THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL Letter

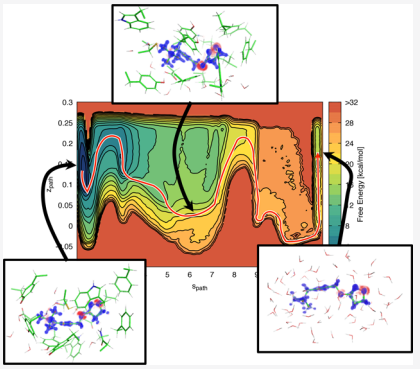
Accuracy of Molecular Simulation-Based Predictions of k_{off} Values: A Metadynamics Study

Riccardo Capelli,*,◇ Wenping Lyu,◇ Viacheslav Bolnykh, Simone Meloni, Jógvan Magnus Haugaard Olsen, Ursula Rothlisberger, Michele Parrinello, and Paolo Carloni

Cite This: J. Phys. Chem. Lett. 2020, 11, 6373–6381 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The k_{off} values of ligands unbinding to proteins are key parameters for drug discovery. Their predictions based on molecular simulation may under- or overestimate experiment in a system- and/or technique-dependent way. Here we use an established method—infrequent metadynamics, based on the AMBER force field—to compute the k_{off} of the ligand iperoxo (in clinical use) targeting the muscarinic receptor M_2. The ligand charges are calculated by either (i) the Amber standard procedure or (ii) B3LYP-DFT. The calculations using (i) turn out not to provide a reasonable estimation of the transition-state free energy. Those using (ii) differ from experiment by 2 orders of magnitude. On the basis of B3LYP DFT QM/MM simulations, we suggest that the observed discrepancy in (ii) arises, at least in part, from the lack of electronic polarization and/or charge transfer in biomolecular force fields. These issues might be present in other systems, such as DNA−protein complexes.



The efficacy and safety of drugs depend critically on their residence time.[1,2] Indeed, k_{off} values—the drug unbinding rate constant, corresponding to the inverse of the residence time—correlates with clinical efficiency even more than binding affinity.[3,4] Hence, the k_{off} value is one of the crucial parameters that current drug design strives to improve.[5,6] While experiments face challenges to identify and characterize rate-limiting transition state(s), simulation approaches are able to predict free energy landscapes and residence times.[7] Techniques devoted to this aim range from long-time molecular dynamics (MD) with specialized hardware[8] to a variety of different enhanced sampling methods such as random acceleration MD (RAMD),[9,10] hyperdynamics,[11] conformational flooding,[12] Markov state models (MSMs),[13] dissipation-corrected targeted MD,[14] and infrequent[15] or frequency adaptive metadynamics.[16] The latter three approaches have also predicted k_{off} values for ligands binding to cytoplasmatic proteins. The values differ by 1 or 2 orders of magnitude from experiments.[13,17−19] The discrepancy, irrespective of the force field used (either Amber99SB[20]/GAFF[21] or CHARMM22*[22]), could be caused by a variety of factors, including force field accuracy, molecular modeling procedures, and sampling issues. Here, we use a multistep simulation approach to address this important issue. We focus on a ligand, iperoxo (Figures 1 and SI1), routinely used in neuroimaging in the clinics. The ligand targets the human muscarinic acetylcholine receptor 2 (M_2). Overall, the system consists of ∼150 000 atoms (figure 1).

First, we attempt to calculate the k_{off} value of the ligand by Amber14SB force field[23]-based well-tempered[24] and frequency-adaptive MetaD. To minimize errors due to the modeling procedure, we use the same pH and ionic strength as in the experimental conditions.[25] We use two approaches to calculate the drug RESP charges. The first one is the Amber standard methodology, based on HF/6-31G* calculations (**RESP-HF**).[21] This has been used to predict the free energy landscape associated with ligand binding to the protein,[26] with a calculated binding affinity in excellent agreement with experiment. The second methodology is based on density functional theory (DFT) with the B3LYP[27−29] exchange−correlation functional (**RESP-B3LYP**).

Then, we perform quantum mechanics/molecular mechanics (QM/MM) calculations. The QM region consists of the ligand and its interacting residues, and it is described by the same B3LYP functional as in the second parametrization. The remaining part is treated as in all the previous calculations performed here, with the Amber14SB force field.[23]

The simulations based on **RESP-HF** charges turn out to face difficulties in obtaining a reasonable estimation of the transition-state free energy (and thus k_{off} value). Those with **RESP-B3LYP** charges lead to a calculated k_{off} of $3.7 \pm 0.7 \times 10^{-4}$ s^{-1}. This value is more reasonable but still much smaller than the experimental value ($1.0 \pm 0.2 \times 10^{-2}$ s^{-1}).

Received: March 30, 2020
Accepted: July 16, 2020
Published: July 16, 2020

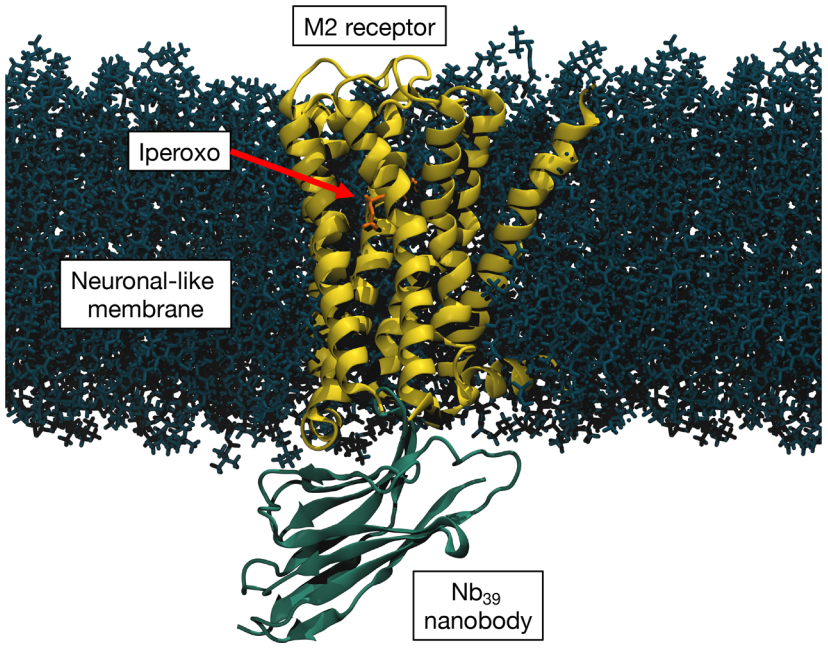


Figure 1. Representation of the simulated system. The M_2 receptor (yellow), with its agonist iperoxo (orange), is embedded in a neuronal-like membrane[30] (dark blue), and it is bound to the Nb_{39} nanobody (green) which keeps it in its active state. Water and Na^+/Cl^- solvated ions are removed for the sake of clarity.

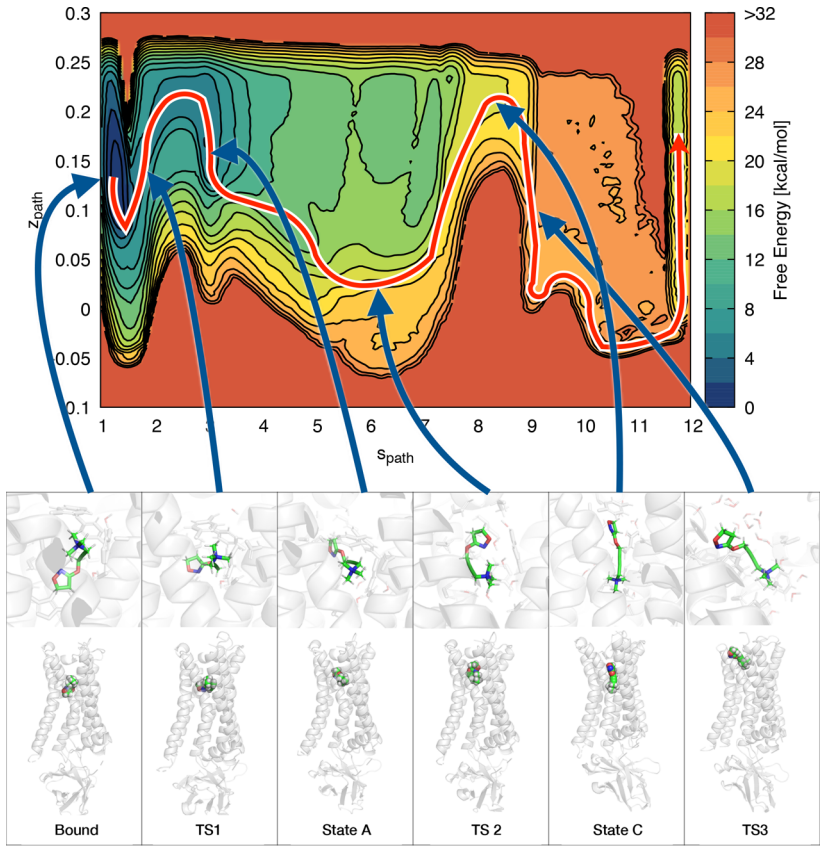


Figure 2. Free energy surface of binding with the observed unbinding pathway and representative structures. The top part shows the M_2/iperoxo free energy surface as a function of s_{path} and z_{path}, with the unbinding pathway followed by the ligand. The bottom part represents the bound, TS1, A, TS2, C, and TS3 states: in the upper panel, both the ligand and its surrounding atoms within 4.5 Å are rendered; in the lower panel, the ligand is rendered in sphere mode with all the receptor and the nanobody. Water and ions are not represented for the sake of clarity.

Comparison of the force-field-based simulations with those based on QM/MM shows a remarkable agreement between the *ab initio* and the force field estimation of the ligand/protein binding energy (Figure 1). However, this is not the case for the

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381

transition state. Our analysis indicates that the lack of polarization may be one of the key factors causing this discrepancy, which in turn affects the accuracy of the k_{off} calculation.

Unbinding Process Using the RESP-HF Parametrization. MetaD-based calculations of k_{off} values require the free energy landscape of ligand unbinding pathways. Previously, we have predicted this for the ligand iperoxo and its target M_2 muscarinic receptor in its active state by using well-tempered metadynamics.[26] The calculations were based on the Amber14SB force field[23] for the protein. The RESP charges for the ligand were calculated at the HF/6-31G* level of theory (**RESP-HF** hereafter). Two different unbinding pathways emerged (here noted with **I** and **II**)[26] (see the Supporting Information, Figure 2). Pathway **I** is the lowest in free energy. Here, the ligand starts from the bound state, rotates around the axis formed by the alkyne bond, passing through the transition state 1 (TS1, Figure 2), to finally reach state **A**. After this step, we observe a rotation of the entire ligand with the trimethylammonium group as a pivot (transition state 2, TS2); TS2 is the rate-limiting step. A salt bridge between the trimethylammonium group and ASP103, present in the bound state, is broken here, and the overall number of intramolecular hydrogen bonds between protein and the ligand is reduced (Figure 7 in the Supporting Information) and substituted with other H-bonds with solvent. A slightly different rotation around the same pivot can lead to state **B**. This second rotation does not lead to unbinding and is not further considered here. After reaching state **C**, the trimethylammonium group breaks the salt bridge formed with ASP103 (transition state 3, TS3) and moves toward the extracellular part, reaching the fully solvated state. The rate-limiting step is TS2. Pathway **II** is identical to **I** until the ligand reaches state **C**. Here, the rearrangement of the extracellular loop 2 (ECL2) of the receptor limits the possibility of the ligand to reach the solvated state right after visiting state **C**, forcing it to perform a further rotation, reaching the last metastable state **D**. After that rotation, the ligand breaks the salt bridge formed with ASP103 and it moves toward the solvent, completing its pathway to solvation.

Here we evaluate the k_{off} by a multistep approach as in Casasnovas et al.[17] Because kinetic calculations are very expensive, we explore only the lowest free energy pathway, **I**.[26] For this, we identify the path collective variables (pathCVs) s_{path} and z_{path}.[31] These are particularly appropriate to study a single pathway between two reference states, limiting the motion of the system only around this predefined path. s_{path} defines the progression along the pathway, while z_{path} samples deviation from the reference path (in our case **I**). Next, we perform well-tempered-MetaD[24] to calculate the free energy as a function of s_{path} and z_{path}. Finally, we use frequency adaptive MetaD,[16] for the actual calculation of k_{off}.

The definition of a pathCV is based on a metric that measures the distance of instantaneous configurations from the path. In the first applications of pathCV,[31] the metric chosen was based on root-mean-square displacement (RMSD) with respect to the initial bound state. In our case, having seen the presence of multiple intermediate states along pathway **I**[26] (Figure 5), we prefer to define our metrics based on a contact map based on the ligand−protein atom pairs that are crucial for the stabilization of the intermediate states (see the Supporting Information for detail). To obtain a sequence of conformations along the unbinding trajectory, we employed Ratchet&Pawl MD[32,33] in a two step-approach. First, we forced our system to perform its unbinding transition using as CV the distance between the binding pocket and the center of mass of the ligand (in the same spirit as a previous work[34]); from these structures we built a first pathCV. As a second step, we apply Ratchet&Pawl to this first pathCV. From the set of conformations obtained during this run, we built the final pathCVs that we employed in the MetaD simulations.

Unbinding Process Using the RESP-B3LYP Parametrization. Multiple-walkers[35] well-tempered[24] MetaD along the pathCV previously obtained lead to a large number of recrossing events in all 10 replicas. The calculated free energy surface along pathway **I** (Figure 2) shows that the unbinding process was the same as that obtained by **RESP-HF**[26] calculations (Figure 2 in the Supporting Information). In particular, all the intermediate states identified by the two setups are the same. Therefore, we use the same pathCVs for both **RESP-HF** and **RESP-B3LYP** frequency adaptive MetaD[16] calculations of k_{off}.

Calculations of k_{off}. We performed 10 different frequency-adaptive MetaD[16] runs, biasing both s_{path} and z_{path}. For **RESP-HF**, the bias needed to perform the first ring rotation from bound state to state **A** in the first run exceeded 140 kJ/mol. This bias corresponds roughly to a residence time in the order of years (because the acceleration factor is exponentially proportional to the bias deposited; see the Supporting Information). Therefore, this parametrization could not be used for the calculations of the k_{off}. For **RESP-B3LYP**, 5 production runs could be collected. They covered 0.9 to 1.7 μs, for a total simulation time of ∼8 μs. (The other 5 runs had to be removed because they deposited bias on the transition state, invalidating the sampling performed.[15])

The resulting distribution of calculated residence times (Figure 3) was fitted with a Poisson distribution. From this, we

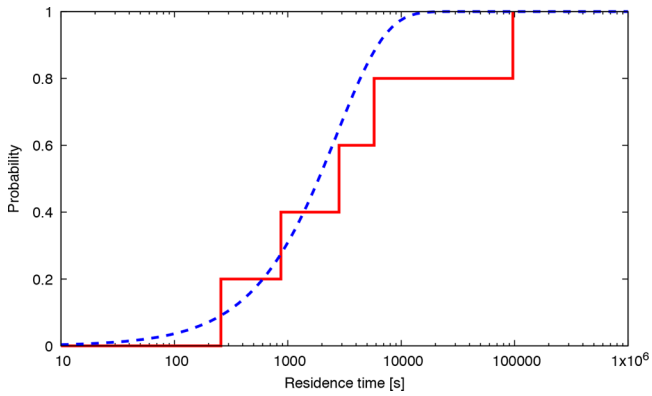


Figure 3. Comparison between the calculated (red line) and the theoretical Poisson distributions (blue dashed curve).

obtained a residence time $2.7 \pm 0.5 \times 10^3$ s and a $k_{off} = 3.7 \pm 0.7 \times 10^{-4}$ s^{-1}. To validate the correctness of the calculation performed, we performed a Kolmogorov−Smirnov test between the obtained distribution and the theoretical one.[36] The p-value turned out to be 0.87. This shows that the obtained distribution is statistically indistinguishable from a theoretical rare event distribution.

In conclusion, for **RESP-B3LYP** calculations we observe a rare event (as confirmed by that displaying a free energy barrier that is higher with respect to the one experimentally observed). For **RESP-HF**, the free energy barrier between the bound and the first intermediate state **A** is so high that the

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381

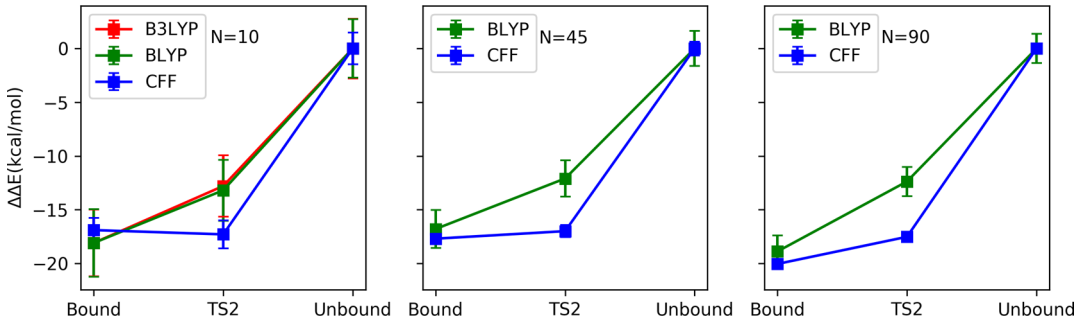


Figure 4. Relative interaction energies ($\Delta\Delta E$) of the ligand to the M_2 receptor at QM/MM level (with BLYP and B3LYP exchange−correlation functionals) and CFF level. N is the number of conformations considered in statistics for each state. $\Delta\Delta E$ is not directly related to the free energy of the binding−unbinding process.

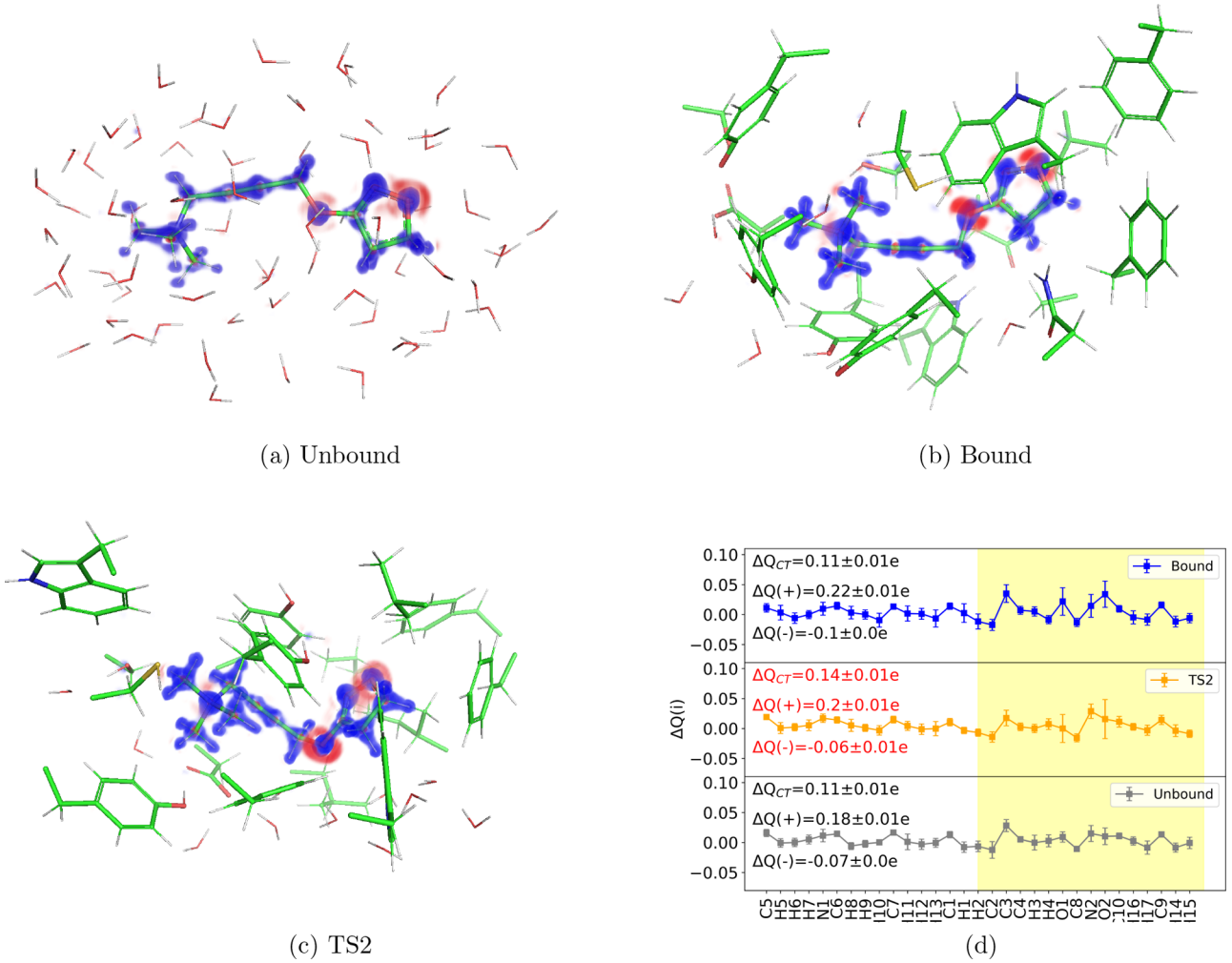


Figure 5. (a−c) Change of electronic density of the ligand on passing from *in vacuo* to in water ($\Delta\rho = \rho^{cmplx} - \rho^{lig} - \rho^{rest}$) for the unbound (a), bound (b), and TS2 (c) states (blue, $\Delta\rho > 0$; red, $\Delta\rho < 0$). The atoms are displayed in stick mode. (d) Corresponding change in atomic charge for each atom ($\Delta Q(i)$, $i = 0-30$). ΔQ_{CT} is the overall charge transferred from the protein to the ligand, ranging from 0.11 to 0.14 electrons. $\Delta Q(+)$ and $\Delta Q(-)$ are the polarizations of the ligand contributed from atoms with positive and negative $\Delta Q(i)$, respectively, ranging from −0.1 to 0.22 electrons. The yellow color band highlights the most significant differences among the three states. The calculations have been carried out using the Voronoi partition scheme.[42]

residence time related to this transition is at least in the order of years. The discrepancy of the **RESP-B3LYP** and, even more, of the **RESP-HF** results, does not seem to originate from the enhanced sampling approach adopted, but it is likely to be ascribed mostly to the underlying potential energy function, associated either with the ligand distortions and/or with ligand−protein interactions along the pathway. As for the first issue, we find that small and similar distortions of the ligand's internal degrees of freedom [such as bond lengths (Figure 3 in the Supporting Information), bond angles (Figure 4 in the Supporting Information), and dihedrals (Figure 5 in the Supporting Information)] are present in both TS2 and bound

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381

states. [However, the ligand undergoes larger structural fluctuation at the TS2 than in the bound state (Figure 6 in Supporting Information).] Thus, the distortion of the molecules is not likely to contribute greatly to the increase of potential energy on passing from the bound state to TS2. We address the second issue by recalculating the ligand−protein energetics at the quantum mechanics level for selected configurations in the two states, considering the unbound state—the ligand in solution—as our reference state (i.e., computing the $\Delta\Delta E$ with respect to the solvated state). We use a QM/MM approach as implemented in the MiMiC multiscale interface.[37,38] The QM part of the system consists of the ligand and its interacting groups (Figure 5a−c; see Computational Methods for details). It is treated at the DFT level, using either the B3LYP[27−29] or the BLYP[27,28] exchange−correlation functionals. The rest is described by the Amber14SB force field,[23] as in all the other calculations presented in this Letter. The energy of the total system is calculated by adding the QM/MM interaction energy to the QM and MM energies, while that of the ligand is given by the QM energy. The $\Delta\Delta E$ values are also calculated at a purely classical force field (CFF) level using **REST-B3LYP**.

In the bound state, $\Delta\Delta E$ values, calculated for the same 10 representative conformations, turn out to be very similar using the force field and DFT (-17.3 ± 1.5 kcal/mol (CFF) and -18.1 ± 3.1 (B3LYP)/-18.1 ± 3.1 (BLYP) kcal/mol); that is, the values are not significantly affected by the exchange−correlation functional (Figure 4 and Table 2 in the Supporting Information). Increasing the number of representative conformations to 45 and 90, calculated for CFF and BLYP, does not modify significantly this picture (-17.7 ± 0.5 (CFF) and -16.8 ± 1.8 (BLYP) and -20.1 ± 0.3 (CFF) and -18.9 ± 1.5 (BLYP) kcal/mol, respectively; Figure 4 or Table 2 in the Supporting Information).

Taken together, these results are consistent with previous well-tempered-MetaD-based free energy calculations using **RESP-HF**, which showed excellent agreement between calculated and experimental affinities for this[26,34] and other systems. This is expected, and it is indeed confirmed by countless examples, in both protein−ligand[39] and protein−protein[40] interactions. This further supports the conclusion that modern force-field-based calculations accurately predict ligand binding affinities.

A dramatically different scenario takes place at the transition state of the unbinding process, TS2. Here the $\Delta\Delta E$ values obtained with the force field (-17.8 ± 1.3 kcal/mol) differ significantly from the DFT values, while the latter are still similar to each other (-12.8 ± 2.9 (B3LYP)/-13.2 ± 2.8 (BLYP) kcal/mol). The trend is preserved when the number of conformations is increased (-17.0 ± 0.6 (CFF) and -12.1 ± 1.7 (BLYP) and -17.5 ± 0.4 (CFF) and -12.4 ± 1.4 (BLYP), for 45 and 90 conformations, respectively; see Figure 4 or Table 2 in the Supporting Information). These results suggest that, within the limitation of our statistics, CFF overestimates the potential energy increase (associated with disruption of several intermolecular interactions) at TS2. This suggestion is fully consistent with our calculations of residence times, which are longer with respect to the experimental data.

Finally, we investigate if changes in electronic polarization[41] and charge-transfer (CT) effects (which are lacking in routinely used biomolecular force fields) occur in the unbinding process. We calculate, using QM/MM, the rearrangement of electronic density of the ligand while passing from *in vacuo* to the bound state, to the transition state TS2, and to the unbound state (Figure 5a−c). The calculations are carried out using both the Voronoi[42] and Bader[43] approaches to the partition of the atomic charges (Figure 8 in the Supporting Information). Polarization effects turn out to involve mostly the positively charged trimethylammonium group (forming a salt bridge with ASP103 or interacting with water) and the two oxygen atoms (Figure 5a−c). The overall CT (ΔQ_{CT}) as well as the polarization (ΔQ_{Pol}), albeit small in magnitude (a fraction of an elementary charge), differ slightly on passing from the bound state to the TS (Figure 5d and Figure 8 in the Supporting Information). Obviously, the fixed charge schemes of commonly used force fields cannot take into account these changes in charge redistribution during the unbinding processes.

Correcting this overestimation would lead to a better representation of the energetics and lower the residence time. However, simply adding a correction might not be enough. Indeed, using more accurate potential energy functions (coming, for instance, from apt polarizable force fields[44]) could give a different pathway and thus a different transition state, as already suggested.[45] In addition, the entropic contribution to the free energy of the transition state may be affected by the accuracy of the potential energy surface of force fields.

The k_{off} constant is exponentially related to the height of the free energy barrier for dissociation. Small errors in the force field may therefore introduce large errors in the k_{off} of drugs, a key parameter in pharmacology. Here, we take iperoxo, a superagonist, and we find that the bound state to its target M_2 receptor is excellently described by both the **RESP-HF** and the **RESP-B3LYP** parametrization of the ligand effective point charges. This confirms the adequacy of the classical force-field representation in the minimum-energy states of the system. Thus, although the electrostatic parametrization in AMBER currently lacks polarization of the system, a high level of accuracy has been achieved by tuning all the parameters of the force field for more than 4 decades.[46,47] (Although the total energy function is highly accurate, its single contribution to electrostatics may not provide the electrostatic energy.) In contrast, the accuracy of the force field at the transition state appears to be limited.[45] This may be caused, at least in part, by the fact that standard nonpolarizable force fields cannot capture the slight changes of electronic polarization and charge-transfer effects (Figure 5) on passing from the bound to the transition state. Furthermore, a more detailed representation of the interaction could in principle find different unbinding pathways, which can change the energetics of the intermediate states. Polarizable force fields,[44,48,49] reactive MD,[50] and/or corrections of the free energy landscape derived from quantum mechanical calculations[45] might alleviate this problem. Similar and possibly even more severe issues may be expected in other unbinding processes, such as those involving protein−protein and protein−DNA complexes.

COMPUTATIONAL METHODS

System Preparation. We followed a protocol similar to that of ref 26. We obtained the structure from the protein data bank (PDB code: 4MQS[51]), parametrized the ligand using GAFF,[21] and embedded it in a neuronal-like[30] membrane. The drug charges were obtained, after an initial geometry optimization, by restrained electric potential fitting method (RESP)[52] with HF/6-31G* (**RESP-HF**) and B3LYP/6-31G* (**RESP-B3LYP**)

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381

levels of theory. The system was then solvated; ions were added to reach the experimental ionic force, and finally the system was minimized and equilibrated (see details in the Supporting Information). All the simulations were performed using GROMACS 2018.4[53] patched with PLUMED 2.5.[54]

rMD Simulation and pathCV Identification. To identify a collection of conformations to set up our pathCV variable,[31] we used Ratchet&Pawl MD.[32,33] As an initial ratcheting coordinate, we considered the distance between the center of mass of our ligand and the center of mass of the pocket (defined by TYR104, SER107, VAL111, PHE195, and TYR239), projected along the axis normal to the membrane. We fixed the bias factor to k = 500 kJ/mol/nm and the final ratchet coordinate to r_{final} = 2.5 nm. After 10 different 20 ns long rMD runs with these parameters, we selected 11 frames that describe well the progression of the ligand toward the solvent, and we used them to define a pathCV based on the contact map[55] between a list of atom pairs of the ligand and the receptor (list in the Supporting Information). With this variable, we performed 10 new 20 ns long rMD runs to verify and eventually refine the new variable, choosing again by visual inspection 11 different frames from the unbinding trajectories and redefining a final pathCV that was then used in our MetaD simulation.

Well-Tempered MetaD along the pathCV. Using the pathCV as identified above, we performed multiple-walkers[35] well-tempered MetaD[24] along the z_{path} and s_{path} using 10 different walkers. We set the bias factor to 24 and used an initial Gaussian height of 1.2 kJ/mol and a frequency deposition of 1 ps^{-1}. To limit the phase space exploration to path **I** only, we prevented our system from reaching pathway **II** by putting a restraint to avoid its motion along that pathway (i.e., we put a harmonic wall at z_{path} = 0.25, where paths **I** and **II** diverge). The total simulation time was 1.8 μs. The free energy surface was reweighted *a posteriori* with the Tiwary and Parrinello algorithm.[56]

Frequency Adaptive MetaD. We carried out 10 different frequency adaptive MetaD[16] runs. The approach is a variant of I-MetaD[15] that speeds up the calculations (details in Supporting Information). We performed 10 different frequency-adaptive MetaD runs, with a bias factor of 24, an initial Gaussian height of 1.2 kJ/mol, an initial frequency deposition of 1 ps^{-1}, an acceleration parameter θ = 100 (details in the Supporting Information), and a minimum frequency deposition of 10^{-2} ps^{-1}. Out of all simulation performed, 5 have deposited bias on the transition state, and thus, we discarded them, obtaining the 5 residence times from the remaining simulations.

QM/MM Single-Point Calculations. A selection of N = 45 or 90 structures associated with the bound, TS2, and unbound states (Figure 2) underwent 1000 steps of energy minimization using the steepest descent algorithm at the CFF level. Then, for each structure, we considered the *total system*, the *rest* (i.e., the system without the ligand), and the *ligand* without the systems (i.e., in vacuum). Overall, 270 structures were considered.

The QM regions in the *total system* consisted of the ligand and the side-chains (up to the -C_β) directly interacting with it as well as water molecules within 4.5 Å from it. They ranged from 196 to 308 atoms (see Table 3 in the Supporting Information). The QM regions of the *rest* were the same except that the ligand was not included. They ranged from 165 to 277 atoms. Those of the *ligand* included only the latter (31 atoms).

The QM part was described at the DFT level (QM part, Figure 2), using either the B3LYP or BLYP exchange–correlation functional.[57] A plane-wave basis set with a cutoff of 90 Ry was used. The core electrons were described through norm-conserving Troullier–Martins pseudopotentials.[58] Isolated system conditions were achieved by using the Martyna–Tuckerman scheme.[59] For the Bound and TS2 states, (i) covalent C_α–C_β bonds across the QM/MM interface were described by an adapted monovalent carbon pseudopotential;[60] (ii) the net charge of the residues considered at QM level were reweighted to their side chain atoms by neutralizing the sum of the partial charges of the remaining backbone atoms in MM region. For the *total system* and *rest*, the atoms other than those in the QM regions ("MM region") were described using exactly the same setup and force field as for the metadynamics. The interactions between the QM and MM parts were described as in ref 37: the electrostatic interactions were calculated explicitly for MM atoms that are within 30 Å of the centroid of the QM part, whereas the interactions with the rest of the system were evaluated using a fifth-order multipole expansion of the electrostatic potential.[37] In all the calculations, the Grimme's correction[61] was used to describe dispersion interactions.

Single-point electronic structure calculations were performed, with a convergence criteria of 10^{-7} au, using the highly scalable MiMiC-based QM/MM interface,[37,38] which combines CPMD 4.1[62] with GROMACS 2019.4.[63]

The ligand binding energy ΔE was calculated either at QM/MM (ΔE(B3LYP/BLYP)) or at the classical force field (CFF) (that is, at the **RESP-B3LYP**) level (ΔE(CFF)). It reads

$$\Delta E = E_{\text{total}} - E_{\text{ligand}} - E_{\text{rest}} \tag{1}$$

where E_{total} and E_{rest} are the potential energies of the *total* system and of the *rest* (given by summing the QM energy with the MM energy and the QM/MM interaction energies), and E_{ligand} the potential energy of the *ligand*.

We computed ΔE at BLYP and CFF levels for N = 45 and N = 90 conformations for each state, in order to verify the consistency of our evaluation. To verify the same effect at a higher level of theory, we chose 10 structures covering the same spreading range of the calculated energies at the BLYP level for the more expensive and accurate B3LYP calculations. (For comparison, the statistical estimate at BLYP and CFF were re-evaluated for the 10 structures.)

The change in electron density upon ligand binding was calculated at the B3LYP level for N = 10.

$$\Delta\rho = \rho^{\text{total}} - \rho^{\text{ligand}} - \rho^{\text{rest}} \tag{2}$$

Here, ρ^{total} is the electron density of the QM part embedded in the MM part of the *total system*, ρ^{ligand} that of the *ligand*, and ρ^{rest} that of the *rest*.

The electron charge transfer (CT) associated with atom i of the ligand reads:[42]

$$\Delta Q(i) = \int_{VP_i} \Delta\rho(\mathbf{r})d\mathbf{r} \tag{3}$$

The integral is solved numerically over the grid points within the Voronoi partition[42] or Bader's atom in molecules partition[64] of atom i (VP_i). An in house code (cpmd-cube-tools: https://pypi.org/project/cpmd-cube-tools/) and the

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373–6381

Bader code (Bader charge analysis: http://theory.cm.utexas.edu/henkelman/code/bader/) were used, respectively.

The CT effect of the whole ligand molecule reads

$$\Delta Q_{CT} = \sum_i \Delta Q(i) \tag{4}$$

An estimation of the change in charge distribution is given by electric polarization as

$$\Delta Q_{Pol} = |\Delta Q(+)| + |\Delta Q(-)| \tag{5}$$

where, $\Delta Q(+) = \sum_i \Delta Q(i)$, $i \in \{\Delta Q(i) > 0\}$ and $\Delta Q(-) = \sum_i \Delta Q(i)$, $i \in \{\Delta Q(i) < 0\}$.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.0c00999.

Details on Ratchet&Pawl MD, FA MetaD, and system preparation and choice for the CV and QM/MM calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Riccardo Capelli − *Computational Biomedicine Section, IAS-5/INM-9, Forschungzentrum Jülich, D-52425 Jülich, Germany; JARA-HPC, Forschungszentrum Jülich, D-54245 Jülich, Germany;* orcid.org/0000-0001-9522-3132; Email: r.capelli@fz-juelich.de, riccardo.capelli@polito.it

Authors

Wenping Lyu − *Computational Biomedicine Section, IAS-5/INM-9, Forschungzentrum Jülich, D-52425 Jülich, Germany;* orcid.org/0000-0001-9176-5444

Viacheslav Bolnykh − *Computational Biomedicine Section, IAS-5/INM-9, Forschungzentrum Jülich, D-52425 Jülich, Germany; Laboratory of Computational Chemistry and Biochemistry, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland;* orcid.org/0000-0002-7090-1993

Simone Meloni − *Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Ferrara, I-44121 Ferrara, Italy;* orcid.org/0000-0002-3925-3799

Jógvan Magnus Haugaard Olsen − *Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway; Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark;* orcid.org/0000-0001-7487-944X

Ursula Rothlisberger − *Laboratory of Computational Chemistry and Biochemistry, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland;* orcid.org/0000-0002-1704-8591

Michele Parrinello − *Department of Chemistry and Applied Biosciences, ETH Zürich, CH-6900 Lugano, Ticino, Switzerland; Facoltà di Informatica, Istituto di Scienze Computazionali, Università della Svizzera Italiana (USI), CH-6900 Lugano, Ticino, Switzerland; Istituto Italiano di Tecnologia, I-16163 Genova, Italy*

Paolo Carloni − *Computational Biomedicine Section, IAS-5/INM-9 and JARA-Institute INM-11: Molecular Neuroscience and Neuroimaging, Forschungzentrum Jülich, D-52425 Jülich, Germany;* orcid.org/0000-0002-9010-0149

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.0c00999

Author Contributions

◇R.C. and W.L. share first authorship.

Notes

The authors declare no competing financial interest.

All the data and PLUMED input files required to reproduce the results reported in this Letter are available on PLUMED-NEST (www.plumed-nest.org), the public repository of the PLUMED consortium,[66] as plumID:20.005.

■ ACKNOWLEDGMENTS

The authors thank Anna Bochicchio and Rodrigo Casasnovas for the initial preparation of the system and Emiliano Ippoliti, Luca Maggi, and GiovanniMaria Piccini for useful discussions. The authors gratefully acknowledge the computing time granted through VSR on the supercomputer JURECA[65] at Forschungszentrum Jülich (Project ID: jias5d) and acknowledge the JSC for the computing time on the supercomputer JURECA Booster module. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 785907 (HBP SGA2) and the Marie Sklodowska-Curie Grant Agreement No. 642069. W.L. appreciates the National Natural Science Foundation of China No. 21505134. J.M.H.O. acknowledges financial support from the Research Council of Norway through its Centres of Excellence scheme (Project ID: 262695). U.R. acknowledges funding from the Swiss National Science Foundation via the NCCR MUST and individual grants.

■ REFERENCES

(1) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug−target residence time and its implications for lead optimization. *Nat. Rev. Drug Discovery* **2006**, *5*, 730−739.

(2) Swinney, D. C. The role of binding kinetics in therapeutically useful drug action. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 31−39.

(3) Lu, H.; Tonge, P. J. Drug−target residence time: critical information for lead optimization. *Curr. Opin. Chem. Biol.* **2010**, *14*, 467−474.

(4) Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular determinants of drug−receptor binding kinetics. *Drug Discovery Today* **2013**, *18*, 667−673.

(5) Schmidtke, P.; Luque, F. J.; Murray, J. B.; Barril, X. Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J. Am. Chem. Soc.* **2011**, *133*, 18903−18910.

(6) Dickson, A.; Tiwary, P.; Vashisth, H. Kinetics of ligand binding through advanced computational approaches: a review. *Curr. Top. Med. Chem.* **2017**, *17*, 2626−2641.

(7) Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Recent progress in molecular simulation methods for drug binding kinetics. *arXiv*, **2020**, 2002.08983.

(8) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13118−13123.

(9) Kokh, D. B.; Kaufmann, T.; Kister, B.; Wade, R. C. Machine learning analysis of tauRAMD trajectories to decipher molecular determinants of drug-target residence times. *Frontiers in molecular biosciences* **2019**, *6*, 36.

(10) Kokh, D. B.; Doser, B.; Richter, S.; Ormersbach, F.; Cheng, X.; Wade, R. C. A Workflow for Exploring Ligand Dissociation from a Macromolecule: Efficient Random Acceleration Molecular Dynamics Simulation and Interaction Fingerprints Analysis of Ligand Trajectories. *arXiv*, **2020**, 2006.11066

(11) Voter, A. F. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.* **1997**, *78*, 3908.

(12) Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1995**, *52*, 2893.
(13) Plattner, N.; Noé, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **2015**, *6*, 7653.
(14) Wolf, S.; Lickert, B.; Bray, S.; Stock, G. Multisecond ligand dissociation dynamics from atomistic simulations. *Nat. Commun.* **2020**, *11*, 2918.
(15) Tiwary, P.; Parrinello, M. From metadynamics to dynamics. *Phys. Rev. Lett.* **2013**, *111*, 230602.
(16) Wang, Y.; Valsson, O.; Tiwary, P.; Parrinello, M.; Lindorff-Larsen, K. Frequency adaptive metadynamics for the calculation of rare-event kinetics. *J. Chem. Phys.* **2018**, *149*, 072309.
(17) Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding kinetics of a p38 MAP kinase type II inhibitor from metadynamics simulations. *J. Am. Chem. Soc.* **2017**, *139*, 4780−4788.
(18) Wang, Y.; Martins, J. M.; Lindorff-Larsen, K. Biomolecular conformational changes and ligand binding: from kinetics to thermodynamics. *Chemical science* **2017**, *8*, 6466−6473.
(19) Wang, Y.; Ribeiro, J. M. L.; Tiwary, P. Past−future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics. *Nat. Commun.* **2019**, *10*, 3573.
(20) 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*, 712−725.
(21) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(22) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **2011**, *100*, L47−L49.
(23) 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*, 3696−3713.
(24) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(25) Schrage, R.; Holze, J.; Klöckner, J.; Balkow, A.; Klause, A. S.; Schmitz, A.-L.; De Amici, M.; Kostenis, E.; Tränkle, C.; Holzgrabe, U.; et al. New insight into active muscarinic receptors with the novel radioagonist (3H) iperoxo. *Biochem. Pharmacol.* **2014**, *90*, 307−319.
(26) Capelli, R.; Bochicchio, A.; Piccini, G.; Casasnovas, R.; Carloni, P.; Parrinello, M. Chasing the full free energy landscape of neuroreceptor/ligand unbinding by metadynamics simulations. *J. Chem. Theory Comput.* **2019**, *15*, 3354−3361.
(27) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098−3100.
(28) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(29) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(30) Chan, R. B.; Oliveira, T. G.; Cortes, E. P.; Honig, L. S.; Duff, K. E.; Small, S. A.; Wenk, M. R.; Shui, G.; Di Paolo, G. Comparative lipidomic analysis of mouse and human brain with Alzheimer disease. *J. Biol. Chem.* **2012**, *287*, 2678−2688.
(31) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **2007**, *126*, 054103.
(32) Marchi, M.; Ballone, P. Adiabatic bias molecular dynamics: a method to navigate the conformational space of complex molecular systems. *J. Chem. Phys.* **1999**, *110*, 3697−3702.
(33) Tiana, G.; Camilloni, C. Ratcheted molecular-dynamics simulations identify efficiently the transition state of protein folding. *J. Chem. Phys.* **2012**, *137*, 235101.
(34) Saleh, N.; Ibrahim, P.; Saladino, G.; Gervasio, F. L.; Clark, T. An efficient metadynamics-based protocol to model the binding affinity and the transition state ensemble of G-protein-coupled receptor ligands. *J. Chem. Inf. Model.* **2017**, *57*, 1210−1217.
(35) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B* **2006**, *110*, 3533−3539.
(36) Salvalaglio, M.; Tiwary, P.; Parrinello, M. Assessing the reliability of the dynamics reconstructed from metadynamics. *J. Chem. Theory Comput.* **2014**, *10*, 1420−1425.
(37) Olsen, J. M. H.; Bolnykh, V.; Meloni, S.; Ippoliti, E.; Bircher, M. P.; Carloni, P.; Rothlisberger, U. MiMiC: A Novel Framework for Multiscale Modeling in Computational Chemistry. *J. Chem. Theory Comput.* **2019**, *15*, 3810−3823.
(38) Bolnykh, V.; Olsen, J. M. H.; Meloni, S.; Bircher, M. P.; Ippoliti, E.; Carloni, P.; Rothlisberger, U. Extreme Scalability of DFT-Based QM/MM MD Simulations Using MiMiC. *J. Chem. Theory Comput.* **2019**, *15*, 5601−5613.
(39) Perez, A.; Morrone, J. A.; Simmerling, C.; Dill, K. A. Advances in free-energy-based simulations of protein folding and ligand binding. *Curr. Opin. Struct. Biol.* **2016**, *36*, 25−31.
(40) Siebenmorgen, T.; Zacharias, M. Computational prediction of protein−protein binding affinities. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, e1448.
(41) Gao, J.; Xia, X. A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* **1992**, *258*, 631−635.
(42) Carloni, P.; Andreoni, W.; Hutter, J.; Curioni, A.; Giannozzi, P.; Parrinello, M. Structure and Bonding in Cisplatin and Other Pt(II) Complexes. *Chem. Phys. Lett.* **1995**, *234*, 50−56.
(43) Bader, R. F. Atoms in molecules. *Acc. Chem. Res.* **1985**, *18*, 9−15.
(44) Li, H.; Chowdhary, J.; Huang, L.; He, X.; MacKerell, A. D., Jr; Roux, B. Drude polarizable force field for molecular dynamics simulations of saturated and unsaturated zwitterionic lipids. *J. Chem. Theory Comput.* **2017**, *13*, 4535−4552.
(45) Haldar, S.; Comitani, F.; Saladino, G.; Woods, C.; van der Kamp, M. W.; Mulholland, A. J.; Gervasio, F. L. A multiscale simulation approach to modeling drug−protein binding kinetics. *J. Chem. Theory Comput.* **2018**, *14*, 6093−6101.
(46) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science* **1987**, *235*, 574−576.
(47) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. Free energy perturbation calculations on binding and catalysis after mutating Asn 155 in subtilisin. *Nature* **1987**, *328*, 551−554.
(48) Villa, F.; MacKerell, A. D., Jr; Roux, B.; Simonson, T. Classical drude polarizable force field model for methyl phosphate and its interactions with Mg2+. *J. Phys. Chem. A* **2018**, *122*, 6147−6155.
(49) Li, H.; Ngo, V.; Da Silva, M. C.; Salahub, D. R.; Callahan, K.; Roux, B.; Noskov, S. Y. Representation of ion−protein interactions using the drude polarizable force-field. *J. Phys. Chem. B* **2015**, *119*, 9401−9416.
(50) Meuwly, M. Reactive molecular dynamics: From small molecules to proteins. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, *9*, No. e1386.
(51) Kruse, A. C.; Ring, A. M.; Manglik, A.; Hu, J.; Hu, K.; Eitel, K.; Hübner, H.; Pardon, E.; Valant, C.; Sexton, P. M.; et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **2013**, *504*, 101−106.
(52) Bayly, C. I.; Cieplak, P.; Cornell, W.; 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.
(53) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19−25.

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381

(54) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(55) Bonomi, M.; Branduardi, D.; Gervasio, F. L.; Parrinello, M. The unfolded ensemble and folding mechanism of the C-terminal GB1 β-hairpin. *J. Am. Chem. Soc.* **2008**, *130*, 13938−13944.
(56) Tiwary, P.; Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **2015**, *119*, 736−742.
(57) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(58) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *43*, 1993−2006.
(59) Martyna, G. J.; Tuckerman, M. E. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* **1999**, *110*, 2810−2821.
(60) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. Variational optimization of effective atom centered potentials for molecular properties. *J. Chem. Phys.* **2005**, *122*, 014113.
(61) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(62) Hutter, J. A.; Deutsch, T.; Bernasconi, M.; Goedecker, S.; Marx, D.; Tuckerman, M.; Parrinello, M. *CPMD*; Copyright IBM Corp 1990−2019; MPI für Festkörperforschung: Stuttgart, 1997−2012.
(63) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(64) Tang, W.; Sanville, E.; Henkelman, G. A grid-based Bader analysis algorithm without lattice bias. *J. Phys.: Condens. Matter* **2009**, *21*, 084204.
(65) Krause, D.; Thörnig, P. JURECA: general-purpose supercomputer at Jülich supercomputing centre. *Journal of large-scale research facilities JLSRF* **2016**, *2*, 62.
(66) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G.; Bonas, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.

https://dx.doi.org/10.1021/acs.jpclett.0c00999
J. Phys. Chem. Lett. 2020, 11, 6373−6381