution of orientations, indicating that intramolecular fluctuations of the substrate proline were not the dominant source of the observed variance in electric field. Equilibrium active site structural fluctuations of the receptor, and distal residues in the substrate, therefore engender the majority of the observed variance.

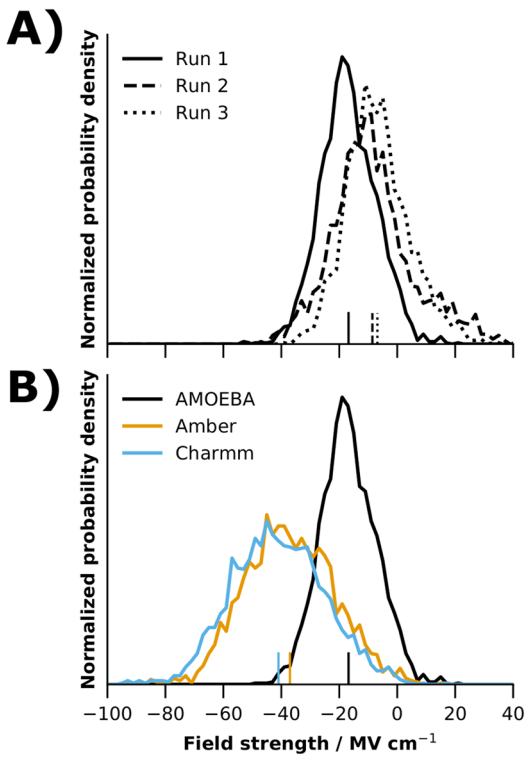


Figure 2. Distributions of environmental electric fields projected perpendicular to the proline ring, observed in 25 ns simulations of the WT, cis-proline CypA complex. (A) Field distributions observed at the isomerized peptide bond in three replicate AMOEBA simulations (solid, dashed, and dotted lines). (B) Field distributions observed for the structures of the AMOEBA Run 1 ensemble (black, identical simulation to Run 1 in panel (A)), but with fields recalculated using either the Amber (orange), and Charmm (blue) force fields. In both panels, ticks on the x-axis denote the mean of each distribution.

The magnitude of field fluctuations is affected by the sensitivity of the force field electrostatic model to small structural changes. Therefore, the AMOEBA field distributions were compared with those of the Amber ff14SB and Charmm C36m additive models, with a TIP3P water model in both cases. For each of the additive force fields, the electric fields at the substrate peptide C and N atoms were recalculated for each trajectory frame of the AMOEBA structural ensemble from the first replicate.

Fields calculated with the additive models were systematically larger (more negative) than those of AMOEBA and showed a broader distribution (Figure 2B). Distributions with Amber and Charmm force fields were remarkably similar, although slightly larger fields were calculated with the Charmm partial charges ($\bar{x}$ = −41.0 MV cm^{-1}), compared to Amber ($\bar{x}$ = −37.0 MV cm^{-1}). For the Charmm partial charges, increasing the real-space electrostatic cutoff from 8 Å to the commonly used 12 Å made a negligible difference (<0.02%) to the total electrostatic energies and calculated field strengths, indicating that the differences between each force field family were systematic, rather than caused by real-space cutoff choice. For an appropriately balanced Ewald sum, this is of course what would be expected. To investigate whether the force field used to create the structural ensemble biased the calculated fields, we generated an alternative structural ensemble of the same system with the Amber ff14SB force field. This ensemble showed analogous results, with a very broad distribution of

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

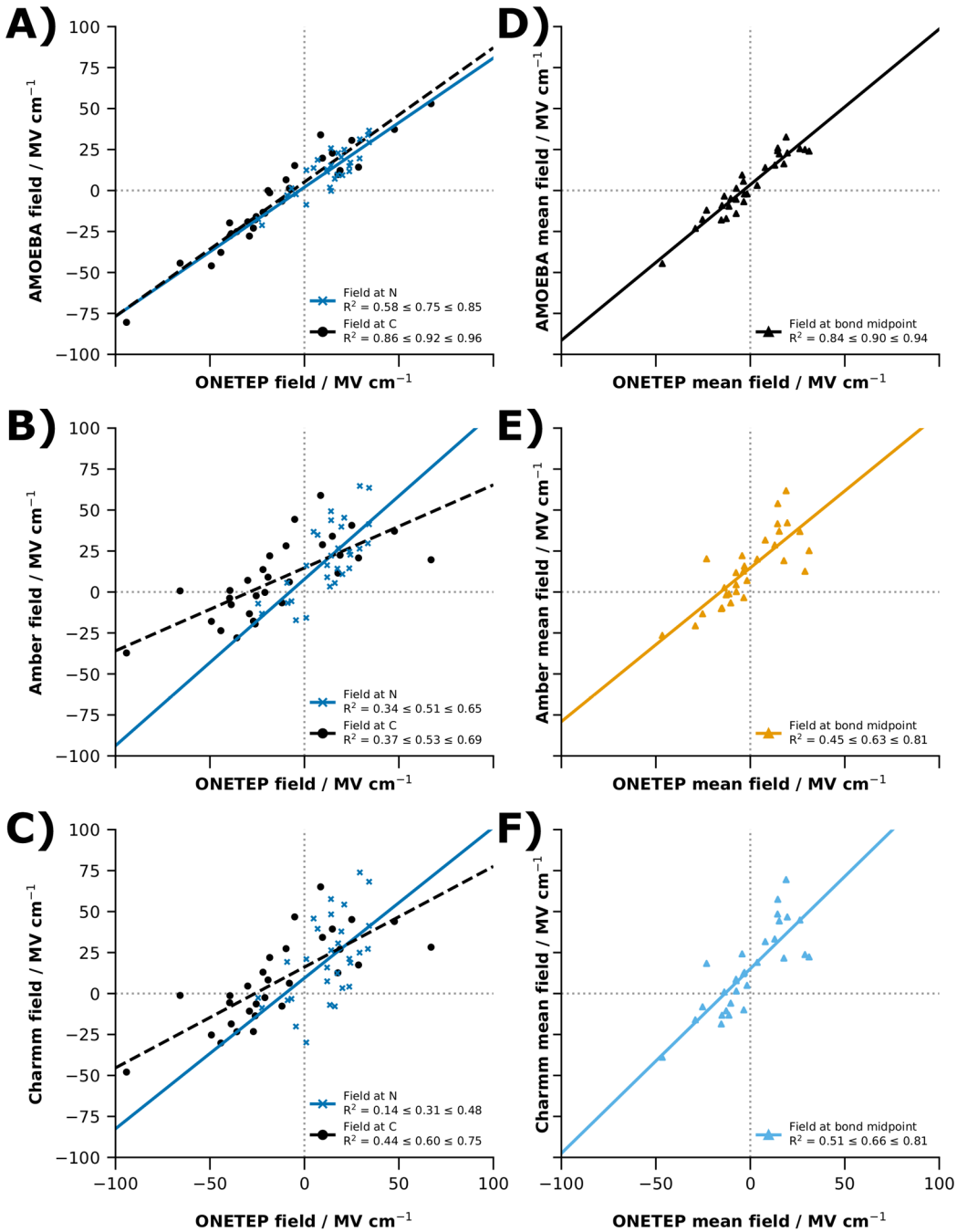


Figure 3. Comparison of ONETEP DFT electric fields inside the CypA WT cis-proline system with those of MM force fields. (A−C) Comparison of the Cartesian *x*, *y*, and *z* field components at the cis-proline N atom (dark blue crosses) and preceding C atom (black circles), for ONETEP and AMOEBA (panel (A)), Amber (panel (B)), and Charmm force fields (panel (C)). (D−F) Equivalent comparison of the Cartesian *x*, *y*, and *z* field components at the bond midpoint, taken as the mean of the N and C atom components, for ONETEP and AMOEBA (panel (D)), Amber (panel (E)), and Charmm force fields (panel (F)). In all cases, uncertainty is estimated from resampling the data with replacement (n = 2000) and is provided as the 95% confidence intervals in R^2.

Amber field strengths centered at −40.9 MV cm^{-1}, and a comparatively tight AMOEBA distribution centered at −18.2 MV cm^{-1} (see Figure S1 in the Supporting Information). Therefore, the differences observed between the fields calculated with additive and polarizable models were not simply caused by the structural ensemble generated with the AMOEBA force field. Instead, the broader distribution of field strengths observed with Amber and Charmm force fields supports the notion that an additive Coulombic electrostatics model is more sensitive to the instantaneous environmental conformations than the AMOEBA polarizable model.

DFT Fields at the CypA Active Site Are Consistent with AMOEBA. In our simulations of the WT cis-proline complex, the AMOEBA force field consistently showed an offset of ca. +15 MV cm^{-1} to +25 MV cm^{-1} to Amber or Charmm electric field strengths. Although this trend was

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

internally consistent across simulations and between force fields (see Figure S2 and Table S2 in the Supporting Information), the AMOEBA electric fields appeared to differ from those calculated by Camilloni and co-workers in CypA, which were estimated to be at least −40 MV cm^{-1} at the active site in both the cis-proline and trans-proline ground states.[65] However, our system and methodology also differed substantially from this previous study. Simulations here were performed with a different peptide substrate, applied no NMR restraints to the protein structural ensemble, and calculated fields at both peptide C and N atoms during the projection of the field vector along the peptide normal, rather than only at the N atom.[65] As such, we tested the accuracy of our MM-calculated field strengths by comparison to field strengths calculated with ONETEP DFT on the entire simulation system, which consisted of 19 018 atoms. A subset of 10 frames extracted at 2.5 ns intervals from the first replicate AMOEBA WT cis-proline simulation was used for comparison. Field vectors were compared directly, using their Cartesian x, y, and z components in the simulation frame, rather than any projection along a specific vector.

Correlations between the ONETEP DFT and AMOEBA fields were relatively high, with R^2 values of 0.75 and 0.92 for fields calculated at the substrate peptide N and C atoms, respectively (see Figure 3A). Correlations between Amber or Charmm fields and ONETEP were much smaller, with R^2 no higher than 0.60 (see Figures 3B and 3C). Moreover, regression lines for the additive force fields appeared to have a large positive intercept for fields calculated at the peptide N and C atoms individually. An identical trend was observed when comparing QM and MM results after averaging the N and C atom field components for each frame (see Figures 3D−F). Again, AMOEBA fields gave close agreement to those calculated with ONETEP (R^2 = 0.90), with a slope close to unity and without the substantial positive y-intercept and broad confidence intervals exhibited by Amber and Charmm fields (see Table 1).

Table 1. Regression Statistics for CypA WT cis Electric Fields Calculated with AMOEBA, Amber and Charmm Force Fields, and Compared to ONETEP DFT Calculations[a]

force field	slope	y-intercept
AMOEBA	$0.86 \leq 0.95 \leq 1.05$	$1.7 \leq 3.5 \leq 5.3$
Amber	$0.70 \leq 0.94 \leq 1.23$	$11.0 \leq 14.8 \leq 18.9$
Charmm	$0.87 \leq 1.13 \leq 1.45$	$10.6 \leq 15.1 \leq 20.0$

[a]Fields calculated as the mean of those at the C and N peptide atoms (see Figures 3D−F for the regression lines). Uncertainty estimated as the 95% confidence interval in regression statistics calculated by resampling the data with replacement (n = 2000).

The large, positive y-intercepts observed when comparing the additive model field components to DFT suggest that Amber and Charmm may systematically estimate more positive electric fields than AMOEBA (see Table 1). This may appear counterintuitive, given that Amber and Charmm field magnitudes, when projected along the vector perpendicular to the proline ring, were substantially more negative than those of AMOEBA (Figure 2B). However, by projecting fields along the vector perpendicular to the proline ring, we only compare the catalytically relevant dimension of field strength, and do not compare the full differences in field magnitude and direction between the force fields. The comparison of individual Cartesian components with DFT more fairly characterizes both properties. In addition, the relatively limited data used in each comparison (30 points, one for each Cartesian component of the field in 10 separate trajectory frames) could potentially result in dataset bias, and we have not attempted to ensure structural diversity between the analyzed frames, preferring instead to use the entire trajectory length to extract time-separated frames at regularly spaced intervals. The confidence intervals in the R^2 values (Figure 3, estimated by resampling with replacement), almost never overlap for the ONETEP/AMOEBA and ONETEP/additive models, suggesting that the improved correlation observed with AMOEBA is likely to be statistically significant if the entire population were compared. To broaden the structural diversity of the dataset while restricting the overall size to a computationally tractable number of full-DFT electric field calculations, we similarly analyzed alternative CypA-peptide substrate complexes.

AMOEBA Reproduction of DFT Fields Is Insensitive to Substrate Isomer and CypA Mutation. Having established that AMOEBA fields were consistent with those of DFT in the WT cis-proline system, we performed an identical analysis of the WT CypA bound to a trans-proline substrate (PDB 1M9C,[57] prepared as described in the Methods section). After simulation with the AMOEBA force field, the AMOEBA, Amber, and Charmm MM fields at the bond midpoint, projected perpendicular to the proline ring, were compared across all trajectory snapshots (see Figure S2 in the Supporting Information). Taking the first replicate as representative, the ONETEP and MM x, y, and z field components were then compared for a subset of 10 equally spaced frames extracted from the full 25 ns simulation (see Figures 4A and 4B).

Again, AMOEBA showed a narrow distribution of projected field strengths, ∼15−20 MV cm^{-1} more positive than Amber or Charmm fields calculated from the same structures (Figure 4A). In terms of the absolute x, y, and z field components, AMOEBA also exhibited strong correlation with the DFT fields, achieving an R^2 of 0.95 across the 10 frames extracted from the WT trans-proline trajectory. In contrast, Amber and Charmm fields for the same structural ensemble showed much lower correlations (R^2 = 0.82 and 0.73, respectively), and, again, 95% confidence intervals in R^2 for the AMOEBA and additive model comparisons did not overlap, despite the limited dataset size (Figure 4B). Next, we performed an identical comparison of electric fields in two mutant CypA systems to add further structural diversity to the dataset and evaluate the transferability of the observed trends to the functionally relevant R55A mutation. We evaluated R55A systems with either cis-proline or trans-proline substrates in order to be equivalent to the WT systems, and fields from three 25-ns AMOEBA simulation of each mutant system were calculated and compared across the AMOEBA, Amber ff14SB, and Charmm C36m force fields (see Figure S2 in the Supporting Information). For each simulated system, the first trajectory was taken to be representative, and a subset of 10 equally spaced frames were again used to correlate Cartesian field components with ONETEP.

Upon mutation of Arg55 in the CypA protein to alanine, there appeared to be a large reduction in the projected field magnitude experienced at the substrate proline bond (see Figures 4C and 4E). The absolute magnitude of this reduction differed in the AMOEBA force field, compared to the two

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

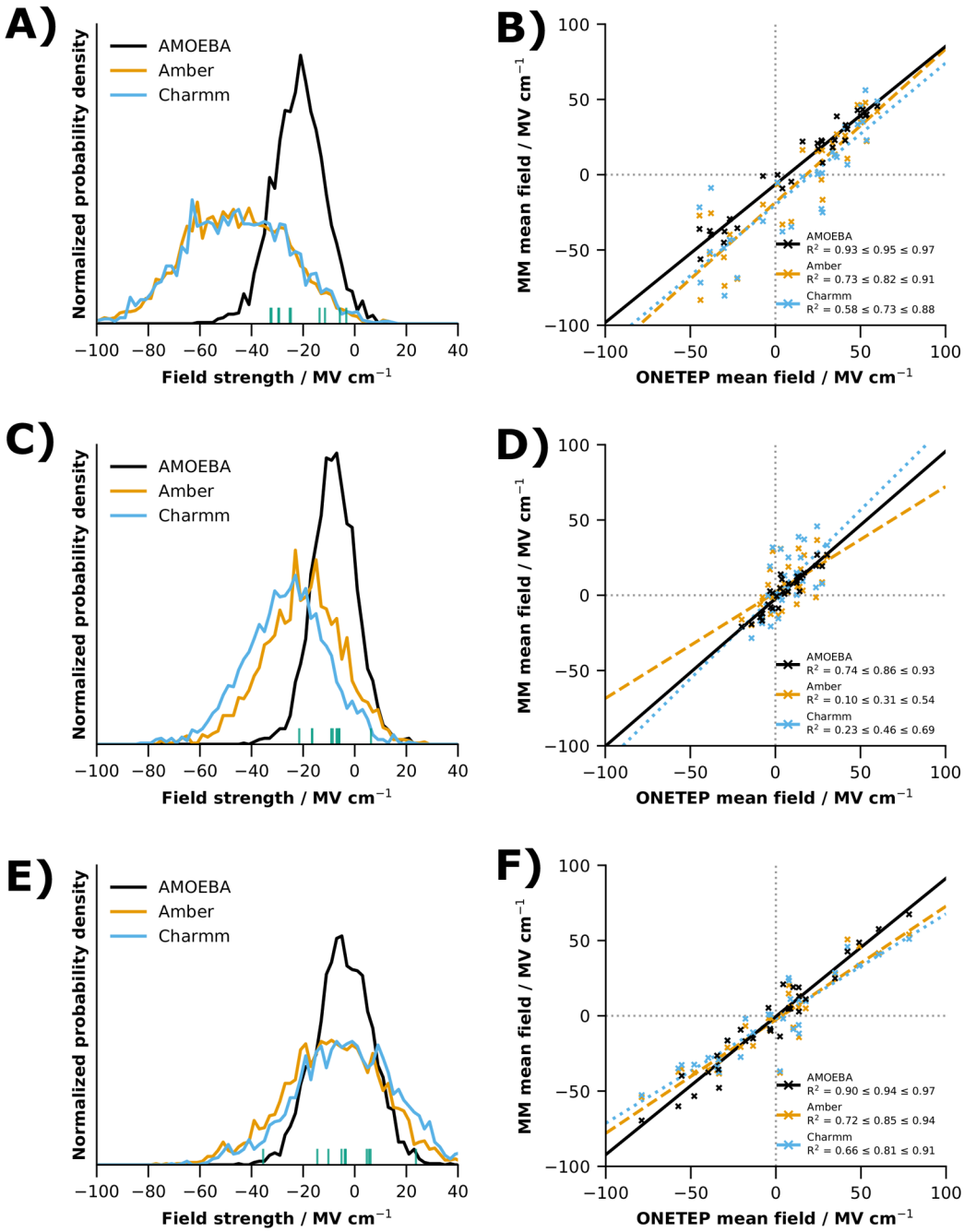


Figure 4. Comparison of MM and DFT fields calculated for additional CypA systems. (A, C, E) Distributions of electric field projected perpendicular to the proline ring from 25 ns AMOEBA simulations of the WT trans-proline (panel (A)), R55A cis-proline (panel (C)), or R55A trans-proline system (panel (E)). Fields were evaluated at the substrate peptide bond midpoint using the AMOEBA (black), Amber (orange), or Charmm (blue) force fields. Field magnitudes of the frames extracted for ONETEP analysis are denoted by ticks (teal) on the *x*-axis. (B, D, F) Comparison of ONETEP DFT *x*, *y*, and *z* electric field components at the peptide bond center with those of AMOEBA (black, solid lines), Amber (orange, dashed lines), and Charmm (blue, dotted lines) for a subset of frames from the 25 ns AMOEBA trajectory of the WT trans-proline (panel (B)), R55A cis-proline (panel (D)), or R55A trans-proline system (panel (F)). Uncertainty is estimated from resampling the data with replacement (n = 2000) and is provided as the 95% confidence intervals in R^2.

additive force fields, but was similar in relative terms—fields were reduced by ∼40%−50% upon CypA mutation with the cis-proline substrate, and by ∼80%−90% upon mutation with the trans-proline substrate. Notably, the fields calculated with the additive models always showed a broader distribution with higher variance (see Table S2 in the Supporting Information), again suggesting that the additive potentials were more sensitive to small structural changes in instantaneous conformation than the polarizable model. This led to greater populations of frames with comparatively high and low field strengths, when calculated with the additive models.

Comparison of the MM field components to their DFT equivalents (Figures 4D and 4F) highlighted trends identical to those observed for the WT systems—AMOEBA consistently showed greater correlation and lower systematic errors to DFT fields. Across the four CypA systems individually, no clear

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

MOST WIEDZY Downloaded from mostwiedzy.pl

Table 2. Regression Statistics for CypA WT trans and R55A Electric Fields, Compared to ONETEP DFT Calculations[a]

	WT trans system		R55A cis system		R55A trans system	
force field	slope	*y*-intercept	slope	*y*-intercept	slope	*y*-intercept
AMOEBA	$0.86 \leq 0.92 \leq 0.98$	$-9.0 \leq -6.5 \leq -3.8$	$0.88 \leq 0.98 \leq 1.08$	$-4.0 \leq -2.4 \leq -0.5$	$0.86 \leq 0.92 \leq 0.98$	$-2.8 \leq -0.4 \leq 2.1$
Amber	$0.84 \leq 1.02 \leq 1.21$	$-25.6 \leq -18.4 \leq -12.3$	$0.41 \leq 0.70 \leq 0.99$	$-2.0 \leq 1.8 \leq 6.3$	$0.68 \leq 0.75 \leq 0.84$	$-6.5 \leq -2.6 \leq 0.8$
Charmm	$0.73 \leq 0.94 \leq 1.16$	$-27.9 \leq -20.0 \leq -12.6$	$0.75 \leq 1.12 \leq 1.50$	$-3.7 \leq 0.6 \leq 5.6$	$0.63 \leq 0.70 \leq 0.77$	$-5.8 \leq -1.8 \leq 1.9$

[a]Fields calculated as the mean of those at the C and N peptide atoms (see Figures 4B, 4D, and 4F for the regression lines). Uncertainty estimated as the 95% confidence interval in regression statistics calculated by resampling the data with replacement (n = 2000).

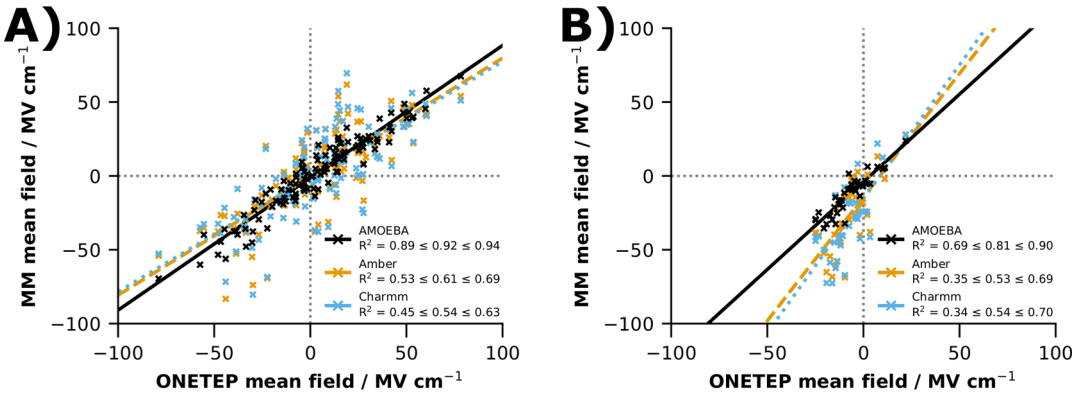


Figure 5. Comparison of DFT and MM calculated fields across all four simulated CypA systems. (A) Comparison of x, y, and z field components at the peptide bond midpoint. AMOEBA (black, solid line, linear regression $y = 0.90x - 1.3$) shows far greater correlation to ONETEP DFT fields than either Amber ff14SB (orange, dashed line) or Charmm C36m (blue, dotted line), and without substantially overestimating or underestimating field strengths at extreme values. (B) Comparison of field magnitudes projected along the vector perpendicular to the substrate proline ring. The color scheme is identical to that observed in panel (A).

Table 3. Regression Statistics for All CypA Electric Fields Calculated with AMOEBA, Amber, and Charmm Force Fields and Compared to ONETEP DFT Calculations[a]

	x, *y*, *z*-components		Perpendicular to proline	
force field	slope	*y*-intercept	slope	*y*-intercept
AMOEBA	$0.86 \leq 0.90 \leq 0.94$	$-2.4 \leq -1.3 \leq -0.1$	$1.07 \leq 1.19 \leq 1.32$	$-5.5 \leq -4.0 \leq -2.5$
Amber	$0.72 \leq 0.80 \leq 0.90$	$-3.2 \leq -0.5 \leq 2.2$	$1.30 \leq 1.68 \leq 2.08$	$-19.2 \leq -14.4 \leq -9.8$
Charmm	$0.70 \leq 0.78 \leq 0.88$	$-3.5 \leq -0.5 \leq 2.6$	$1.40 \leq 1.81 \leq 2.22$	$-20.9 \leq -15.5 \leq -11.0$

[a]Fields calculated as the mean of those at the C and N peptide atoms (see Figures 5A and 5B for the regression lines). Uncertainty estimated as the 95% confidence interval in regression statistics calculated by resampling the data with replacement (n = 2000).

trend in terms of overestimation or underestimation was observed with particular force fields. Slopes, intercepts, and R^2 values varied (see Tables 1 and 2), but AMOEBA consistently demonstrated high correlation to DFT fields (R^2 = 0.86−0.95). Combining all four CypA systems together, however, highlighted a slight tendency for the additive force fields to underestimate the magnitude of the x, y, and z electric field components, compared to DFT, in both the positive and negative directions (see Figure 5A), that is, negative field strengths were calculated to be too positive and positive field strengths were calculated to be too negative. In addition, the broad distribution of the Amber and Charmm fields remained evident in the comparison of the individual Cartesian field components with DFT, consistent with earlier observations of the field strengths projected perpendicular to the proline ring.

The observation that additive force fields slightly underestimated the magnitude of the electric field components compared to DFT again appeared to contradict the fact that, when fields were projected perpendicular to the proline ring, field strengths were almost always larger (more negative) with Amber and Charmm than with AMOEBA (see Figures 2B and Figures 4A, 4C, 4E, as well as Figure S2 in the Supporting Information). When the projected electric fields were compared to DFT fields; however, the additive force fields demonstrated a clear systematic error, with y-intercepts of approximately −15 MV cm^{-1}, and slopes much greater than 1 (see Figure 5B and Table 3). In contrast, projected AMOEBA electric fields showed a small systematic offset of −4 MV cm^{-1} from those of DFT. Hence, the majority of the observed difference in field magnitudes between AMOEBA and additive force fields is caused by projecting the total electric field along a catalytically relevant dimension. Therfore, this comparison may not assess the general accuracy of an MM force field; instead, our conclusions are better drawn taking into account both magnitude and directionality of the total field, by comparing individual field components. With this in mind, Cartesian field components calculated with AMOEBA correlated well with those of DFT (R^2 = 0.92) and with slope close to unity and intercept close to zero (see Table 3).

■ DISCUSSION

Electric Fields in Polarizable and Additive Models. The development of electrostatic parameters for additive biomolecular force fields is problematic, because of the

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

fundamentally different environments in which the atomic partial charges are first derived (generally by gas-phase QM calculations upon individual amino acids or fragments) and then applied (condensed-phase simulations of large macromolecular complexes, including water, ions, and other solutes). The Amber and Charmm force field families, like many others, have compensated for the effects of transferring parameters to the condensed phase by fitting partial charges that are intrinsically overpolarized.[10,11,76,77] This approximation is ultimately inadequate, and recognition of this fact has contributed to the drive to develop accurate explicitly polarizable models.[8,24,31,33]

Nevertheless, parametrization processes for additive models are generally intended to re-create intermolecular energetics (determined by the electrostatic potential at a given site) and forces (determined by the electric field) as accurately as possible. However, our results for CypA suggest that the intrinsic overpolarization of atomic partial charges is insufficient to re-create the effects of induction upon intramolecular electric fields, and that this issue may be generalizable to multiple force field families. The polarization model in AMOEBA is both a more accurate electrostatic model overall, with reduced random error and improved correlation to DFT fields, and does not show a sizable systematic error in electric fields, with a near-unity slope, in comparison to DFT (Figure 5).

The implications of poorly estimating intramolecular electric fields are likely to vary by application. Properties reliant on a balance of intermolecular interactions and conformational propensities, such as solvation or binding free energies, exploration of structural and dynamical ensembles, or kinetics of transitions, are likely to be re-created successfully with well-parametrized additive models through compensatory effects of the remaining potential function terms. Both the Amber ff14SB and Charmm C36m force fields have been well-validated for their ability to accurately re-create biomolecular structure and dynamics.[2,3] AMOEBA, or other equally accurate polarizable models, may provide significant improvement in accuracy when the electric field is the crucial property of interest.[47,48,64] Notably, our results suggest that this is true, even for electric fields encountered in "normal" protein environments—the maximum electric field magnitudes encountered in CypA (approximately ±75 MV cm^{-1} from Figure 5) are well below those measured in more "extreme" environments, such as the active site of ketosteroid isomerase.[78]

Beyond electrostatics calculations, intramolecular electric fields, and how proteins conformationally respond to changes in electric fields, are also increasingly implicated in biophysical processes, from enzyme reaction mechanisms to voltage sensing in ion channels.[78−82] Nevertheless, the increased computational cost of polarizable models is still likely to limit the scope of dynamics-based applications to large biomolecular systems. Advances in both software and hardware will therefore be crucial to further extend the range of applications for polarizable simulations, although our simulations with CypA demonstrate that nanosecond-length dynamics are currently routinely accessible via the use of consumer-grade GPU hardware, at relatively modest computational cost (for CypA, an approximately 40-fold increase for AMOEBA with Tinker-OpenMM over equivalent additive simulations). Clearly, the excellent agreement of AMOEBA with DFT observed here does not negate the potential for inaccuracy in other systems or even alternative states along the CypA reaction pathway, but it does highlight the potential for polarizable MM models to achieve quantitative accuracy in biological environments where the electric field is crucial for structure or function.

Implications for CypA Mechanism. The "electrostatic handle" mechanism of CypA proposes that catalytic activity is driven by the alignment of a large electric field (provided by the R55 residue in WT CypA) with the carbonyl dipole of the residue preceding the substrate proline in the transition state. Our results with both AMOEBA and additive force fields are consistent with a significant contribution of R55 to the overall electric field in both CypA ground states. However, the magnitude of this contribution (taken as the difference between mean fields in the WT and R55A simulations) appears to be smaller than that proposed by Camilloni and coworkers, with a maximum of −8.7 MV cm^{-1} for the cis-proline or −23.5 MV cm^{-1} for the trans-proline system across all AMOEBA simulations (Table S2 in the Supporting Information), rather than the ca. −30 MV cm^{-1} estimated previously.[65]

The comparisons of QM and MM electric fields performed here were not designed specifically to probe the CypA mechanism, and there may be multiple reasons for the observed discrepancies in field strengths. First, the 25 ns MD simulations performed with AMOEBA, without replica-averaged NMR restraints, are unlikely to fully explore the equilibrium ground-state dynamics of CypA. Field distributions varied between repeat simulations of all four WT and R55A CypA states (see Figure S2 and Table S2), suggesting a lack of convergence of the absolute field strengths. However, in all but one simulation, AMOEBA fields, projected perpendicular to the proline ring, were relatively smaller than those of either Amber or Charmm. On average, AMOEBA fields also showed a decrease in field strength in the R55A mutant simulations, generally supporting the conclusion that R55 makes significant contributions to the active site field in the ground states. In addition, the control calculation of AMOEBA field strengths taken from a structural ensemble generated with Amber suggests that the use of AMOEBA for MD does not fundamentally change the distribution of configurations explored, in the context of electric fields.

It is also possible that the contribution of R55 to electric fields in CypA was previously overestimated. First, field strengths were previously calculated at the proline N atom only, rather than averaged across both peptide bond atoms, and we consistently see a larger electric field at the proline N than the preceding C atom in all force fields tested in our calculations (see Table S2). Second, although fields were previously calculated using DFT ONIOM calculations at the B3LYP/6-31G** level, only a subset of the CypA protein (250 atoms) was included in the QM region, and without long-range electrostatic effects of protein and solvent estimated by electrostatic embedding. Moreover, electric fields were calculated taking into account the contribution of only the CypA receptor residues, not using the environment fields approach of Fried and Wang that included back-polarization of the receptor by the substrate.[64] It is difficult to predict the modulation of electric field caused by back-polarization at a specific site along a specific vector, but the effect should strictly be included as far as practicable. Finally, the previous WT CypA simulations were performed incorporating experimentally derived restraints, and should therefore be a reliable representation of the WT structural ensembles. However, trajectory frames were geometry optimized using the Amber

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

ff99SB-ILDN force field prior to the DFT calculations, and structures of the R55A mutant were generated directly from the WT structural ensemble rather than a separate set of simulations. Therefore, the previously reported effects of the R55A mutation are purely electrostatic, and they do not consider the potential for any structural reconfiguration upon mutation. Both approximations are conceptually likely to overstabilize and hence overestimate the electrostatic effects within the CypA active site.

Therefore, our results are broadly consistent with the electrostatic mechanism proposed by Camilloni and co-workers, but may not have captured all catalytically relevant CypA motions too, because the magnitude of the overall field contribution is small. The limitations of the short simulations performed here, notwithstanding the fact that the AMOEBA 2013 protein force field has not been as widely tested and validated as either of the two additive models investigated, mean that we cannot draw unambiguous conclusions regarding the CypA mechanism. Recent detailed investigations of CypA dynamics have required both innovative biophysical techniques, and lengthy advanced sampling methodologies for simulations, and the use of polarizable force fields does not negate this necessity for adequate sampling.[74,83−85] However, the accuracy of AMOEBA fields, compared to DFT, suggests that polarizable simulations provide a promising avenue to better explore the electrostatic contributions to mechanism across members of the cyclophilin family, in a computationally efficient way.

■ CONCLUSION

AMOEBA's use of higher-order multipoles and an induced dipole representation of induction has been developed to accurately reproduce the electrostatic potential around small molecules and amino acid fragments. The results presented here demonstrate that this accurate representation of electrostatics transfers directly to a typical condensed phase biomolecular environment. The ability of MM force fields to reproduce intramolecular electric fields is an experimentally relevant property of interest, and increasingly experimentally measurable via, for example, Stark effect spectroscopy or redox potential measurements.[49,86] The use of linear-scaling ONETEP DFT calculations allowed us to evaluate electric fields inside fully atomistic, fully solvated representations of CypA under periodic boundary conditions and in the presence of explicit ions, without simplified or truncated structural models. In the absence of experimental measurements of electric fields in CypA, we compared electric fields calculated with AMOEBA, Amber, and Charmm biomolecular force fields with these extensive QM calculations. AMOEBA showed significantly better agreement with DFT-calculated fields across the full range of systems, structures, and field magnitudes tested, with nonoverlapping 95% confidence intervals in R^2 in the majority of cases, despite relatively small sample sizes.

The high accuracy of AMOEBA electrostatics in condensed-phase, intramolecular environments suggests that it (and potentially other polarizable force fields) may be ideal for applications where electrostatic accuracy is key. Electric field strengths in the CypA simulations are well within the range of those typically found in biomolecular environments, which therefore suggests a wide range of potential applications. Promisingly, however, studies of enzyme activity and mechanism with polarizable force fields have already begun to allow quantitative insights into electrostatically driven biological mechanisms that additive force fields appear unable to provide.[47,48] The results we present from simulations of CypA cannot unambiguously determine a reaction mechanism, nor were they intended to. They are supportive of, but not limited to, an electrostatic role for R55, consistent with what is known for CypA. However, the fact that AMOEBA can determine electric fields in CypA with almost DFT-level accuracy indicates that, with suitably equilibrated structural ensembles, provided for example by high-performance software implementations,[54,55] AMOEBA could be used to accurately probe the ground-state dynamics and electric fields of the wider cyclophilin family within single simulations, and at substantially lower computational cost than is currently required for QM/MM or QM post-processing approaches.

■ ASSOCIATED CONTENT

Supporting Information

Underlying data and code used to support the results of this study, as detailed in the Data Availability section, is made freely available as DOI: 10.5281/zenodo.3678278. The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00217.

Tables describing all simulations performed and the mean electric fields projected perpendicular to the substrate proline ring, for all simulations generated with the AMOEBA force field. Figures depicting the distributions of electric field strengths in all simulations performed with either the Amber ff14SB or AMOEBA force fields (PDF)

Movie depicting the electric fields and vectors perpendicular to the proline ring at the proline N atom in the first WT cis-proline CypA simulation (MPG)

Movie depicting the electric fields and vectors perpendicular to the proline ring at the alanine C atom, in the first WT cis-proline CypA simulation (MPG)

■ AUTHOR INFORMATION

Corresponding Authors

Jonathan W. Essex − *School of Chemistry, University of Southampton, Southampton SO17 1BJ, United Kingdom;* orcid.org/0000-0003-2639-2746; Email: J.W.Essex@soton.ac.uk

Richard T. Bradshaw − *School of Chemistry, University of Southampton, Southampton SO17 1BJ, United Kingdom;* orcid.org/0000-0002-8652-4301; Email: R.T.Bradshaw@soton.ac.uk

Authors

Jacek Dziedzic − *School of Chemistry, University of Southampton, Southampton SO17 1BJ, United Kingdom; Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, 80-233 Gdańsk, Poland;* orcid.org/0000-0003-4786-372X

Chris-Kriton Skylaris − *School of Chemistry, University of Southampton, Southampton SO17 1BJ, United Kingdom;* orcid.org/0000-0003-0258-3433

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.0c00217

MOST WIEDZY Downloaded from mostwiedzy.pl

Author Contributions

[∇]These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Calculations in this work made use of the Iridis4 supercomputer at the University of Southampton, and the TASK Academic Computer Centre in Gdansk, Poland. We thank all members of the EPSRC/NSF-funded APES consortium, particularly Jay Ponder, Teresa Head-Gordon, Martin Head-Gordon, David Case, Jason Swails, Mark Tuckerman, Paul Nerenberg, Lorna Smith, and Ilian Todorov, for helpful discussions throughout. We gratefully acknowledge support from EPSRC (Grant No. EP/K039156/1).

■ REFERENCES

(1) Huggins, D. J.; Biggin, P. C.; Dämgen, M. A.; Essex, J. W.; Harris, S. A.; Henchman, R. H.; Khalid, S.; Kuzmanic, A.; Laughton, C. A.; Michel, J.; Mulholland, A. J.; Rosta, E.; Sansom, M. S. P.; van der Kamp, M. W. Biomolecular Simulations: From Dynamics and Mechanisms to Computational Assays of Biological Activity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2019**, *9* (3), e1393.

(2) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.

(3) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71−73.

(4) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11* (7), 3499−3509.

(5) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65* (3), 712−725.

(6) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257−3273.

(7) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12* (1), 281−296.

(8) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. *Adv. Protein Chem.* **2003**, *66*, 27−85.

(9) Riniker, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *J. Chem. Inf. Model.* **2018**, *58* (3), 565−578.

(10) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.

(11) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586−3616.

(12) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225−11236.

(13) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(14) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25* (13), 1656−1676.

(15) Amaro, R. E.; Mulholland, A. J. Multiscale Methods in Drug Design Bridge Chemical and Biological Complexity in the Search for Cures. *Nat. Rev. Chem.* **2018**, *2* (4), 0148.

(16) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129−1143.

(17) Bottaro, S.; Lindorff-Larsen, K. Biophysical Experiments and Biomolecular Simulations: A Perfect Match? *Science* **2018**, *361* (6400), 355−360.

(18) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78* (8), 1950−1958.

(19) Li, P.; Merz, K. M. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**, *10* (1), 289−297.

(20) Yoo, J.; Aksimentiev, A. New Tricks for Old Dogs: Improving the Accuracy of Biomolecular Force Fields by Pair-Specific Corrections to Non-Bonded Interactions. *Phys. Chem. Chem. Phys.* **2018**, *20* (13), 8432−8449.

(21) Cerutti, D. S.; Swope, W. C.; Rice, J. E.; Case, D. A. Ff14ipq: A Self-Consistent Force Field for Condensed-Phase Simulations of Proteins. *J. Chem. Theory Comput.* **2014**, *10* (10), 4515−4534.

(22) Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the Road Less Traveled: AMBER Ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *J. Chem. Theory Comput.* **2016**, *12* (8), 3926−3947.

(23) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119* (16), 5113−5123.

(24) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. Classical Electrostatics for Biomolecular Simulations. *Chem. Rev.* **2014**, *114* (1), 779−814.

(25) Bereau, T.; Kramer, C.; Monnard, F. W.; Nogueira, E. S.; Ward, T. R.; Meuwly, M. Scoring Multipole Electrostatics in Condensed-Phase Atomistic Simulations. *J. Phys. Chem. B* **2013**, *117* (18), 5460−5471.

(26) Kramer, C.; Gedeck, P.; Meuwly, M. Atomic Multipoles: Electrostatic Potential Fit, Local Reference Axis Systems, and Conformational Dependence. *J. Comput. Chem.* **2012**, *33* (20), 1673−1688.

(27) Rocklin, G. J.; Boyce, S. E.; Fischer, M.; Fish, I.; Mobley, D. L.; Shoichet, B. K.; Dill, K. A. Blind Prediction of Charged Ligand Binding Affinities in a Model Binding Site. *J. Mol. Biol.* **2013**, *425* (22), 4569−4583.

(28) Caleman, C.; Van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; Van Der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8* (1), 61−74.

(29) Fennell, C. J.; Wymer, K. L.; Mobley, D. L. A Fixed-Charge Model for Alcohol Polarization in the Condensed Phase, and Its Role in Small Molecule Hydration. *J. Phys. Chem. B* **2014**, *118* (24), 6438−6446.

(30) Riquelme, M.; Lara, A.; Mobley, D. L.; Verstraelen, T.; Matamala, A. R.; Vöhringer-Martinez, E. Hydration Free Energies in the FreeSolv Database Calculated with Polarized Iterative Hirshfeld Charges. *J. Chem. Inf. Model.* **2018**, *58* (9), 1779−1797.

(31) Albaugh, A.; Boateng, H. A.; Bradshaw, R. T.; Demerdash, O. N.; Dziedzic, J.; Mao, Y.; Margul, D. T.; Swails, J.; Zeng, Q.; Case, D.

MOST WIEDZY Downloaded from mostwiedzy.pl

A.; Eastman, P.; Wang, L.-P.; Essex, J. W.; Head-Gordon, M.; Pande, V. S.; Ponder, J. W.; Shao, Y.; Skylaris, C.-K.; Todorov, I. T.; Tuckerman, M. E.; Head-Gordon, T. Advanced Potential Energy Surfaces for Molecular Simulation. *J. Phys. Chem. B* **2016**, *120* (37), 9811−9832.

(32) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annu. Rev. Biophys.* **2019**, *48* (1), 371−394.

(33) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116* (9), 4983−5013.

(34) Ponder, J. W.; Wu, C. J.; Ren, P. Y.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549−2564.

(35) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; Mackerell, A. D. Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9* (12), 5430−5449.

(36) Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *J. Phys. Chem. Lett.* **2014**, *5* (18), 3144−3150.

(37) Mu, X.; Wang, Q.; Wang, L.-P.; Fried, S. D.; Piquemal, J.-P.; Dalby, K. N.; Ren, P. Modeling Organochlorine Compounds and the σ-Hole Effect Using a Polarizable Multipole Force Field. *J. Phys. Chem. B* **2014**, *118* (24), 6456−6465.

(38) Ngo, V.; da Silva, M. C.; Kubillus, M.; Li, H.; Roux, B.; Elstner, M.; Cui, Q.; Salahub, D. R.; Noskov, S. Y. Quantum Effects in Cation Interactions with First and Second Coordination Shell Ligands in Metalloproteins. *J. Chem. Theory Comput.* **2015**, *11* (10), 4992−5001.

(39) Jing, Z.; Liu, C.; Qi, R.; Ren, P. Many-Body Effect Determines the Selectivity for Ca2+ and Mg2+ in Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (32), E7495−E7501.

(40) Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. TINKTEP: A Fully Self-Consistent, Mutually Polarizable QM/MM Approach Based on the AMOEBA Force Field. *J. Chem. Phys.* **2016**, *145* (12), 124106.

(41) Mao, Y.; Shao, Y.; Dziedzic, J.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Performance of the AMOEBA Water Model in the Vicinity of QM Solutes: A Diagnosis Using Energy Decomposition Analysis. *J. Chem. Theory Comput.* **2017**, *13* (5), 1963−1979.

(42) Dziedzic, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. Mutually Polarizable QM/MM Model with in Situ Optimized Localized Basis Functions. *J. Chem. Phys.* **2019**, *150* (7), 074103.

(43) Bradshaw, R. T.; Essex, J. W. Evaluating Parameterization Protocols for Hydration Free Energy Calculations with the AMOEBA Polarizable Force Field. *J. Chem. Theory Comput.* **2016**, *12* (8), 3871−3883.

(44) Mohamed, N. A.; Bradshaw, R. T.; Essex, J. W. Evaluation of Solvation Free Energies for Small Molecules with the AMOEBA Polarizable Force Field. *J. Comput. Chem.* **2016**, *37* (32), 2749−2758.

(45) Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, *86* (1), 387−415.

(46) Vaissier Welborn, V.; Head-Gordon, T. Computational Design of Synthetic Enzymes. *Chem. Rev.* **2019**, *119* (11), 6613−6630.

(47) Vaissier Welborn, V.; Ruiz Pestana, L.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nat. Catal.* **2018**, *1* (9), 649−655.

(48) Vaissier, V.; Sharma, S. C.; Schaettle, K.; Zhang, T.; Head-Gordon, T. Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase. *ACS Catal.* **2018**, *8* (1), 219−227.

(49) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.* **2015**, *48* (4), 998−1006.

(50) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. Introducing ONETEP : Linear-Scaling Density Functional Simulations on Parallel Computers. *J. Chem. Phys.* **2005**, *122* (8), 084119.

(51) Skylaris, C.-K.; Mostofi, A. A.; Haynes, P. D.; Diéguez, O.; Payne, M. C. Nonorthogonal Generalized Wannier Function Pseudopotential Plane-Wave Method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2002**, *66* (3), 035119.

(52) Wilkinson, K. A.; Hine, N. D. M.; Skylaris, C.-K. Hybrid MPI-OpenMP Parallelism in the ONETEP Linear-Scaling Electronic Structure Code: Application to the Delamination of Cellulose Nanofibrils. *J. Chem. Theory Comput.* **2014**, *10* (11), 4782−4794.

(53) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9* (9), 4046−4063.

(54) Lagardère, L.; Jolly, L.-H.; Lipparini, F.; Aviat, F.; Stamm, B.; Jing, Z. F.; Harger, M.; Torabifard, H.; Cisneros, G. A.; Schnieders, M. J.; Gresh, N.; Maday, Y.; Ren, P. Y.; Ponder, J. W.; Piquemal, J.-P. Tinker-HP: A Massively Parallel Molecular Dynamics Package for Multiscale Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields. *Chem. Sci.* **2018**, *9* (4), 956−972.

(55) Harger, M.; Li, D.; Wang, Z.; Dalby, K.; Lagardère, L.; Piquemal, J.-P.; Ponder, J.; Ren, P. Tinker-OpenMM: Absolute and Relative Alchemical Free Energies Using AMOEBA on GPUs. *J. Comput. Chem.* **2017**, *38* (23), 2047−2055.

(56) Huang, J.; Lemkul, J. A.; Eastman, P. K.; MacKerell, A. D. Molecular Dynamics Simulations Using the Drude Polarizable Force Field on GPUs with OpenMM: Implementation, Validation, and Benchmarks. *J. Comput. Chem.* **2018**, *39* (21), 1682−1689.

(57) Howard, B. R.; Vajdos, F. F.; Li, S.; Sundquist, W. I.; Hill, C. P. Structural Insights into the Catalytic Mechanism of Cyclophilin A. *Nat. Struct. Mol. Biol.* **2003**, *10* (6), 475−481.

(58) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, CA, 2014.

(59) Ren, P. Y.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933−5947.

(60) Ponder, J. TINKER: Software Tools for Molecular Design. *TINKER: Software Tools for Molecular Design*; Washington University School of Medicine: St. Louis, MO, 2001.

(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(62) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112* (30), 9020−9041.

(63) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(64) Fried, S. D.; Wang, L.-P.; Boxer, S. G.; Ren, P.; Pande, V. S. Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013**, *117* (50), 16236−16248.

(65) Camilloni, C.; Sahakyan, A. B.; Holliday, M. J.; Isern, N. G.; Zhang, F.; Eisenmesser, E. Z.; Vendruscolo, M. Cyclophilin A Catalyzes Proline Isomerization by an Electrostatic Handle Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (28), 10203−10208.

(66) Hill, Q.; Skylaris, C.-K. Including Dispersion Interactions in the ONETEP Program for Linear-Scaling Density Functional Theory Calculations. *Proc. R. Soc. London, Ser. A* **2009**, *465* (2103), 669−683.

(67) Nigro, P.; Pompilio, G.; Capogrossi, M. C. Cyclophilin A: A Key Player for Human Disease. *Cell Death Dis.* **2013**, *4* (10), e888.

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144

(68) Davis, T. L.; Walker, J. R.; Campagna-Slater, V.; Finerty, P. J.; Paramanathan, R.; Bernstein, G.; MacKenzie, F.; Tempel, W.; Ouyang, H.; Lee, W. H.; Eisenmesser, E. Z.; Dhe-Paganon, S. Structural and Biochemical Characterization of the Human Cyclophilin Family of Peptidyl-Prolyl Isomerases. *PLoS Biol.* **2010**, *8* (7), e1000439.
(69) Fanghänel, J.; Fischer, G. Insights into the Catalytic Mechanism of Peptidyl Prolyl Cis/Trans Isomerases. *Front. Biosci., Landmark Ed.* **2004**, *9*, 3453−3478.
(70) Leone, V.; Lattanzi, G.; Molteni, C.; Carloni, P. Mechanism of Action of Cyclophilin A Explored by Metadynamics Simulations. *PLoS Comput. Biol.* **2009**, *5* (3), e1000309.
(71) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. Intrinsic Dynamics of an Enzyme Underlies Catalysis. *Nature* **2005**, *438* (7064), 117−121.
(72) Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T. Hidden Alternative Structures of Proline Isomerase Essential for Catalysis. *Nature* **2009**, *462* (7273), 669−673.
(73) Tork Ladani, S.; Hamelberg, D. Intricacies of Interactions, Dynamics and Solvent Effects in Enzyme Catalysis: A Computational Perspective on Cyclophilin A. *Mol. Simul.* **2014**, *40* (10−11), 765−776.
(74) Wapeesittipan, P.; Mey, A. S. J. S.; Walkinshaw, M. D.; Michel, J. Allosteric Effects in Cyclophilin Mutants May Be Explained by Changes in Nano-Microsecond Time Scale Motions. *Commun. Chem.* **2019**, *2* (1), 41.
(75) Zydowsky, L. D.; Etzkorn, F. A.; Chang, H. Y.; Ferguson, S. B.; Stolz, L. A.; Ho, S. I.; Walsh, C. T. Active Site Mutants of Human Cyclophilin A Separate Peptidyl-Prolyl Isomerase Activity from Cyclosporin A Binding and Calcineurin Inhibition. *Protein Sci.* **1992**, *1* (9), 1092−1099.
(76) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA, and Proteins. *J. Comput. Chem.* **1995**, *16* (11), 1357−1377.
(77) Mackerell, A. D. Empirical Force Fields for Biological Macromolecules: Overview and Issues. *J. Comput. Chem.* **2004**, *25* (13), 1584−1604.
(78) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346* (6216), 1510−1514.
(79) Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N. Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem. Inf. Model.* **2019**, *59* (5), 2367−2373.
(80) Kasimova, M. A.; Lindahl, E.; Delemotte, L. Determining the Molecular Basis of Voltage Sensitivity in Membrane Proteins. *J. Gen. Physiol.* **2018**, *150* (10), 1444−1458.
(81) Sun, R.-N.; Gong, H. Simulating the Activation of Voltage Sensing Domain for a Voltage-Gated Sodium Channel Using Polarizable Force Field. *J. Phys. Chem. Lett.* **2017**, *8* (5), 901−908.
(82) Crnjar, A.; Comitani, F.; Hester, W.; Molteni, C. *Trans* − *Cis* Proline Switches in a Pentameric Ligand-Gated Ion Channel: How They Are Affected by and How They Affect the Biomolecular Environment. *J. Phys. Chem. Lett.* **2019**, *10* (3), 694−700.
(83) Otten, R.; Liu, L.; Kenner, L. R.; Clarkson, M. W.; Mavor, D.; Tawfik, D. S.; Kern, D.; Fraser, J. S. Rescue of Conformational Dynamics in Enzyme Catalysis by Directed Evolution. *Nat. Commun.* **2018**, *9* (1), 1314.
(84) Thompson, M. C.; Barad, B. A.; Wolff, A. M.; Sun Cho, H.; Schotte, F.; Schwarz, D. M. C.; Anfinrud, P.; Fraser, J. S. Temperature-Jump Solution X-Ray Scattering Reveals Distinct Motions in a Dynamic Enzyme. *Nat. Chem.* **2019**, *11* (11), 1058−1066.
(85) Juárez-Jiménez, J.; Gupta, A. A.; Karunanithy, G.; Mey, A. S. J. S.; Georgiou, C.; Ioannidis, H.; De Simone, A.; Barlow, P. N.; Hulme, A. N.; Walkinshaw, M. D.; Baldwin, A. J.; Michel, J. Dynamic Design: Manipulation of Millisecond Timescale Motions on the Energy Landscape of Cyclophilin A. *Chem. Sci.* **2020**, *11*, 2670−2680.
(86) Murgida, D. H.; Hildebrandt, P. Heterogeneous Electron Transfer of Cytochrome c on Coated Silver Electrodes. Electric Field Effects on Structure and Redox Potential. *J. Phys. Chem. B* **2001**, *105* (8), 1578−1586.

MOST WIEDZY Downloaded from mostwiedzy.pl

https://dx.doi.org/10.1021/acs.jcim.0c00217
J. Chem. Inf. Model. 2020, 60, 3131−3144