Supporting Information for Determinants of Alanine Dipeptide Conformational Equilibria on Graphene and Hydroxylated Derivatives

Horacio Poblete[1], Ingrid Miranda-Carvajal[2], and Jeffrey Comer*[1]

[1]Institute of Computational Comparative Medicine, Nanotechnology Innovation Center of Kansas State, Department of Anatomy and Physiology, Kansas State University, Manhattan, Kansas, 66506-5802
[2]Universidad Nacional de Colombia

Uncertainty in the Free Energies Determined by ABF. Uncertainties were computed by splitting the ABF calculations in half and comparing the gradients of the potential of mean force from the first half to those of the second half. For the first half, the gradients and counts were obtained directly from the history files produced by the Colvars module [1]. The final gradients and counts include force samples from the complete simulations, i.e. from both the first and second halves. To separate the contribution from only the second half of the simulations, for each bin i, we calculated

$$g_i^{(2)} = \frac{n_i^{\mathrm{F}} g_i^{\mathrm{F}} - n_i^{(1)} g_i^{(1)}}{n_i^{\mathrm{F}} - n_i^{(1)}}, \tag{1}$$

where $n_i^{(1)}$ and $g_i^{(1)}$ were the counts and gradients for the first half of the simulation and n_i^{F} and g_i^{F} were the final counts and gradients. Assuming that the two halves of the simulation are statistically independent, the uncertainty in the gradient were then assumed to be

$$\delta[g_i] = \left| g_i^{(2)} - g_i^{(1)} \right|. \tag{2}$$

As described by Comer et al. [2], the uncertainty in changes in the free energy (Δw) from point a to b was calculated by numerically integrating over the uncertainties of the gradients between the points. The uncertainties of the gradient are assumed to be independent; thus, we used the square root of the sum of the squares of $\delta[g_i]$, rather than the simple sum:

$$\delta[\Delta w_{a\to b}] = \Delta\xi \left(\sum_{i=i_a}^{i_b} \delta[g_i]^2 \right)^{1/2}, \tag{3}$$

*E-mail: jeffcomer@ksu.edu

where $\Delta\xi$ is the bin width along the transition coordinate. $\Delta\xi$ is also the transition coordinate increment for discrete integration of the gradient to obtain the potential of mean force. For the potential of mean force as a function of distance between Ac-Ala-NHMe and the top graphene layer, the value of this potential $w_{1D}(Z)$ and its uncertainty are conventionally taken to be zero at large separations. Thus, we choose the position a to be $Z =$ 12 Å. The uncertainty of the potential of mean force grows monotonically as the distance from point a increases, i.e. as a larger number of gradient uncertainties are summed over. This effect can be seen clearly in Fig. S2B further below.

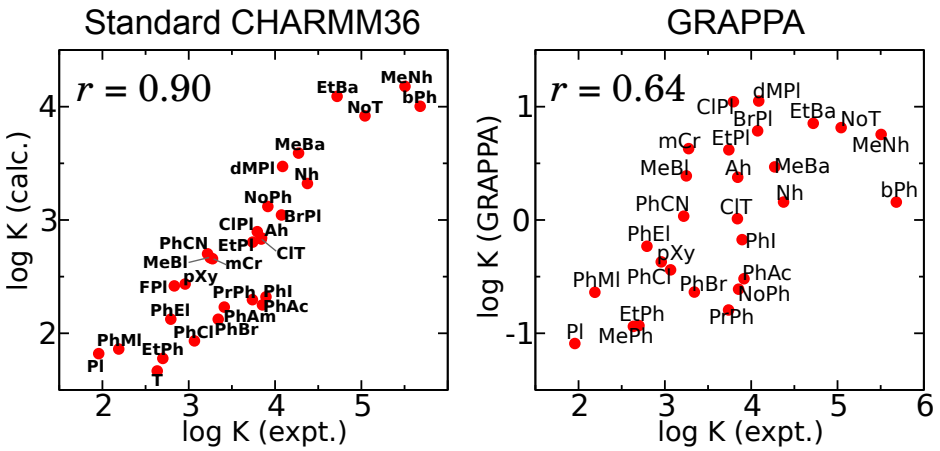

Figure S1: Force field comparison between standard CHARMM and GRAPPA parameters for graphene. Comparison of the logarithm of the adsorption equilibrium constants derived from experiment and simulation for small aromatic adsorbates on large-diameter naked carbon nanotubes. The left panel graphs results from simulations performed using the standard aromatic type (CG2R61) of the CHARMM General Force Field version 3.0.1 [3] for the graphenic surface, while the right uses the polarizable GRAPPA [4] model for graphene. The Pearson correlation coefficient (r) is shown in each graph. Note that the vertical scales are different and GRAPPA appears to considerably underestimate the adsorption affinity. See Comer *et al.* [5] for the key to the abbreviations of the adsorbate molecules.

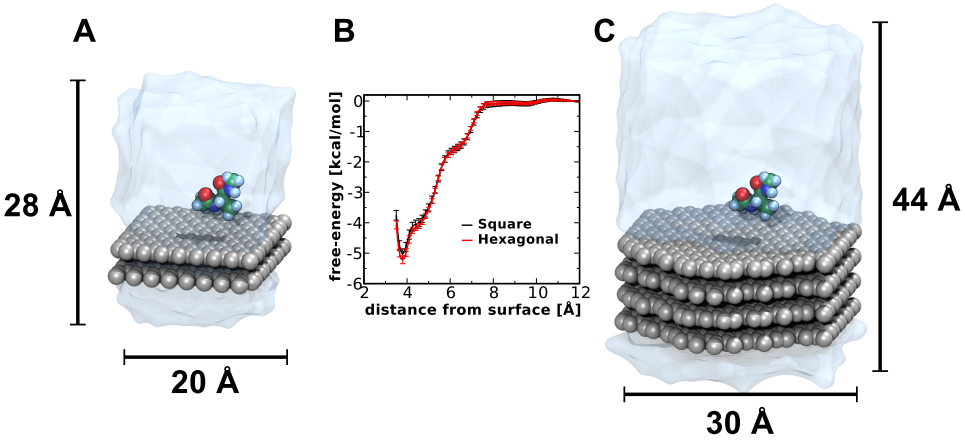

Figure S2: Atomistic simulation systems. (A) Minimal rectangular system containing 1218 atoms. (B) Free energy of Ac-Ala-NHMe as function of Z for systems shown in panels A and C. (C) Initial hexagonal system containing 3448 atoms. The free energy profiles are consistent within the statistical uncertainty of the calculation, while the smaller system gives 1.8 times the performance (318 ns/day on a 6-core workstation with 1 NVIDIA GPU).

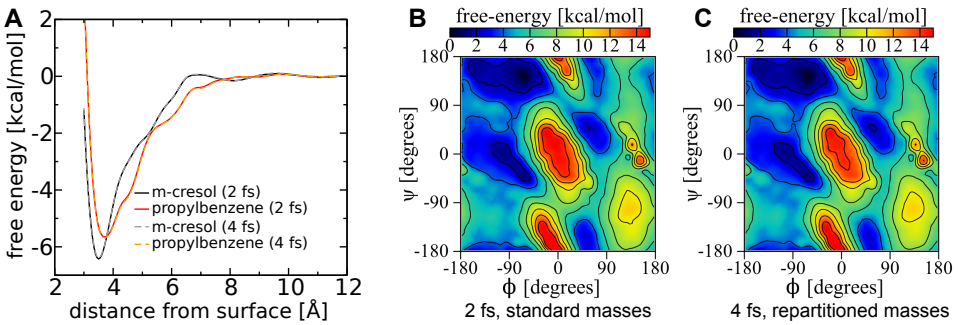

Figure S3: Sanity check for the mass-repartitioning scheme. (A) Calculated free energy as a function of distance between the top layer of graphene and *m*-cresol or propylbenzene (Z) in an aqueous medium. The graphene model is the same as that shown in Fig. S2A and the other simulation parameters are similar to those described in the main text. The dashed curves are for simulations where mass repartitioning[6] of non-water hydrogen atoms was applied, allowing the use of a 4 fs time step. These dashed curves very closely overlap the solid curves, which result from standard-mass simulations at a 2 fs time step. The simulations total about 800 ns each. (B,C) Free-energy as a function of Ac-Ala-NHMe (ϕ, ψ) dihedral angles for Ac-Ala-NHMe at the graphene–water interface using standard masses (B) or mass repartitioning (C). In principle, modification of masses cannot affect the free energy (at least when gravitational forces are neglected); however, the larger time step could potentially introduce greater truncation errors in the molecular dynamics integration. Supporting the validity of the mass repartitioning scheme, the results here demonstrate high precision agreement between conventional models with 2 fs time steps and mass-repartitioned models using 4 fs. All apparent discrepancies are within the estimated statistical uncertainty. The larger time step results in approximately twice the efficiency, allowing a given level of convergence in the free energy to be obtained in about half the real time.

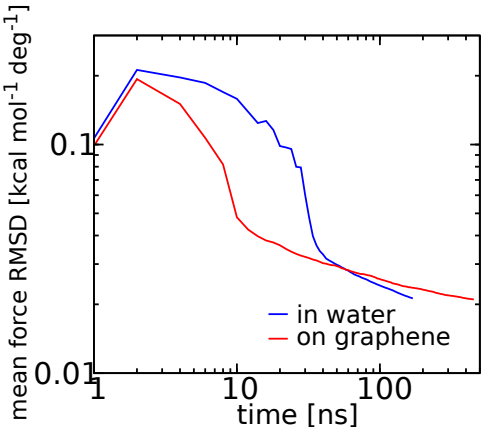

Figure S4: Convergence of ABF calculations in (ϕ, ψ). Each curve is the root-mean-square difference between the mean forces estimated in two independent ABF runs of $w_{2\mathrm{D}}(\phi, \psi)$ as a function of simulated time per run. The independent calculations began from different random velocity distributions. Steady statistical convergence is obtained after about 40 ns. Long-timescale convergence appears to be faster for Ac-Ala-NHMe in water than for Ac-Ala-NHMe on graphene. Comparing the gradients of the potentials of mean force (or equivalently, the mean forces), rather than the potential itself, avoids the ambiguity in the reference potential.

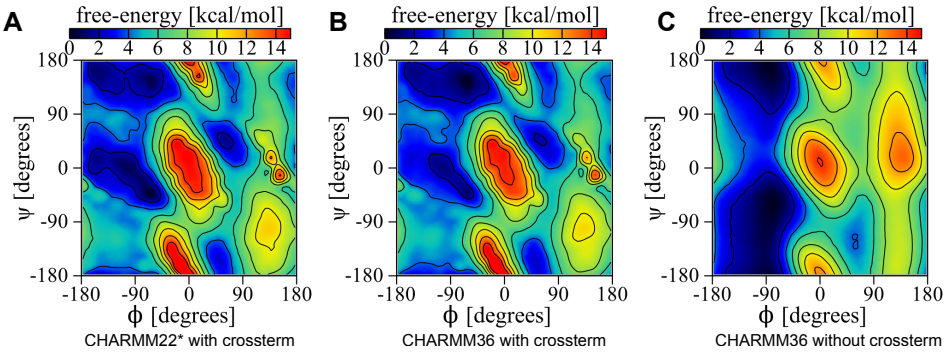

Figure S5: Force-field dependence of Ramachandran free-energy landscape for Ac-Ala-NHMe. Free-energy landscape for Ac-Ala-NHMe using (A) CHARMM22 with tabulated crossterm (CMAP correction) [7], (B) CHARMM36 with tabulated crossterm (CMAP correction) [8], and (C) CHARMM36 without crossterm.

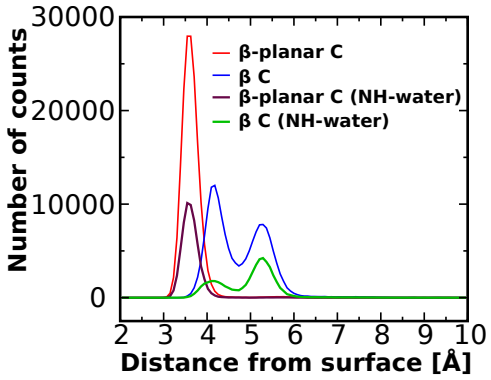

Figure S6: Histograms of the distance between the first layer of graphene and the carbon atom adjacent to the C_α atom of Ac-Ala-NHMe. This figure is similar to Fig. 2A of the main text except separate histograms have been made (burgundy and green curves) over only those frames of the simulation for which the C-terminal NH group is hydrogen bonded to a water molecule. This graph demonstrates that the second peak ($Z \approx 5.3$ Å) for the β conformation is mostly associated with such a hydrogen bond, while the first peak is not.

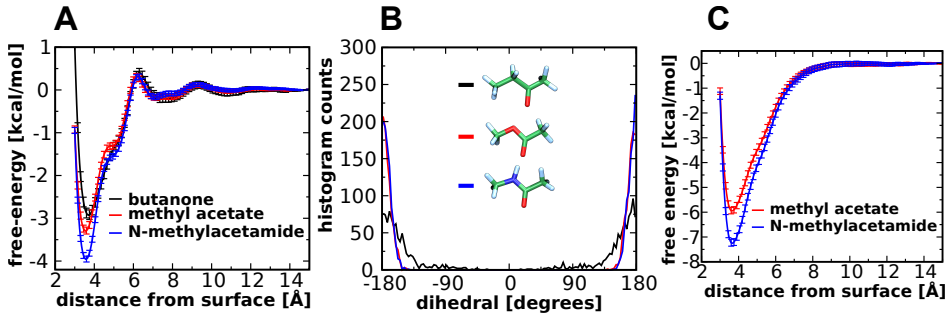

Figure S7: Interaction of N-methylacetamide and its analogues with graphenic surfaces. (A) Calculated free energy as a function of distance between the the top layer of the graphene and the center of mass of butanone (black), methyl acetate (red), or N-methylacetamide (blue) in an aqueous environment. (B) Distribution of the of the dihedral angle of the backbone atoms (C–X–C–C) of butanone (X = C), methyl acetate (X = O), and N-methylacetamide (X = N) in bulk aqueous solution. Both methyl acetate and N-methylacetamide have a high propensity for nearly planar structures. (C) Calculated free energy as function of distance between the top layer of graphene and center of mass of methyl acetate (red) or N-methylacetamide (blue) to in a vacuum environment.

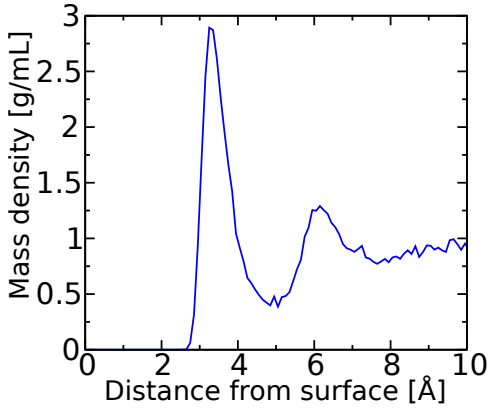

Figure S8: Mass density of water molecules as function of distance from the graphene layer. A high density interfacial layer appears at a distance of 2.9 Å.

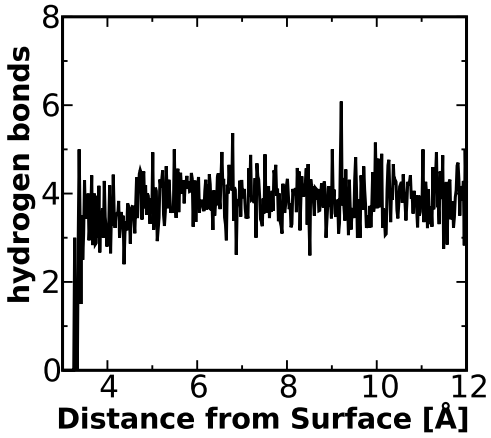

Figure S9: Number of hydrogen bonds between Ac-Ala-NHMe and water as a function of distance from the graphene layer. The number of hydrogen bonds remains relatively constant over all normally accessible distance values. The reduction of the number hydrogen bonds for small values of the distance correspond to unfavorable configurations where steric clashes between the Ac-Ala-NHMe and graphene are present.

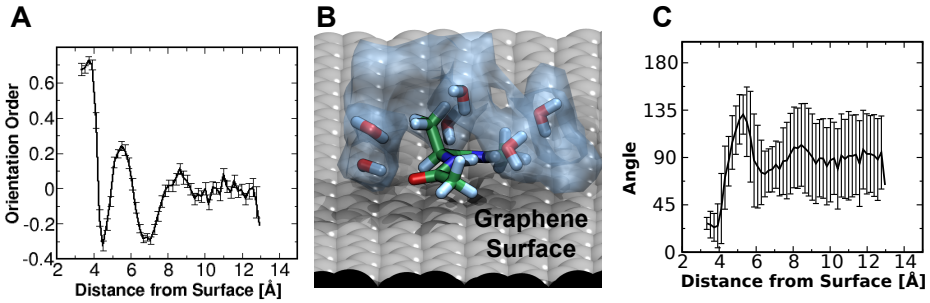

Figure S10: Orientation of the Ac-Ala-NHMe side chain as a function of distance from the graphene layer. (A) Orientational order parameter $(3\cos^2\theta - 1)/2$, where θ is the angle between the z axis and the vector from atoms C_α to C_β for Ac-Ala-NHMe as a function of the distance to the graphene layer. Error bars show the standard error of this parameter. (B) Representative conformation showing the organization of water molecules surrounding Ac-Ala-NHMe adsorbed to graphenic surface. (C) Angle of the side chain θ as a function of the distance to the surface. Error bars show the standard deviation of θ. Large standard deviations for distances > 7 Å imply little orientational order.

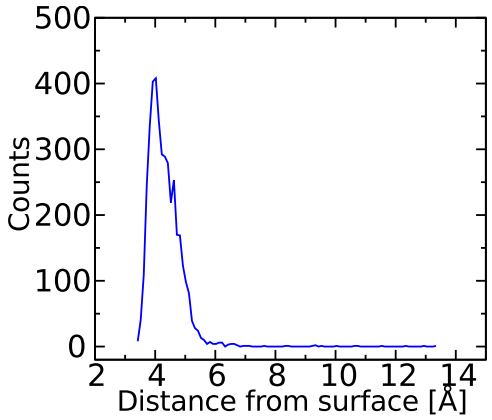

Figure S11: Distance distribution of Ac-Ala-NHMe from graphene layer in $w_{2D}(\phi, \psi)$ free energy calculations. The calculated (ϕ, ψ) PMF represents predominantly adsorbed configurations.

Example simulation files. The file `alad_graphene_oh.zip`, included as part of this Supporting Information, is a ZIP archive containing all of the files needed to calculate the free energy as a function of (ϕ, ψ) for Ac-Ala-NHMe on a hydroxylated graphene surface using NAMD 2.12. The results of an identical calculation are shown in Fig. 4C of the main text. The included files are detailed below.

- **Preparation scripts:**
  - `useful.tcl`: General utilities.
  - `vector.tcl`: Vector algebra utilities.
  - `repartitionHydMasses.tcl`: VMD Tcl script for repartitioning the masses of hydrogen atoms as described in Hopkins *et al.* [6].
  - `markStructure.tcl`: VMD Tcl script for marking the inner layers of graphene so that NAMD can restrain them.
  - `doPrep.sh`: Bash script that executes the previous VMD Tcl scripts.
- **Structure files:**
  - `graphOh1_alad.psf`: The molecular structure file for the model including atom types, charges, and covalent topology (bonds, angles, dihedrals, impropers, and crossterms).

  - `repart_graphOh1_alad.psf`: The same file after repartitioning of the hydrogen atom masses.
  - `graphOh1_alad.pdb`: PDB format file for the model.
  - `rest_graphOh1_alad.pdb`: PDB format file with the inner layers of graphene marked.

- **Parameter files:**
  - `par_all36_prot.prm`: Official CHARMM36 [8] parameter files for proteins.
  - `par_all36_cgenff.prm`: Official CHARMM General Force Field parameter file (version 3.0.1).
  - `toppar_water_ions_prot_cgenff.str`: Official CHARMM parameter file for water and ions with some changes for compatibility with NAMD.
  - `cgenff_circumcoronene_oh4.str`: Some additional bonded parameters for the hydroxylated graphene surface generated by the ParamChem web interface [9] for hydroxylated circumcoronene.

- **Configuration files:**
  - `abf_phi_psi.colvars`: Colvars [1] configuration file implementing adaptive biasing force along the (ϕ, ψ) angles and an additional restraint to keep Ac-Ala-NHMe near the surface. Compatible with NAMD 2.12.
  - `template_abf_graph_4fs.namd`: Template NAMD configuration file containing all of the general molecular dynamics parameters.
  - `namd/abf_graph_oh_alad.0.namd`: Executable NAMD configuration file containing some customizable parameters. Can be executed by `namd2 abf_graph_oh_alad.0.namd`. A previous NAMD run can be continued by incrementing the variable `run`.

- **Pre-equilibrated simulation restart files:**
  - `output/graphOh1_alad.2d_psiphi.0.restart.coor`: Initial coordinates.
  - `output/graphOh1_alad.2d_psiphi.0.restart.vel`: Initial velocities.
  - `output/graphOh1_alad.2d_psiphi.0.restart.xsc`: Initial system geometry.

Colvars [1] configuration file for 3D ABF calculation in (ϕ, ψ, Z) in NAMD 2.12.

```
colvarsTrajFrequency    5000
colvarsRestartFrequency 10000

colvar {
    name phi
    width 10.0
    lowerBoundary -180.0
    upperBoundary 180.0

    dihedral {
       oneSiteTotalForce
       group1 {
           atomNumbers 5
       }
       group2 {
           atomNumbers 7
       }
       group3 {
           atomNumbers 9
       }
       group4 {
           atomNumbers 15
       }
    }
}

colvar {
    name psi
    width 10.0
    lowerBoundary -180.0
    upperBoundary 180.0

    dihedral {
       oneSiteTotalForce
       group1 {
           atomNumbers 17
       }
       group2 {
           atomNumbers 15
       }
       group3 {
           atomNumbers 9
```

```
       }
       group4 {
           atomNumbers 7
       }
    }
}

colvar {
    name z

    width 0.15

    lowerBoundary 3.45
    lowerWallConstant 5.0
    upperBoundary 12
    upperWallConstant 5.0

    outputValue
    outputAppliedForce

    distanceZ {
       main {
           atomNumbersRange 1-22
       }
       ref {
           atomNumbers { 27 31 36 40 44 48 53 57 61 65 69 73 78 82 86 90 94 98 102
           106 110 114 118 122 126 130 134 138 142 146 150 154 158 162 166 170 174
           178 182 186 190 194 198 202 206 210 214 218 222 226 230 234 238 242 246
           250 254 258 262 266 269 273 277 281 285 289 293 297 300 304 308 312 316
           320 323 327 331 335 338 342  }
      }
    }
}


abf {
    colvars phi psi z
    fullSamples 200
    historyFreq 500000
}
```

References

(1) Fiorin, G.; Klein, M. L.; Hénin, J. Using Collective Variables to Drive Molecular Dynamics Simulations. *Math. Probl. Eng.* **2013**, *111*, 3345–3362.

(2) Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. The Adaptive Biasing Force Method: Everything You Always Wanted to Know But Were Afraid to Ask. *J. Phys. Chem. B* **2015**, *119*, 1129–1151.

(3) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(4) Hughes, Z. E.; Tomásio, S. M.; Walsh, T. R. Efficient Simulations of the Aqueous Bio-interface of Graphitic Nanostructures with a Polarisable Model. *Nanoscale* **2014**, *6*, 5438–5448.

(5) Comer, J.; Chen, R.; Poblete, H.; Vergara-Jaque, A.; Riviere, J. E. Predicting Adsorption Affinities of Small Molecules on Carbon Nanotubes Using Molecular Dynamics Simulation. *ACS Nano* **2015**, *9*, 11761–11774.

(6) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-time-step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874.

(7) MacKerell Jr, A.; Feig, M.; Brooks III, C. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.

(8) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D. Optimization of the Additive CHARMM All-atom Protein Force Field Targeting Improved

Sampling of the Backbone ϕ, ψ and Side-chain $\chi 1$ and $\chi 2$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(9) Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.