Simulating Protein–Ligand Binding with Neural Network Potentials

Shae-Lynn Lahey and Christopher N. Rowley

Drug molecules adopt a range of conformations both in solution and in their protein-bound state. The strain and reduced flexibility of bound drugs can partially counter the intermolecular interactions that drive protein–ligand binding. To make accurate computational predictions of drug binding affinities, computational chemists have attempted to develop efficient empirical models of these interactions, although these methods are not always reliable. Machine learning has allowed the development of highly-accurate neural-network potentials (NNPs), which are capable of predicting the stability of molecular conformations with accuracy comparable to state-of-the-art quantum chemical calculations but at a billionth of the computational cost. Here, we demonstrate that these methods can be used to represent the intramolecular forces of protein-bound drugs within molecular dynamics simulations. These simulations are shown to be capable of predicting the protein–ligand binding pose and conformational component of the absolute Gibbs energy of binding for a set of drug molecules. Notably, the conformational energy for anti-cancer drug erlotinib binding to its target was found to considerably overestimated by a molecular mechanical model, while the NNP predicts a more moderate value. Although the ANI-1ccX NNP was not trained to describe ionic molecules, reasonable binding poses are predicted for charged ligands, although this method is not suitable for modeling the ligands in solution.

Introduction

Molecular simulation of the binding of small molecules to proteins has provided computational prediction and rationalization of the affinity and selectivity of drugs with their targets. These simulations rely on molecular mechanical (MM) force fields to describe the intra and intermolecular interactions of the solvent, protein, and ligand. These force fields are constructed from simple mathematical functions that approximate the potential energy surface of the protein–ligand complex. These “force fields” require the definition of a large set of parameters, which are typically chosen to yield the closest agreement with empirical or quantum chemical data.

Development of accurate models of potential energy terms of protein–ligand binding and their optimal parameters is a longstanding objective in computational chemistry. The electrostatic, repulsive, and dispersion interaction terms have been developed actively; however, accurate representation of intramolecular potential energy of the ligand is particularly challenging and no complete, general solution has been developed. Most current force fields use the same terms as the first models that were developed more than 50 years ago, where bond angles and stretches are described with harmonic potentials (i.e., spring-like) and torsional barriers are defined as the sum of a handful of cosine functions. Force fields for drug-like compounds are particularly difficult to develop because of the enormous variety of chemical motifs, which often feature complex chemical effects like conjugation, hyperconjugation, and aromaticity. This is compounded by the enormous variety of chemical motifs that are possible in chemical drug space, where each could require a distinct set of parameters. For example, the proprietary OPLS3 force field defines 124 atom types and 48142 torsional parameters. Other methods provide options to reparameterize force fields automatically using ab initio calculations, although this complicates the simulation workflow and can be computationally expensive.

Recently, machine-learned neural network potentials (NNPs) have emerged as an alternative to conventional MM force fields. The ANI models developed by Roitberg and coworkers define the atomic positions in terms of a set of “symmetry functions”, which are constructed from the position of a given atom relative to nearby atoms. A neural network is trained...
to reproduce the high-level ab initio electronic energies (i.e., CCSD(T)) from these data. These potentials are remarkably robust and predict the structures and relative stabilities of molecular conformations across a broad set of chemical structures with similar accuracy to the high-level ab initio data they were trained to reproduce. The computational cost of NNPs is comparable to molecular mechanical models, so they can be used to perform ns length simulations of molecules containing dozens or hundreds of atoms routinely.

Here, we present a strategy to simulate protein–ligand complexes using a machine-learned NNP to represent the intramolecular interactions of the ligand. This model is embedded inside a conventional MM force field for the protein and solvent, so established models for these components can be used without modification. We call this method NNP/MM, as it functions the same as Quantum Mechanical / Molecular Mechanical (QM/MM) models do, but with the NNP used in place of the QM method. This method is tested for its ability to predict the poses of protein-bound drugs in comparison to electron density distributions determined by X-ray crystallography. The Gibbs energies for restraining the ligands to their bound conformations are calculated using NNP/MM and compared to the CGenFF force field.

Computational Methods

Theory

In this method, the potential energy of the whole system is defined as the sum of the potential energy of the subsystem described by the NNP (i.e., the intramolecular interactions of the ligand) (\(\gamma_{\text{NNP}}\)), the potential energy of the environment around the ligand (\(\gamma_{\text{MM}}\)), and the interactions between the ligand and (\(\gamma_{\text{NNP/MM}}\)) and its environment (Eqn. 1).

\[
\gamma(r) = \gamma_{\text{MM}}(r_{\text{MM}}) + \gamma_{\text{NNP}}(r_{\text{NNP}}) + \gamma_{\text{NNP/MM}}(r)
\]  

(1)

The MM region is represented using a conventional MM force field, so \(\gamma_{\text{MM}}\) is calculated in the normal fashion for an additive force field. For non-covalent protein–ligand binding, the \(\gamma_{\text{NNP/MM}}\) term is the conventional MM non-bonded interactions between the protein and the ligand, which is simply the sum of Lennard-Jones and pairwise Coulombic interactions between the NNP atoms and MM atoms (Eqn. 2).

\[
\gamma_{\text{NNP/MM}}(r) = \sum_{i} \sum_{j} \frac{q_i q_j}{4\pi \varepsilon_{i j} r_{ij}} + 4\varepsilon_{i j} \left[ \left( \frac{\alpha_{i j}}{r_{ij}} \right)^{12} - \left( \frac{\alpha_{i j}}{r_{ij}} \right)^{6} \right]
\]  

(2)

This functions similarly to mechanically-embedded QM/MM models, where the NNP serves as the “QM” model embedded within the MM system. This method can be employed in many established simulation codes without modification because they can take advantage of existing QM/MM features, which allow the energy and forces of a critical subsection of the system to be calculated using an external method. Sample input files and our scripts can be downloaded from our online repository and will be included in future distributions of NAMD.

Technical Details

All molecular dynamics (MD) simulations were performed using NAMD 2.13. The ligand intramolecular energies and forces were calculated using the ANI-1ccX NNP implemented in the TorchANI package. The programs were interfaced through the general-purpose external-force functionality of the NAMD QM/MM code. The CHARMM36m force field was used to represent the protein and the mTIP3P model was used to represent the water molecules. Sample input files and our scripts can be downloaded from our online repository and will be included in future distributions of NAMD. The CGenFF force field was used to calculate the non-bonded ligand–protein interactions (i.e., \(\gamma_{\text{NNP/MM}}\)).

The calculation of the erlotinib potential energy surface was performed using ORCA 4.2.1. Optimizations with constraints on the amine torsional angle were performed using the resolution of identity 2nd-order Möller–Plesset theory (RI-MP2) with the def2-TZVP basis set. Single point energy evaluations were performed at these optimized structures using DLPNO-CCSD(T)/def2-TZVP to generate the QM potential energy surface.

Test Set

To evaluate the ability of the ANI-1ccX potential to predict the pose of a bound ligand, we developed a test set of protein–ligand complexes. We selected a structurally-diverse set of complexes where a high-resolution crystallographic structure of the protein–ligand complex was available, including several where the ligand is in a conformationally-strained pose. The ANI-1ccX NNP is only defined for carbon, nitrogen, hydrogen, and oxygen, so only ligands comprised of these elements were selected. The full details of the structures are included in Supplementary Materials.

Simulations of Ligand Binding Poses

The NNP/MM ligand binding poses were generated by MD simulations of the protein–ligand complexes. The crystallographic structure (including crystallographic water molecules) was placed in a periodic unit cell of liquid water. The protonation states of the protein and ligand were assigned using H++ 3.2 and inspection of the crystallographic structure. A 5 ns equilibration MD simulation using the CGenFF force field for the ligand was performed where all non-hydrogen atoms of the ligand and protein were restrained to their crystallographic positions. The equilibrated structures were used as the initial structures of 2 NNP/MM MD simulations of the complexes. In these simulations, the \(C_\alpha\) of the protein backbone were restrained to their crystallographic positions using harmonic potentials (\(k_c = 10 \text{ kcal/mol Å}^{-2}\)). These simulations were performed with a thermostat temperature set to correspond to the temperature the crystallographic structure was collected for (e.g., 100 K). The ligand electron density was obtained from the crystallographic electron density map (2fo-fc), selecting all points within 2 Å of the ligand atoms in the PDB structure. An isosurface value of 0.5 was used in the renderings.
Calculation of Conformational Gibbs Energy

Confinement and release alchemical free energy perturbation is a popular technique for calculating absolute protein–binding energies. In these methods, the total binding energy is divided into a set of Gibbs energies for each step in a path where the ligand is constrained to its bound conformation and is then decoupled from its environment. The component corresponding to the reversible work required to constrain the ligand to its bound conformation is defined as $\Delta G_{\text{cons}}$. Physically, this energy corresponds to the reduction of conformational freedom and isomerization to a higher energy conformation that occurs when a ligand binds to a protein. In confine-and-release absolute binding energy calculations schemes, this is the only term where the intramolecular interactions of the ligand are significant. Accordingly, it is only necessary to use the NNP/MM method when calculating this term; the remaining terms can be calculated using conventional force fields. Notably, this step does not include any alchemical transformation, so performing the calculation with NNP/MM does not present any special challenges.

This term can be calculated by defining the root-mean-square deviation (RMSD) of the ligand relative to its bound conformation ($\zeta$) and then calculating the Gibbs energy required to impose a harmonic restraint on the RMSD ($\frac{1}{2}k_c\zeta^2$) so that the ligand is restricted to hold its bound conformation. Using umbrella sampling, the potential of mean force (PMF) can be calculated as a function of the RMSD. Integration of this PMF biased by the harmonic restraining function provides the ($\Delta G_{\text{cons}}$). In this work, the PMF for both the ligand within the site was calculated from an umbrella sampling simulation where the windows were separated by 0.5 Å and a harmonic biasing potential with a spring constant of 50 kcal/mol Å$^{-2}$ was used. Each window was sampled by performing a 1 ns equilibration simulation followed by a 4 ns sampling simulation. The PMF was constructed from the umbrella simulation using Weighted Histogram Analysis Method (WHAM) with statistical uncertainties of the profiles estimated by bootstrap analysis.

$$
e^{-\Delta G_{\text{cons}}/k_B T} = \frac{\int e^{-[\zeta(\zeta)+\frac{1}{2}k_c\zeta^2]/k_BT} d\zeta}{\int e^{-[\zeta(\zeta)/k_BT]} d\zeta}$$

where $k_c$ is a harmonic potential to restrain the conformation of the ligand at the reference structure. In this work, a value of $k_c = 10$ kcal/mol Å$^{-2}$ was used.

These calculations are performed for the ligand bound to the protein and in solution. The difference of these energies provides the conformational or “strain” component of the absolute binding energy,

$$\Delta G_{\text{cons}} = \Delta G_{\text{cons,site}} - \Delta G_{\text{cons, solvent}}$$

Results and Discussion

Prediction of Ligand Poses

Figure 1 shows the ligand poses generated from the ANI/MD simulations overlaid with the crystallographic electron density maps of the ligand. Generally, the NNP/MM ligand pose overlaps well with the crystallographic density. The positions of the ligand phenyl rings in the thrombin complex (3DA9) and the biotin carboxylase complex (2W6N) are the most significant deviation. The NNP/MM model still relies on conventional MM parameters for the protein–ligand and water–ligand interactions, so these deviations may not be related to the NNP component of the model.

One notable success of the NNP/MM potential is in predicting the binding pose of erlotinib to the epidermal growth factor receptor (EGFR). The core scaffold of this drug is composed of amine-linked ethynyl-phenyl and quinazoline rings. Crystallographic structures of the protein-bound complex show the quinazoline ring bound in the adenosine-binding site while the ethynyl-phenyl group binds in a pocket formed by the T702, T830, and K721 residues. The binding pose predicted by CGenFF is inconsistent with the XRD data, in which the two rings form a more acute angle relative to each other ($\phi_1 = 63 \pm 1^\circ, \phi_2 = 4 \pm 1^\circ$). The simulation using the NNP/MM model is more consistent with the crystallographic data, whereby ($\phi_1 = 44 \pm 1^\circ, \phi_2 = 4 \pm 1^\circ$).

Surprisingly, the poses predicted for the ligands that contain charged functional groups (2HY, S ETA, and 3EIG) are reasonable even though the ANI-1ecX potential was not designed to describe charged species and none of the molecules this NNP was trained for were charged.

Conformational Free Energies

The conformational strain of the ligand that occurs in protein–ligand binding arises from the need for the ligand to adopt the conformation it holds in its bound form. The bound conformations of binding was endergonic. For example, Roux and coworkers’ calculations of the binding affinity of imatinib to Abl kinase predicted that while the net interaction energy of binding was $-27.7$ kcal/mol, the conformational energy countered this by $11.3$ kcal/mol. The conformational energies for the test set of ligands were estimated by calculating the PMF for the deviation from the bound pose using umbrella-sampling MD simulations with both the CGenFF and NNP/MM models. $\Delta G_{\text{cons}}$ was calculated from these PMFs using Eqn. 3. These energies are collected in Table 1. The PMFs for all complexes are presented in Supplementary Information.

Table 1 Conformational Gibbs energy of binding for protein–ligand complexes calculated using the MM(CGenFF) and NNP/MM methods. All energies are in kcal/mol.

| PDB ID | $\Delta G_{\text{cons}}$ CGenFF | $\Delta G_{\text{cons}}$ NNP/MM | charge  |
|--------|-------------------------------|-------------------------------|---------|
| 1XOZ   | $0.4 \pm 0.0$                 | $0.5 \pm 0.0$                 | 0       |
| 2W6N   | $4.7 \pm 0.1$                 | $5.2 \pm 0.1$                 | 0       |
| 3EYG   | $1.9 \pm 0.1$                 | $1.0 \pm 0.2$                 | 0       |
| 4HJO   | $13.0 \pm 0.1$                | $8.3 \pm 0.1$                 | 0       |
| 4NCCT  | $3.4 \pm 0.1$                 | $2.8 \pm 0.2$                 | 0       |
| 2HY    | $8.1 \pm 0.1$                 | $32.6 \pm 0.1$                | 1       |
| 3EIG   | $11.1 \pm 0.0$                | $37.7 \pm 0.0$                | -2      |
| 3ETA   | $5.6 \pm 0.1$                 | $15.2 \pm 0.2$                | 1       |

Amongst the neutral ligands, the NNP/MM conformational energies are generally similar in magnitude to the CGenFF strain
Fig. 1 Calculated poses of ligands (red) in protein binding sites. The crystallographic electron density of the ligands are shown in blue. The PDB ID, protein name, and ligand name are included beneath the image.
energies. This indicates that the ANI-1ccX model can achieve similar results to the CGenFF model despite the lack of any explicit parameterization for these molecules. The conformational energies of 4HJO (erlotinib bound to EGFR) show the largest difference, with the NNP/MM strain energy being 4.7 kcal/mol smaller than the CGenFF strain energy. The high strain predicted by the CGenFF model is due to the amine functional group of erlotinib holding a pyramidal geometry in the solution simulations, creating a large energetic penalty to force the drug into its bound conformation. In the NNP/MM simulation of erlotinib in solution, the amine group remains close to a co-planar geometry with respect to the quinazoline ring, with a moderate skew in the dihedral angle between the phenyl group and the amine.

The ligands that contain charged functional groups (2HYY, 3ETA, and 3EIG) have anomalously high conformational energies. This issue originates from the use of the ANI-1ccX NNP, which was only trained on neutral molecules. This NNP predicts reasonable geometries of the ammonium and carboxylate groups in these molecules, but these ionic functional groups form spurious intramolecular contacts with other parts of the molecule in the solution NNP/MM MD simulations. This results in the stabilization of regions of the PMF corresponding to large structural deviations from the bound pose. As the NNP(ANI-1ccX) model was not designed for the description of charged molecules like these, it is unsuitable for calculating their conformational energies.

Extensive MD simulations are needed to calculate $\Delta G_{\text{cons}}$ by calculating the PMF of the RMSD, but these simulations were completed at a modest computational cost because of efficient implementations of the ANI model for execution on graphical processing units. For example, the NNP/MM MD simulations of imatinib (69 atoms) executed at a rate of 3.4 ns/day on a single Titan Xp NVIDIA GPU. Even faster performance is anticipated after the planned integration of NNPs directly into NAMD and other molecular simulation codes.

**Torsional Potential Energy Surface of Erlotinib**

The large difference in the ANI-1ccX and CGenFF conformational energies of 4HJO (erlotinib bound to EGFR) originate from the ligand adopting conformations in solution that are drastically different than the bound conformation when the CGenFF model is used, while the NNP/MM model predicts similar conformations in both the binding site and solution. This is evident in the CGenFF PMF of the ligand’s conformation relative to its bound pose in Figure 2, which is considerably broader than the NNP/MM PMF and is higher energy in the crystallographic pose (RMSD=0 Å).

The geometry of the erlotinib amine linker and its aromatic substituents deviates sharply from the bound pose in the CGenFF solution structure; the amine is partially pyramidalized and the aromatic substituents are skew to each other. In contrast, in the NNP/MM simulation, the amine predominantly remains in a planar geometry, conjugated with the quinazoline and phenyl rings.

The potential energy surface corresponds to rotations around the amine torsion angles. The minima on the CGenFF surface corresponds to structures where the amine is significantly pyramidal and the substituent phenyl and quinazoline rings adopt angles that reduce steric repulsion between them. The ANI-1ccX surface is consistent with the DLNPO-CCSD(T) surface, where there is a broad global minimum centered around ($\phi_1 = 0^\circ, \phi_2 = 0^\circ$) and the amine nitrogen holds a planar arrangement with the aromatic groups.

The failure of the CGenFF force field stems from the lack of a distinct atom type for amines conjugated with aromatic rings. While it would be possible to adjust the parameters of the CGenFF force field to improve its description of the arylamine potential energy surfaces, this introduces a new fitting stage and requires computationally demanding QM calculations to provide the target data. Generally, it is not immediately apparent where a general-purpose force field will fail. By using NNPs to calculate these interactions, these issues are avoided entirely because energy surfaces with near-CCSD(T) accuracy can be generated efficiently and without the need to parameterize the intramolecular potential energy surface explicitly.

**Conclusions**

NNPs provide accurate representations of the intramolecular interactions of drug molecules in molecular simulations of protein–ligand binding. These simulations take advantage of established MM models of the protein and solution, while eliminating the need to develop a force field for the intramolecular interactions of each ligand. By employing a NNP that has already been trained on a broad set of molecular species, the fundamental intramolecular interactions that give rise to the molecular energy surface are captured without the need to parameterize a force field. This representation is also free of the harmonic/torsional/improper scheme used in conventional force fields. This allows the sim-
Fig. 3 (a) The fragment of erlotinib used to calculate the potential energy surface. Truncated groups are shown in grey. (b) Representative solution conformations of erlotinib for the CGenFF MM model (green) and NNP/MM model (red) overlaid with the ligand pose from the 4HJO crystal structure (c) The relaxed potential energy surfaces for rotation around the erlotinib fragment amine bonds calculated using (i) DLPNO-CCSD(T)//MP2 (ii) NNP(ANI-1ccX) and (iii) the CGenFF MM model. Energies are in kcal/mol.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank NSERC of Canada for funding through the Discovery Grants program (RGPIN-05795-2016). SLJL thanks the School of Graduate Studies at Memorial University for a graduate fellowship, the Faculty of Science of Memorial University for a Science Undergraduate Research Award, and Dr. Liqin Chen for a scholarship. Computational resources were provided by Compute Canada (RAPI: djk-615-ab). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. CNR thanks the Beckman Institute at the University of Illinois for a travel grant.

Notes and references

1. K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell, Journal of Chemical Information and Modeling, 2012, 52, 3155–3168.
2. M. Riquelme, A. Lara, D. L. Mobley, T. Verstraelen, A. R. Matamala and E. Vöhringer-Martinez, Journal of Chemical Information and Modeling, 2018, 58, 1779–1797.
3. M. Mohebifar, E. R. Johnson and C. N. Rowley, Journal of Chemical Theory and Computation, 2017, 13, 6146–6157.
4. E. T. Walters, M. Mohebifar, E. R. Johnson and C. N. Rowley, J. Phys. Chem. B, 2018, 122, 6690–6701.
5. S. Lifson and A. Warshel, J. Chem. Phys., 1968, 49, 5116–5129.
6. E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyman, M. K. Dahlgren, J. L. Knight and et al., J. Chem. Theory Comput., 2016, 12, 281–296.
7 L. Huang and B. Roux, *J. Chem. Theory Comput.*, 2013, 9, 3543–3556.
8 L.-P. Wang, T. J. Martinez and V. S. Pande, *J. Phys. Chem. Lett.*, 2014, 5, 1885–1891.
9 C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid and J. C. Gumbart, *J. Comput. Chem.*, 2013, 34, 2757–2770.
10 L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives and W. L. Jorgensen, *Nucleic Acids Res.*, 2017, 45, W331–W336.
11 J. Behler, *Angewandte Chemie International Edition*, 2017, 56, 12828–12840.
12 J. S. Smith, O. Isayev and A. E. Roitberg, *Chem. Sci.*, 2017, 8, 3192–3203.
13 J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev and A. E. Roitberg, *Nat. Comm.*, 2019, 10, 2903.
14 J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 2007, 98, 146401.
15 D. Bakowies and W. Thiel, *J. Phys. Chem.*, 1996, 100, 10580–10594.
16 L. Hickey and C. Rowley, *NAMD–TorchANI interface scripts*, https://github.com/RowleyGroup/NNP-MM, 2019.
17 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, 26, 1781–1802.
18 R. Group, *TorchANI*, 2019, https://github.com/aiqm/torchani.
19 M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone and et al., *Nat. Methods*, 2018, 15, 351–354.
20 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and J. MacKerell, Alexander D., *Nat. Methods*, 2017, 14, 71–73.
21 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, 79, 926–935.
22 E. Neria, S. Fischer and M. Karplus, *J. Chem. Phys.*, 1996, 105, 1902–1921.
23 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell Jr., *J. Comput. Chem.*, 2010, 31, 671–690.
24 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, 2, 73–78.
25 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, 7, 3297–3305.
26 Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo and F. Neese, *J. Chem. Phys.*, 2018, 148, 011101.
27 R. Anandakrishnan, B. Aguilar and A. V. Onufriev, *Nucleic Acids Research*, 2012, 40, W537–W541.
28 J. Wang, Y. Deng and B. Roux, *Biophys. J.*, 2006, 91, 2798–2814.
29 D. L. Mobley, J. D. Chodera and K. A. Dill, *J. Chem. Theory Comput.*, 2007, 3, 1231–1235.
30 D. L. Mobley and K. A. Dill, *Structure*, 2009, 17, 489–498.
31 J. C. Gumbart, B. Roux and C. Chipot, *J. Chem. Theory Comput.*, 2013, 9, 3789–3798.
32 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, 13, 1011–1021.
33 B. Roux, *Comput. Phys. Commun.*, 1995, 91, 275–282.
34 A. Grossfield, *WHAM: the weighted histogram analysis method, version 2.0.6*, http://membrane.urmc.rochester.edu/content/wham, 2018.
35 Y.-L. Lin, Y. Meng, W. Jiang and B. Roux, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, 110, 1664–1669.
36 O. Isayev, personal communication.