Supporting Information for

Local anesthetic and anti-epileptic drug access and binding to a bacterial voltage-gated sodium channel.

Céline Boiteux, Igor Vorobyov, Robert J. French, Christopher French, Vladimir Yarov-Yarovoy and Toby W. Allen

Supporting Methods

Full simulation detail:

The Na_vAb protein (residues 1 to 221, x-ray crystallographic coordinates from pdb 3RVY (1)) was embedded in a bilayer of 336 (167 top and 169 bottom) dipalmitoylphosphatidylcholine (DPPC) lipids (the best studied lipid bilayer in Molecular Dynamics (MD) simulations (2)), with explicit water molecules (~20,000 molecules), 150 mM of NaCl and a high concentrations of benzocaine (BZC; Fig.S4a) or phenytoin (PHT; Fig.S4b), as described below, to form 2 simulation boxes of 125x125x76 Å containing 117,821 or 118,904 atoms respectively. Residue C217 was mutated back to the original isoleucine in all four monomers.

In order to enhance partitioning and sampling of binding sites by the drugs on the MD timescale, an initial bulk concentration of 500 mM was first tested for both drugs. At such a concentration, PHT displayed a tendency for aggregation in water, leading to a reduction to 75 mM. In both systems, rapid partitioning of the drugs into the membrane interface was observed, bringing the concentrations of the drugs in solution down to 5-10 mM after ~0.2 μs of simulation, as shown in Fig.S4c. This is the bulk drug concentration in equilibrium with the channel-membrane system, and is still considered to be high (recommended plasma concentration is 40 to 80 μM for PHT (3), and toxicity has been documented for concentrations as low as 9μM BZC (4)), but necessary for unbiased exploration of binding on the multi-μs timescale.

All systems were built and pre-equilibrated with the CHARMM program (5, 6), using the C36 lipid (2) and C22 protein parameters (7) with CMAP corrections (8), TIP3P water (9), modified ion parameters (described below) and newly parameterized BZC and PHT (see *Drugs Parameterization* section below). After 1000 (4x250) steps of steepest descent minimization, MD simulations commenced with a timestep of 1 fs and initial harmonic restraints (10 kcal/mol/$Å^2$) applied to all heavy atoms. These restraints were slowly released over 2.5 ns, followed by 5 ns of simulation without any restraints, using a timestep of 2 fs. The systems were then equilibrated for an additional 100 ns using NAMD (version 2.9) (10) and the same force field. All simulations were performed at constant pressure (1 atm) (11, 12), with fixed lateral area (125.4x125.4 $Å^2$) in order to maintain the correct area per lipid obtained from CHARMM equilibration) and constant temperature of 323K (chosen to avoid the gel phase transition of DPPC lipids), using a Nosé-Hoover thermostat (13, 14). All bonds to H atoms were maintained using the SHAKE algorithm (15). Electrostatic interactions were computed using Particle Mesh Ewald (16), with grid spacing of 1 Å and 6th order B-spline for mesh interpolation. Non-bonded pair lists were updated every 20 steps with a cutoff distance of 15 Å and a real space cutoff of 12 Å with energy switch (switching distance of 10 Å).

Anton software version 2.12.1 from D. E. Shaw Research was used for production runs using the purpose-built Anton supercomputer (17). These simulations were carried out using tetragonal

periodic boundary conditions in the NPT ensemble at 323 K, a 2 fs time step with non-bonded long-range interactions computed every 6 fs using the RESPA multiple time step algorithm (18). The multi-integrator (multigrator) algorithm (19) developed in-house by D. E. Shaw Research was used for temperature and semi-isotropic pressure coupling whereas their Gaussian split Ewald (GSE) method (20) was used for handling long-range electrostatic interactions with Gaussian RMS width parameters σ and σ_s (used for Ewald splitting and charge spreading onto the grid/force interpolation from the grid, respectively), the grid size and spreading radius optimized for each simulation using guesser scripts for an initial structure. For BZC and PHT system production runs, 12.36 and 12.32 Å non-bonded interaction cutoffs were used, optimized by Anton simulation scripts for these systems. Non-bonded interaction tapering (switching or shifting) was not used, as it was not available on Anton. A long-range Lennard-Jones (LJ) correction (beyond cutoff) was not used for those systems either as was suggested for C36 lipid force field (2). A detailed description of the simulation methodology employed in this study can be found at: http://www.deshawresearch.com/downloads/download_desmond.cgi/Desmond_Users_Guide-0.5.3.pdf.

Modified LJ parameters were used to describe the interactions between Na^+ and carbonyl oxygens of residues 175-178 to reproduce their correct free energies of solvation in a protein backbone mimetic, N-methylacetamide (21-23). In Anton production runs modified Lennard-Jones parameters for interactions between Na^+ and aspartate/glutamate carboxyl oxygen atoms, phosphate and carbonyl lipid oxygen atoms were used as was suggested in a recent study (24) to prevent ion overbinding and provide good agreement with experimental osmotic pressure data for aqueous salt solutions as well as electrophoresis data for lipid vesicles. In addition, in production runs, modified Na^+ Lennard-Jones parameters (25) were used for other Na^+ interactions, in order to correctly reproduce experimental free energy difference between bulk aqueous solvation of Na^+ and K^+. Standard CHARMM LJ parameters for Cl^- were used (26).

The Na_vAb/PHT production simulation ran for 2,843 ns following on from the NAMD equilibrated structure (after 47 ns) totaling 2.9 μs with trajectory frames saved every 0.12 ns. For Na_vAb/BZC, ~150 ns of equilibration with NAMD, were extended by a further 301 ns using Anton without latest LJ modification for Na^+ interactions (except for ones with carbonyl oxygen for Na_vAb selectivity filter residues as described below) and then followed by a 1,404 ns production run with all the LJ modifications included, thus totaling ~ 1.9 μs of simulation. Trajectory frames were saved every 0.12 ns as for the Na_vAb/PHT simulation.

Simulations for membrane partitioning:

Simulations of membrane partitioning were performed with dimyristoylphosphatidylcholine (DMPC) bilayers in 0.15 M aqueous NaCl solution at 313 K, chosen to ensure a liquid crystalline phase at the same temperature used experimentally for BZC (27). The systems were simulated with tetragonal periodic boundary conditions (PBC), with 128 lipids, 5870 waters, 15 Na^+ and 15 Cl^-, one drug molecule, with 32,767 or 32,775 atoms in total for BZC or PHT respectively, built using standard procedures as described previously (28).

Umbrella Sampling (29) (US) for BZC and PHT membrane partitioning involved 61 independent simulations (windows) with 1 Å resolution, i.e. from -30 to 30 Å, with the drug center of mass (COM) held with respect to membrane COM near each US window position by a 2.5 kcal/mol/Å^2 force constant. The lateral distance of BZC or PHT from the *z* axis as well as the lipid bilayer COM

along the z axis were constrained using cylindrical and planar constraints of 5 kcal/mol/Å^2 to prevent drifting and thus assist simulation analysis without affecting free energy profiles. Potentials of mean force (PMF) were calculated using weighted histogram analysis method (WHAM) (30). All US simulations were performed in the *NPT* ensemble at a constant pressure of 1 atm using the Langevin piston barostat (12), and constant temperature of 313K using a Nosé-Hoover thermostat(13, 14). All bonds to hydrogen atoms were maintained using the SHAKE algorithm (15) allowing to use a timestep of 2 fs. Electrostatic interactions were computed using Particle Mesh Ewald (16), with grid spacing of 1 Å and 6th order B-spline for mesh interpolation. Non-bonded pair lists were kept up to 16 Å and updated heuristically. And a real space non-bonded cutoff distance of 12 Å was used with atom-based force switch algorithm starting at 8 Å.

For BZC we ran between 11.3 and 17.3 ns per window (16.6 ns on average), while for PHT we ran between 11.5 and 18 ns per window (15.4 ns on average), but extended to 25ns for the central windows (z = -2 to z = 2 Å) to ensure convergence. Simulations were run using the CHARMM program (5, 6) and CHARMM36 parameters for lipids (2). Based on PMF convergence (Fig.S2c&d) for each US window we discarded the first 5 ns for BZC and 7 ns for PHT simulations. Symmetrized PMF results are shown in Fig.S2a, compared with similar results from unbiased simulations (where drugs were in the membrane and away from the protein) in Fig.S2b.

Membrane Partitioning:

PMFs for BZC and PHT are shown in Fig.S2. The PMF minima are -3.3±0.4 kcal/mol at |z|~8.6±0.1 Å for BZC and -4.13±0.09 kcal/mol at |z|~9.59±0.03 Å for PHT. At the membrane center, PHT encounters a barrier of 0.66±0.07 kcal/mol (with respect to bulk aqueous solution) and BZC a plateau of -0.6±0.1 kcal/mol. Partition coefficients were calculated as (31):

$$P_{\mathrm{X}} = \frac{1}{(z_2 - z_1)} \int_{z_1}^{z_2} e^{-\frac{\{W(z) - W(z_1)\}}{k_B T}} dz$$

where $W(z)$ is the PMF, z_1 and z_2 are points in aqueous solution on opposite sides of the membrane, k_{B} is Boltzmann constant, and T is the absolute temperature. Partitioning free energies were calculated as $\Delta G_{\mathrm{X}} = -k_{\mathrm{B}} T \ln P_{\mathrm{X}}$. Error bars were estimated from PMFs by propagation of errors. Calculated partitioning coefficients (and free energies) are P_{b} = 38.7±6.3 (ΔG_{b}= -2.28±0.10 kcal/mol) and P_{p} = 139±25 (ΔG_{p} = -3.07±0.11 kcal/mol) for BZC and PHT, respectively. The experimental values for P_{b} in DMPC from two different studies are 202±6 (27) at 313 K and ~186 (32) at 310 K, being several times greater, but with a difference in ΔG_{b} of only ~1 kcal/mol. This suggests weaker partitioning into the membrane with our model, which might be due to differences between simulations and experimental setup or possibly a slightly overestimated dipole moment (Table SA4, Appendix SA1), but within acceptable energetic differences. Interestingly, the gas-phase dipole in our model (3.13 D) is only marginally larger than QM MP2 one (3.02 D), though can typically be ~20% larger to get right condensed phase properties (7, 33). Moreover, BZC interaction energies with water are in good agreement with scaled QM values (Table SA6). There is a similar discrepancy between PHT calculated (139 ± 25) and experimental (657 ± 33) (34) values, although quantitative comparison in this case might be problematic as the experiments were done using a different membrane (egg phosphatidylcholine) and temperature (298 K).

Bulk Solvent Free Energy Perturbation calculations:

Partitioning free energies of BZC and PHT between TIP3P water, and hexane (Hex) or cyclohexane (cHex) were calculated by the free energy perturbation (FEP) approach (35, 36), using the CHARMM27 force field with new parameters for BZC and PHT. Water-Hex and water-cHex partitioning free energies were calculated as a difference in Hex or cHex and water solvation free energies. Electrostatic and dispersive contributions were computed using the standard linear coupling scheme (with a coupling parameter ranging from 0 to 1 in increments of 0.1), whereas the repulsive term was transformed into a soft-core potential and calculated in multiple stages (for staging parameter values of 0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0), as described previously (36). Systems of 424 Hex/cHex or 2704 water molecules were placed in cubic boxes with BZC or PHT in the center constrained by a weak (0.5 kcal/mol/Å^2) harmonic restraint acting on a solute center of mass. All systems used for FEP calculations were first subjected to initial 200 step steepest descent minimization (to eliminate bad contacts) and then 6 ns of MD equilibration. For each system 3 FEP simulations were performed starting from different equilibrated structures using 40 independent 600 ps runs for each value of a coupling or staging parameter with the 1st 100 ps of each run treated as equilibration (thus totaling 20 ns of production runs per simulation). All equilibration and FEP MD simulations were performed in the *NPT* ensemble at a constant pressure of 1 atm using the Langevin piston barostat (12), a constant temperature of 298K using a Nosé-Hoover thermostat (13, 14), PME for electrostatics (16), and cubic periodic boundary conditions.

Calculated free energies were corrected for finite LJ cutoffs in FEP simulations using average differences in interaction energies between a solute and a solvent from a number of atomic coordinate sets from the FEP simulations. Interaction energies were computed using non-bond cutoffs used in the FEP simulations (LJ energy switch from 10 to 12 Å and no long range correction) as reference values and were compared to those computed with the long-range (LR) cutoff (LJ energy switch from 30 to 32 Å along with the long-range LJ correction) and those computed with the same cutoffs as used in MD membrane partition simulations (force switch from 8 to 12 and no long range correction, as suggested for use with the C36 lipid force field (2). The partitioning free energy values corrected in such a way were labeled FEP (LR) and FEP (raw), respectively (see Table S1).

The calculated water-Hex and water-cHex partitioning free energies for BZC with long-range LJ corrections are -0.01±0.17 kcal/mol and -0.57±0.23 kcal/mol, in agreement within the error with experimental values (-0.26±0.25 kcal/mol (37) and -0.38±0.01 kcal/mol (27) for Hex and cHex, respectively). The calculated water-Hex for PHT is 0.23±0.19 kcal/mol, also in agreement with an experimental value (0.41 (38)) (see Table S1). No experimental data were found for PHT water-cHex partitioning. The use of the same LJ cutoffs used in membrane partitioning simulations results in a slightly larger difference with experimental data (compare FEP (raw) and experimental values in Table S1), although still within the reported errors.

Drug Transition rates through the membrane:

We used Kramer's transition rate theory to compute the transition rates of the drugs through the DMPC lipid bilayer, using trajectories obtained during US calculations. It can be shown that this rate can be approximated by the expression (39, 40):

$$k = \frac{D(z_{barrier})}{2\pi kT}\left[-W''(z_{barrier})W''(z_{well})\right]^{1/2} e^{-\Delta W_{barrier}/kT}$$

where $D(z_{barrier})$ is the diffusion constant measured near the top of the barrier in the z-direction, $W''(z)$ the second derivative of the PMF estimated at the barrier and wells by least squares harmonic fits, and $\Delta W_{barrier}$ is the barrier height for the transition through the membrane. For BZC, $\Delta W_{barrier}$ =

2.66±0.45 kcal/mol, we estimated the curvatures of the wells at the lipid interface and of the barrier at the bilayer center to be 0.088±0.026 and −0.013±0.020 kcal/mol/A^2, respectively. For PHT $\Delta W_{barrier}$ = 4.80±0.12 kcal/mol, with wells and barriers having curvatures 0.084±0.007 and −0.16 ± 0.09 kcal/mol/A^2, respectively.

To estimate the 1D diffusion constant in the z direction near the barrier (z~0), $D(z_{\text{barrier}})$, we analyzed the corresponding central US windows with Hummer's method (41):

$$D(z_i) = \frac{\left\langle \delta z^2 \right\rangle_i}{\tau_i}$$

where $\left\langle \delta z^2 \right\rangle_i$ is the mean square deviation from the average position in US window i, and τ_i is the position correlation time for umbrella sampling window i:

$$\tau_i = \lim_{s\to 0} \tau_i(s) = \lim_{s\to 0} \frac{\hat{C}_z(s;z_i)}{\left\langle \delta z^2 \right\rangle_i} = \lim_{s\to 0} \frac{\int_0^\infty e^{-st} \left\langle \delta z(t)\delta z(0) \right\rangle_i \mathrm{d}t}{\left\langle \delta z^2 \right\rangle_i}$$

$\hat{C}_z(s;z_i)$ is the Laplace transform of the position autocorrelation function $C_z(t;z_i)$:

$$\hat{C}_z(s;z_i) = \int_0^\infty e^{-st} C_z(t;z_i)\,\mathrm{d}t$$

where $C_z(t;z_i) = \left\langle \delta z(t)\delta z(0) \right\rangle_i$, s is inverse time and $\delta z = z - \left\langle z \right\rangle_i$ is the drug COM displacement.

Values of $\tau_i(s)$ were calculated at s values 0.01, 0.02, …, 0.1, 0.2, …,1.0, 2.0,…, 10.0 ps^{-1}. $\tau_i(s)$ were extrapolated to s=0 by fitting the function $a/(s+b)$, where a and b are parameters, in the s range from 0.02 to 1.00 ps^{-1}. See our previous study (42) for more details.

The value of the diffusion constant close to the barrier in the z-direction was determined to be 0.015 and 0.0027 A^2/ps for z = –0.12±0.18 Å and –0.11±0.40 Å, for BZC and PHT, respectively (Fig.S2d). The estimated transition rates through the membrane are 1.77±0.50 μs^{-1} for BZC and 35±3 ms^{-1} for PHT, explaining significantly less events with PHT in the middle of the membrane.

Drug parameterization:

BZC and PHT topology and parameters are not available in CHARMM biomolecular (7) or generalized (CGENFF) (33) force fields (although interfacial lipid interactions of BZC have been studied with GROMACS and QM/MM (43, 44)) and were developed *de novo* using the following methodology, strictly following CHARMM force field standards. Initial guesses for partial atomic charges and other force field parameters for these molecules were obtained using CGENFF program (45, 46) version 0.9.7 beta available at https://www.paramchem.org/. Molecular structures in mol2 format from the ZINC database (47) (ZINC12358719 for BZC and ZINC02510358 for PHT) were used to generate topology and parameters for CGENFF program. For both drugs we only considered neutral forms that predominate at physiological conditions (pK_a=2.51 for BZC NH_3^+/NH_2 equilibrium (48) and pK_a=8.06–8.33 or PHT NH/N$^-$ equilibrium (49, 50)). There is also lactam/lactim (imide/imine) tautomerism for PHT with di-lactam, two lactam-lactim and di-lactim tautomers possible (Appendix SA1). Our QM calculations at HF/6-31G(d) and MP2/6-31G(d) levels in both gas phase and in implicit solvents of different polarity indicate that the di-lactam form is dominant in all cases (>10 kcal/mol lower in energy; Table SA1 in Appendix SA1) consistent with related compounds (50). Consequently, neutral di-lactam PHT and BZC structures were used.

After generating initial topology and parameters for BZC and PHT, we performed validation and optimization using QM as target data, following the suggested CGENFF force field methodology (33). Parameter optimization and validation was focused on initial force field parameters, which were not present in CGENFF and were obtained from existing parameters with high penalty scores (i.e. where chemical analogy was poor) (46). MP2/6-31G(d) molecular dipole magnitude and orientation (Tables SA4 and SA5, Appendix SA1) as well as scaled HF/6-31G(d) orientations with water (Tables SA6 and SA7, Appendix SA1) were used for partial atomic charge optimization (Tables SA2 and SA3, Appendix SA1) for compatibility with biomolecular CHARMM force field (7). The gas-phase MP2/6-31G(d) dipole, along with HF/6-31G(d) interaction energies, have to be overestimated by CHARMM (ideally by ~16%) to account for polarization in aqueous media (7, 33). For PHT we found the CHARMM dipole moment increased by ~17% compared to the QM value (Table SA5, Appendix SA1) and interaction energies with water were in good agreement with 1.16 scaled QM values (Table SA7, Appendix SA1). For BZC, the CHARMM dipole moment was only ~4% larger than the QM value (Table SA4, Appendix SA1). A further increase in CHARMM dipole moment by partial charge modification resulted in poor agreement for BZC-water interactions and was not pursued further. Internal bond and angle parameters were validated or modified based on comparison of MP2/6-31G(d) and CHARMM optimized geometries (see Tables SA8 and SA9, Appendix SA1) and scaled vibrational frequencies (Table SA10 for BZC, not performed for PHT). For bond lengths and angles, 0.01 Å and 1° differences between QM and CHARMM values were sought (see Tables SA8 and SA9, Appendix SA1). Dihedral angle parameters were fitted to reproduce MP2/6-31G(d) potential energy surfaces for rotation around a particular bond (see Figures SA1 and SA2, Appendix SA1). Parameter optimization was iterated several times until a satisfactory agreement between QM and CHARMM data was obtained. A substantial improvement over the CGENFF parameters (in terms of better agreement between CHARMM and QM geometries, vibrational frequencies, and interactions with water) was achieved. Final topology and parameters for BZC and PHT are provided in appendices SA2 and SA3 as stream files for the CHARMM program.

Dissociation constant calculations for channel binding:

For each binding site, the dissociation constant was calculated from the equilibrium PMF $W(\boldsymbol{r})$:

$$K_d^{-1} = N_A \int_{site} \frac{e^{-W(r)/kT}}{e^{-W(r_0)/kT}} dr$$

where $\boldsymbol{r_0}$ is a bulk reference. This quantity can also be expressed in terms of the unbiased 2D PMFs:

$$K_d^{-1} = N_A \Delta z \frac{e^{-W(x_m;y_m)/kT}}{e^{-W(x_0;y_0)/kT}} \iint_{site} \frac{e^{-W(x;y)/kT}}{e^{-W(x_m;y_m)/kT}} dxdy$$

where $W(x_m,y_m)$ is the free energy at a reference position in the membrane, and Δz is the thickness of the slab along z-axis used for calculation of the 2D PMFs projected on the x- and y-axis. As a consequence of the high affinity of the drugs for the membrane, the 2D PMFs show poor sampling of the transition from solvent to membrane, and do not allow for direct comparison with a reference position in the solvent. To overcome this, the dissociation constant was first obtained relative to the membrane interface (at z = 9 and 10 Å for BZC and PHT respectively), and converted to be relative

to solvent using the full sampling provided by the US membrane partitioning calculations, as included in the above expression. Results for each binding site are summarized in Tables S2 and S3 for BZC and PHT, respectively.

Supporting Tables

Table S1: Partitioning free energies of BZC and PHT in hexane (Hex) and cyclohexane (cHex) calculated experimentally[a] and by free energy perturbations with (LR)[c] and without (raw)[b] long-range Lennard-Jones corrections. [d]From reference (37). [e]From reference (27). [f]From reference (38). All values are in kcal/mol and errors in calculated free energies are standard deviations calculated from 3 independent simulations.

Method	*Water → Hex*	*Water → cHex*
	Benzocaine	
Experimental[a]	*-0.26 ± 0.25[d]*	*-0.38 ± 0.01[e]*
FEP (raw)[b]	*0.20 ± 0.17*	*0.00 ± 0.23*
FEP (LR)[c]	*-0.01 ± 0.17*	*-0.57 ± 0.23*
	Phenytoin	
Experimental[a]	*0.41[f]*	*N/A*
FEP (raw)[b]	*0.57 ± 0.19*	*0.12 ± 0.19*
FEP (LR)[c]	*0.23 ± 0.19*	*-0.75 ± 0.19*

Table S2: BZC binding sites. The $^+$ sign following the segment name in brackets indicates that the residue belongs to the next subunit.

Binding site	*Location*	*Residues (Segment)*	*Interactions*	K_d *(μM)*
A_b	*Fenestrations (S5-S6)*	***F203** (S6)*	*π-stacking*	*141±28*
B_b	*Gate (S6)*	*V213-I217 (S6, 4 subunits)*	*Hydrophobic*	*407±54*
C_b	*P-loop, between 2 subunits*	*F167&Y168(P1)-W195(S6$^+$)*	*π-stacking*	*460±62*
D_b	*High interface VSD – S5*	*F37(S1-S2)-Y142(S5$^+$)-F167(P1$^+$)*	*π-stacking*	*811±77*
E_b	*P-loop, within 1 subunit*	*F144(S5)-F198&F201(S6)*	*π-stacking*	*4790 ±830*
F_b	*Low interface VSD – S5/S6*	*F107(S4)-F140(S5$^+$)-F207(S6$^+$)*	*π-stacking*	*2240±370*
G_b	*High VSD*	*N49(S2)-R102(S4)*	*H-bond/cation-π*	*1400±150*

Table S3: PHT binding sites.

Binding site	*Location*	*Residues (Segment)*	*Interactions*	K_d *(μM)*
A_p	*Gate S6, S4-S5*	***N211**(S6)-D219(S6$^+$)*	*H-bond*	*73±19*
B_p	*Low interface VSD – S5*	*V113(S4)-P114(S4-S5)-L131&L136(S5$^+$)*	*Hydrophobic*	*48±26*
C_p	*S5, S4-S5*	*I127(S4-S5)-L131&I134(S5)*	*Hydrophobic*	*294±84*
D_p	*Low VSD*	*F14&I18 (S1)-R108,L109&V113(S4)*	*Hydrophobic/H-bond*	*19±8*
E_p	*High VSD*	*N49&I53(S2)-I97(S3)-L98(S4)*	*Hydrophobic/H-bond*	*25±12*

Supporting Figures

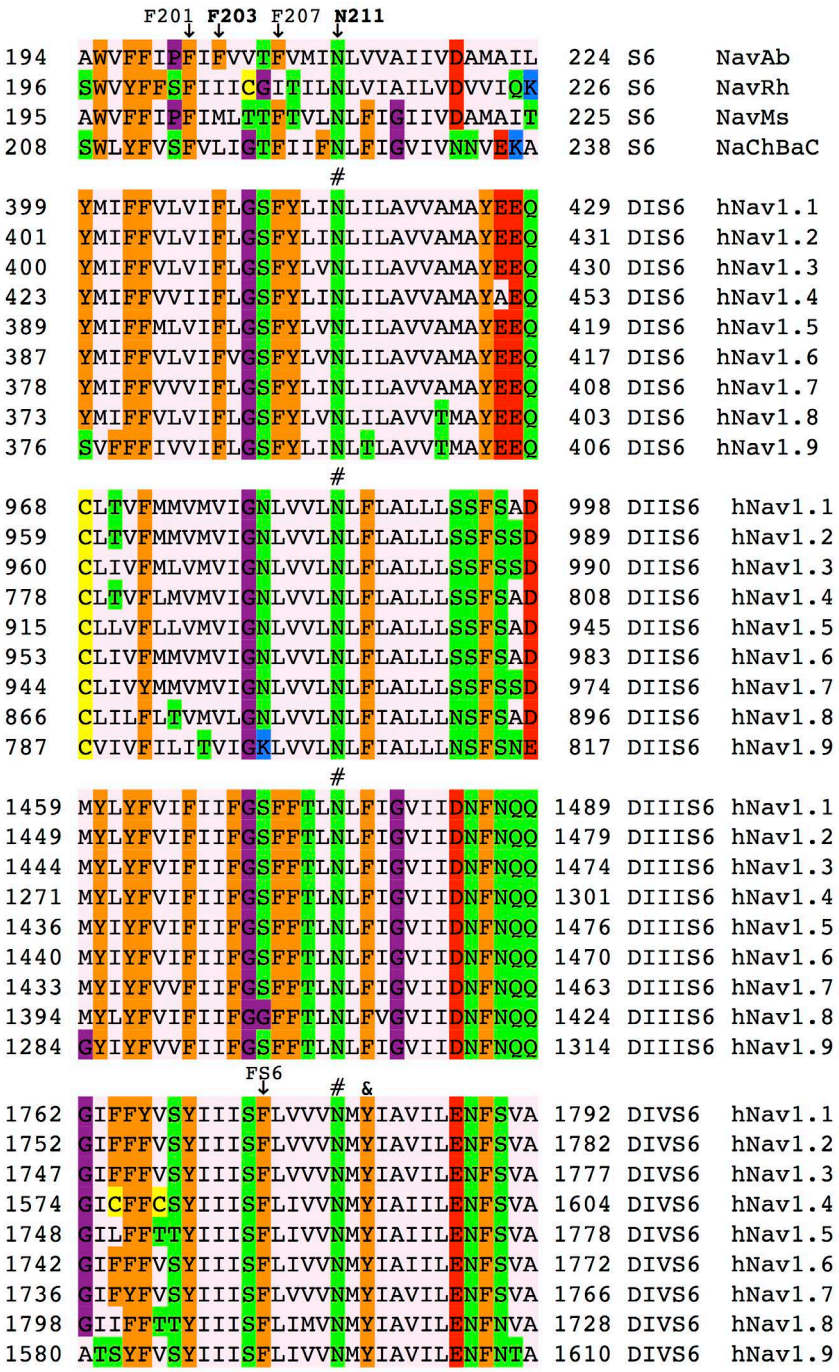

Figure S1. Sequence alignment of segments S6 of bacterial channels Na_vAb, Na_vRh, Na_vMs and NaChBaC, and all 4 domains (DI to DIV) of human $Na_v1.1$ to $Na_v1.9$. Positions of F201, F203 F207 and N211 in Na_vAb are indicated. Positions of conserved residues F1764 (FS6), Y1771 (&) and N418 (DI), N426 (DII), N1466 (DIII) and N1769 (DIV) (#), shown to be involved in drug binding in mammalian channels, are also marked ($Na_v1.2$ numbering). Sequences were aligned with the ClustalO program available on the Uniprot website and amino acids are colored using the Zappo color scheme (hydrophobic residues in pink, aromatic in orange, negatively charged in red, positively charged in blue, hydrophilic in green, P and G in magenta and C in yellow).

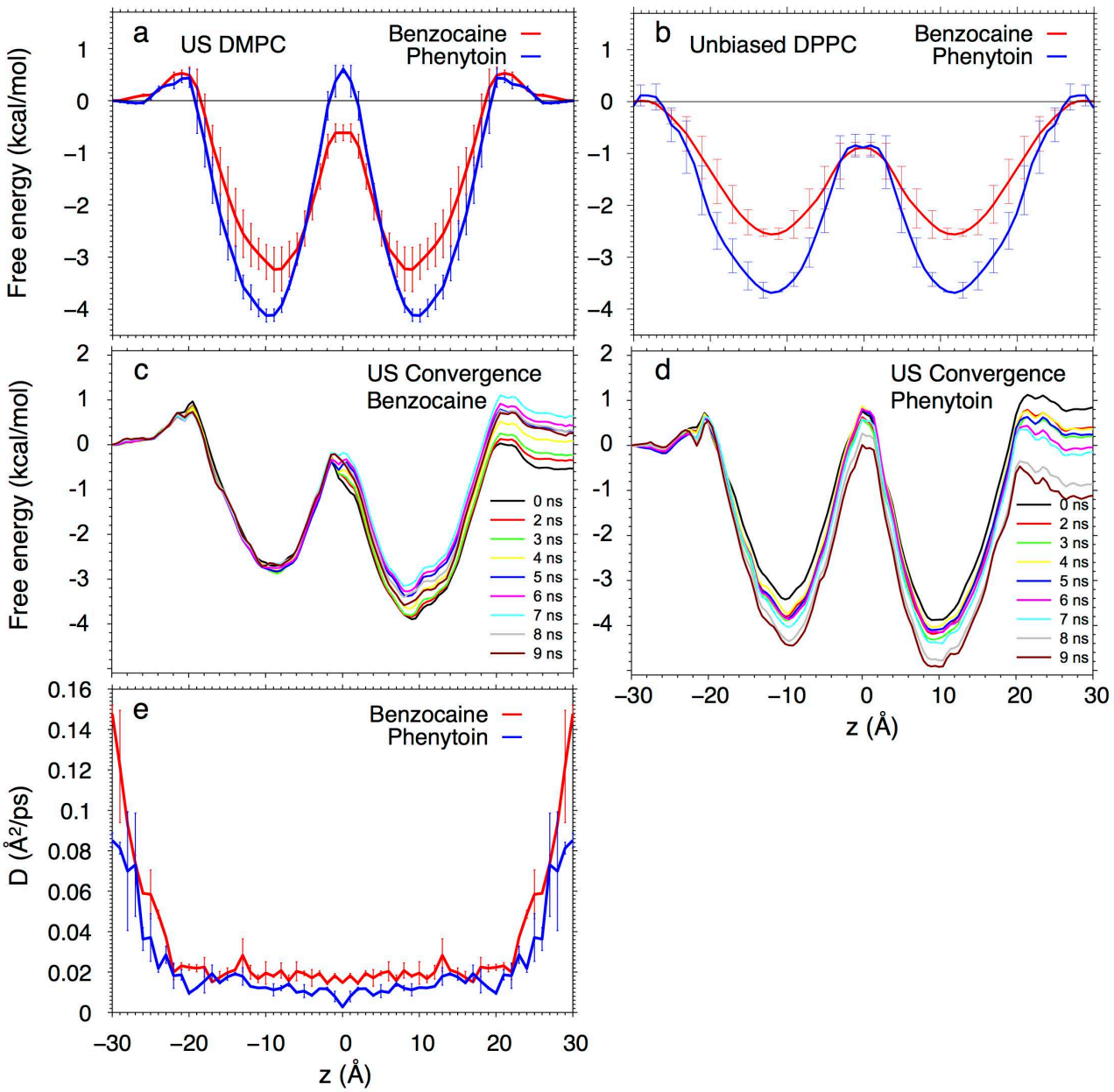

Figure S2: BZC (red) and PHT (blue) (a) PMFs across a DMPC lipid membrane hydrated by 0.15 M aqueous NaCl at 313 K (US calculations) and (b) PMFs across a DPPC lipid membrane from simulations of Na_vAb in the presence of BZC and PHT by sampling only the membrane regions away from the protein, with results consistent with panel a, but with reduced central PHT barrier. (c & d) Convergence of BZC (c) and PHT (d) US PMFs across DMPC bilayers. (e) BZC (red) and PHT (blue) diffusion coefficients across DMPC bilayers. Error bars represent asymmetry across the membrane.

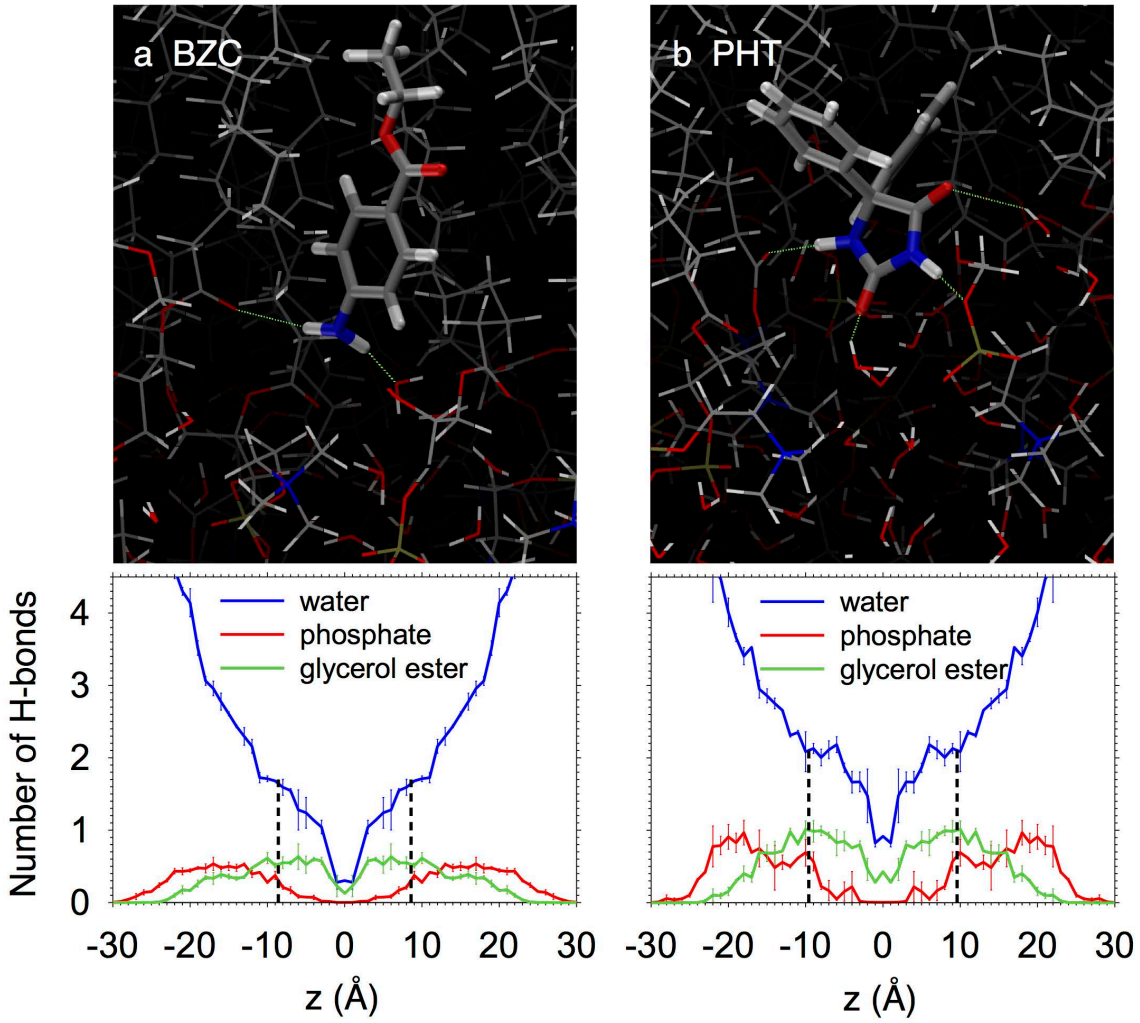

Figure S3: Interaction of BZC (a, top) and PHT (b, top) with lipids and water molecules inside the membrane near the free energy minima for each molecule. BZC and PHT molecules are shown in tube representation, lipid and water molecules in wireframe representation. C atoms are gray, H – white, O – red, N – blue, P – dark yellow. Hydrogen bonds to BZC/PHT are shown by green dotted lines. Bottom panels show mean H-bonding numbers with water, lipid phosphate and carbonyl groups as a function of drug position (H-bond D−H···A defined when distance D···A ≤ 3.8 Å and angle D−H···A ≥ 120°), with dashed black vertical lines representing minimum free energy positions for each molecule.

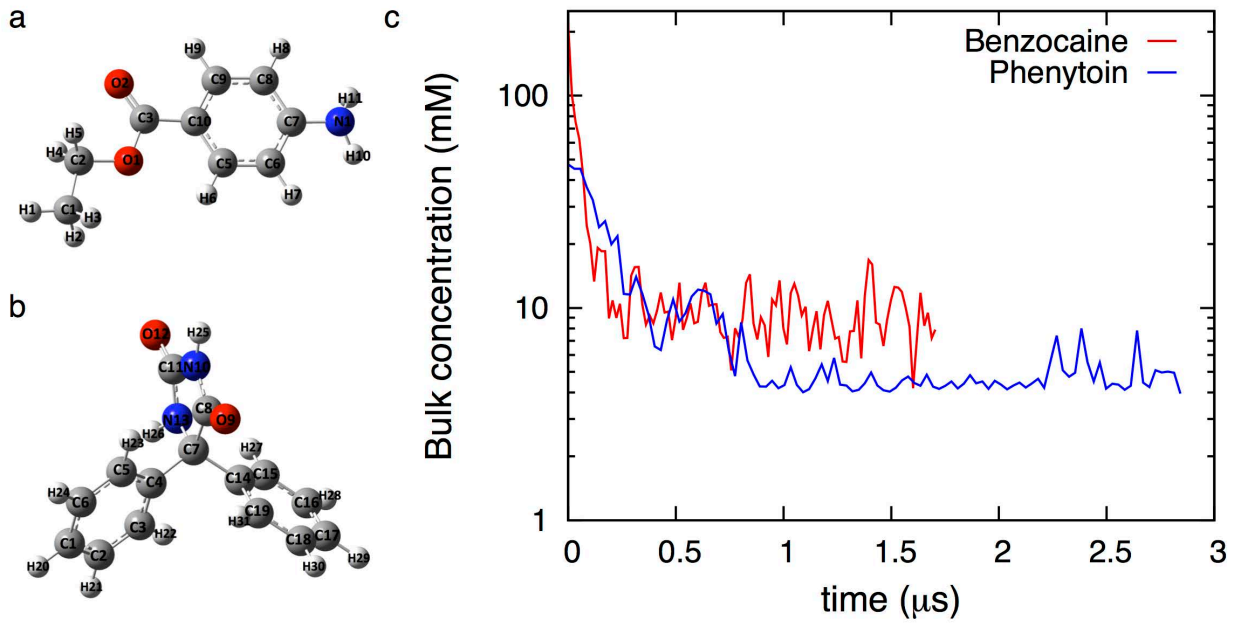

Figure S4: Ball-and-stick representations of (a) BZC and (b) PHT. (c) Bulk solvent concentrations during simulations, BZC (red) and PHT (blue).

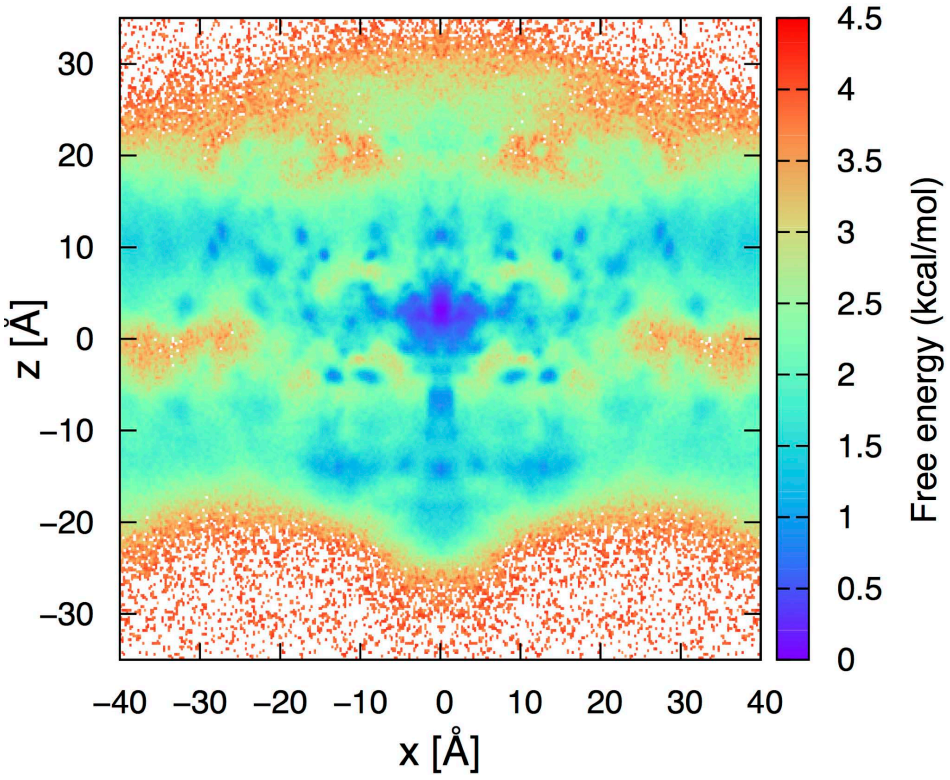

Figure S5: 2D free energy projection on x- and z- axis for BZC in the Na_vAb/BZC system, using full sampling along y.

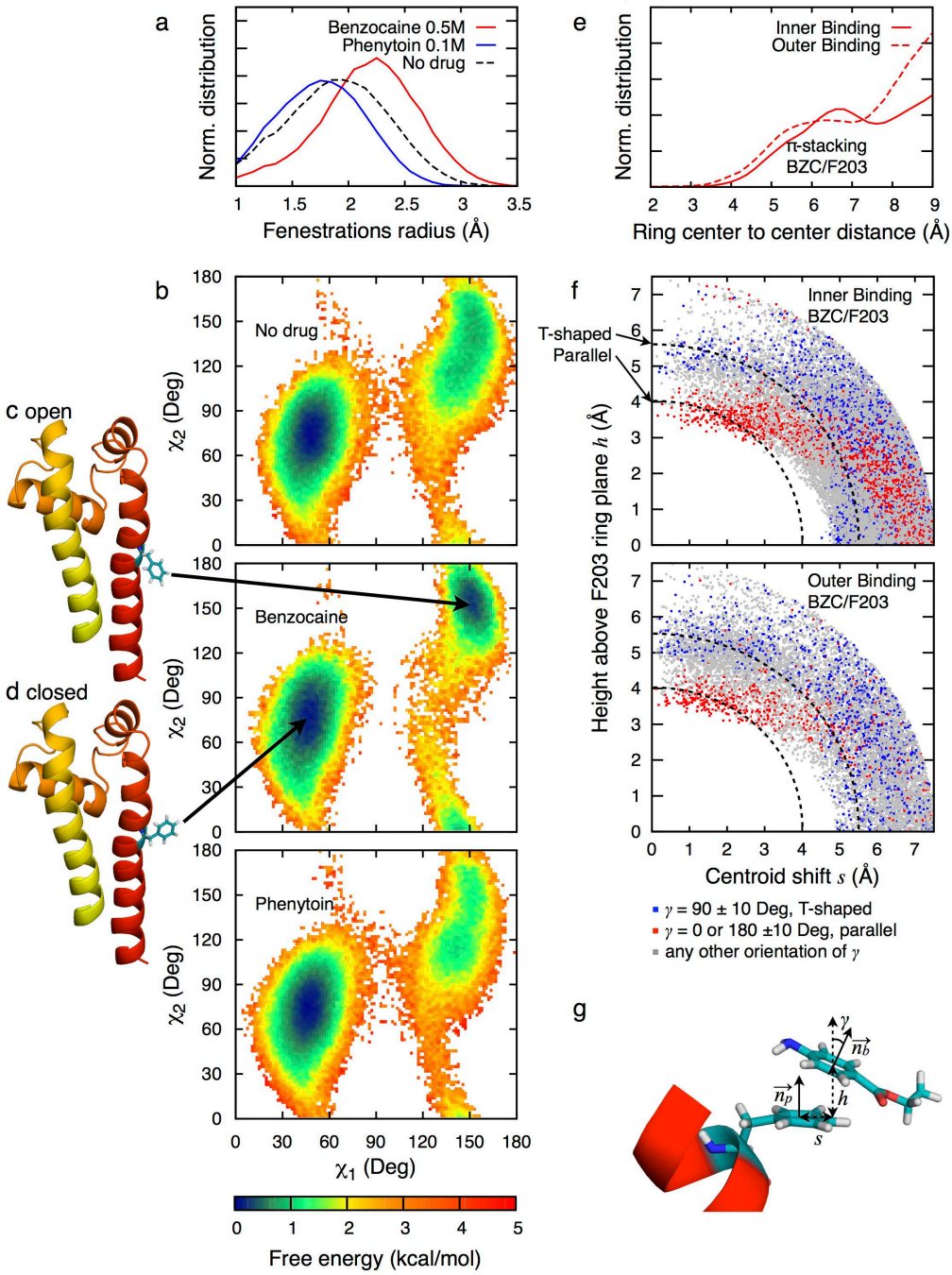

Figure S6: a) Distribution of fenestration radius in the presence of BZC (solid red) and PHT (solid blue) and in the absence of drug (dashed black), revealing an enlargement of the fenestrations in the presence of BZC. Fenestration radius was measured using CAVER 3.0 (51), using a spherical probe with minimum radius 1 Å to explore the opening in the PD (residues 130-221; aligned using backbone atoms), starting in the hydrophobic cavity (at the COM of all residues 203 and 204), using analysis of every 20th frame, with size corresponding to the radius at the bottleneck for each frame. b) Change in helix F203 rotamer free energy minima due to BZC binding. Panels c and d are insets for panel b, revealing different F203 orientations corresponding to closed and open fenestrations. e) Distribution of BZC relative to F203 from the cavity (inner binding, solid line) and membrane (outer binding, dashed line) facing sides, revealing a plateau corresponding to π-stacking region. f) Geometric characterization of the interactions compatible with π-stacking for inner (top) and outer binding (bottom) representing the distance *h* between the center of the benzocaine ring and the plane formed by the F203 ring as a function of the lateral shift *s* between the center of the ring, as described in panel (g). Dot color corresponds to the angle γ between ring normal vectors.

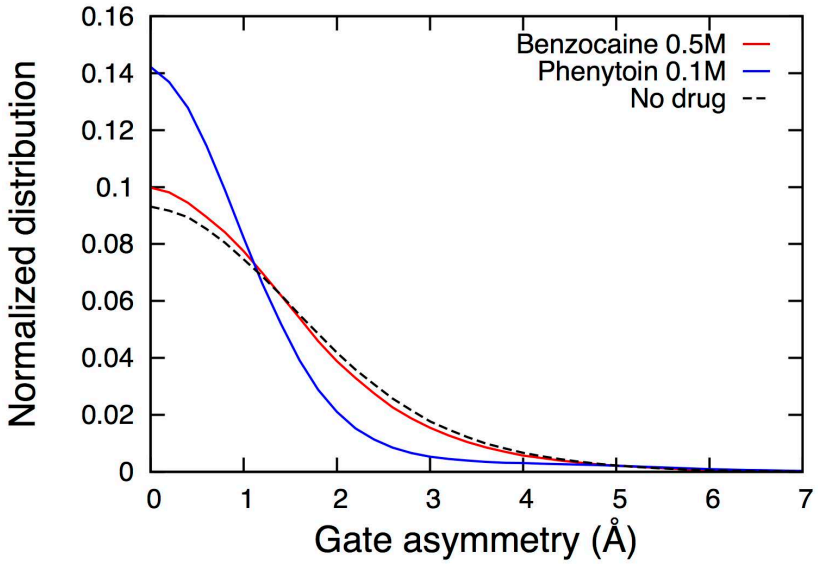

Figure S7: Na_vAb gate asymmetry for simulations in the presence of BZC (solid red) and PHT (solid blue), and absence of drugs (dashed black), revealing a drastic change due to PHT. Gate asymmetry is measured as the difference between distances separating the COM of the Cα of residues 215–218 on pairs of opposite monomers.

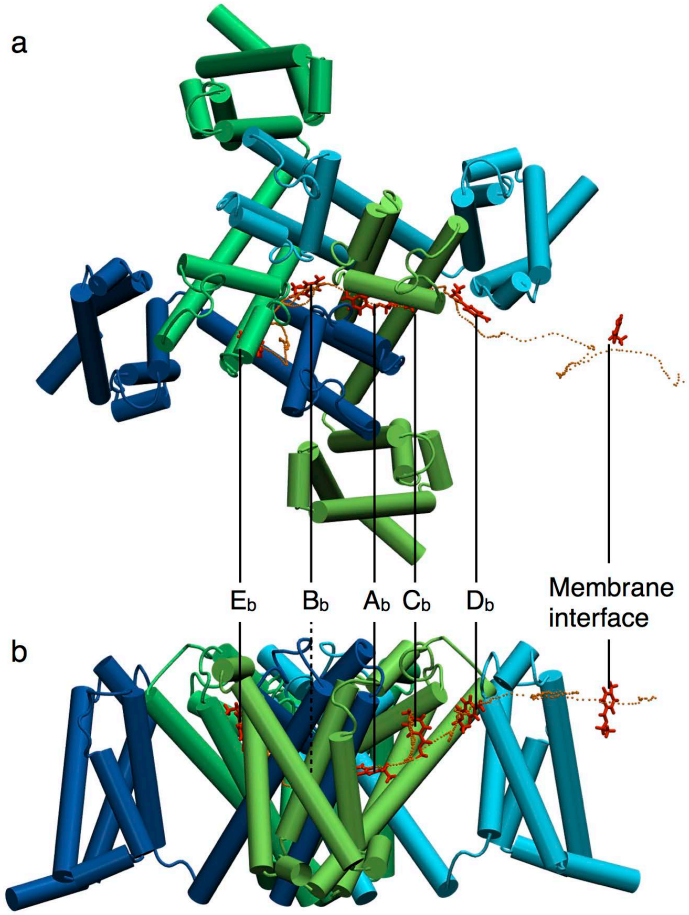

Figure S8: (A) Top-view of the channel showing the trajectory of a BZC molecule from the membrane interface to the cavity along the binding pathway from the membrane interface to site E_b via sites D_b, C_b, A_b and B_b, as indicated. (B) Corresponding side-view.

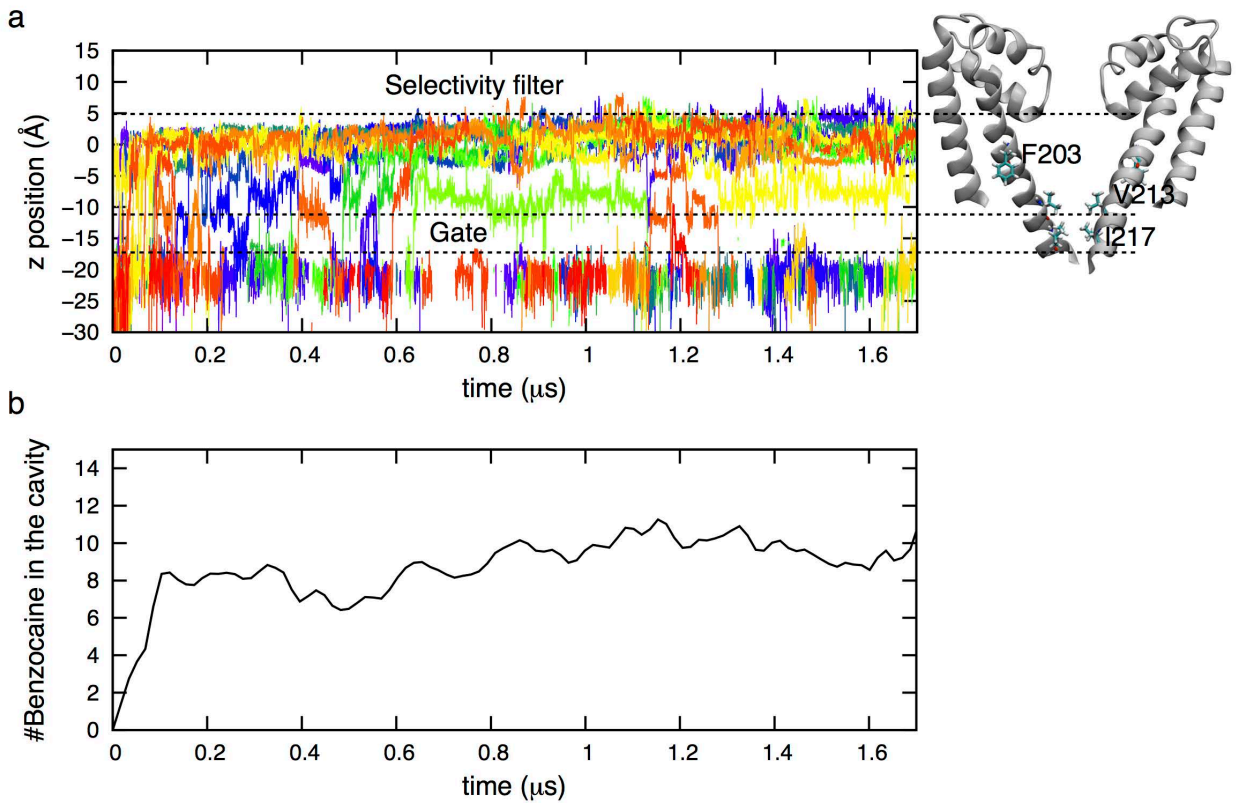

Figure S9: a) Time series showing the *z* position of the COM of BZC molecules entering the cavity, with corresponding representation of the PD with labeling of key residues F203, V213 and I217 (right). Only BZC molecules nearing or entering the pore (r<15 Å) via the gate are shown. b) Number of BZC molecules in the cavity as a function of time.

Supporting References

1. Payandeh J, Scheuer T, Zheng N, & Catterall WA (2011) The crystal structure of a voltage-gated sodium channel. *Nature* 475(7356):353-358.
2. Klauda JB*, et al.* (2010) Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J Phys Chem B* 114(23):7830-7843.
3. Lascelles PT, Kocen RS, & Reynolds EH (1970) The distribution of plasma phenytoin levels in epileptic patients. *J Neurol Neurosurg Psychiat* 33:501-505.
4. Logan BK & Gordon AM (2005) Death of an infant involving benzocaine. *J Forensic Sci* 50(6):1486-1488.
5. Brooks B & Karplus M (1983) Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsine inhibitor. *Proc Natl Acad Sci USA* 80:6571-6575.
6. Brooks BR*, et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545-1614.
7. MacKerell Jr AD*, et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586-3616.
8. Mackerell AD, Jr., Feig M, & Brooks CL, 3rd (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25(11):1400-1415.
9. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, & Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79(2):926-935.
10. Phillips JC*, et al.* (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781-1802.
11. Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72(4):2384.
12. Feller SE, Zhang YH, Pastor RW, & Brooks BR (1995) Constant Pressure Molecular Dynamics Simulation: the Langevin Piston Method. *J Chem Phys* 103(11):4613-4621.
13. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 81:511.
14. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 31(3):1695-1697.
15. Ryckaert J-P, Ciccotti G, & Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys* 23(3):327-341.
16. Darden T, York D, & Pedersen L (1993) Particle mesh Ewald: An N log (N) method for Ewald sums in large systems. *J Chem Phys* 98:10089.
17. Shaw DE*, et al.* (2008) Anton, a special-purpose machine for molecular dynamics simulation. *Comm of the ACM* 51(7):91-97.
18. Tuckerman M, Berne BJ, & Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97(3):1990.
19. Lippert RA*, et al.* (2013) Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *J Chem Phys* 139(16):164106.
20. Shan Y, Klepeis JL, Eastwood MP, Dror RO, & Shaw DE (2005) Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J Chem Phys* 122(5):54101.
21. Berneche S & Roux B (2001) Energetics of ion conduction through the K+ channel. *Nature* 414:73-77.
22. Noskov SY, Berneche S, & Roux B (2004) Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* 431(7010):830-834.
23. Allen TW, Andersen OS, & Roux B (2006) Molecular dynamics - potential of mean force calculations as a tool for understanding ion permeation and selectivity in narrow channels. *Biophys Chem* 124(3):251-267.
24. Venable RM, Luo Y, Gawrisch K, Roux B, & Pastor RW (2013) Simulations of Anionic Lipid Membranes: Development of Interaction-Specific Ion Parameters and Validation Using NMR Data. *J Phys Chem B* 117(35):10183-10192.
25. Noskov SY & Roux B (2008) Control of ion selectivity in LeuT: two Na+ binding sites with two different mechanisms. *J Mol Biol* 377(3):804-818.
26. Beglov D & Roux B (1994) Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J Chem Phys* 100(12):9050-9063.
27. Ávila CM & Martínez F (2003) Thermodynamics of partitioning of benzocaine in some organic solvent/buffer and liposome systems. *Chem Pharm Bull* 51(3):237-240.
28. Li L, Vorobyov I, MacKerell AD, & Allen TW (2008) Is arginine charged in a membrane? *Biophys J* 94(2):L11-13.
29. Torrie GM & Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 23(2):187-199.

30. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, & Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13(8):1011-1021.
31. Vorobyov I, Bennett WFD, Tieleman DP, Allen TW, & Noskov S (2012) The Role of Atomic Polarization in the Thermodynamics of Chloroform Partitioning to Lipid Bilayers. *J Chem Theory Comput* 8(2):618-628.
32. Lukacova V*, et al.* (2007) Drug-membrane interactions studied in phospholipid monolayers adsorbed on nonporous alkylated microspheres. *J Biomol Screen* 12(2):186-202.
33. Vanommeslaeghe K*, et al.* (2010) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem* 31(4):671-690.
34. Pang KY, Braswell LM, Chang L, Sommer TJ, & Miller KW (1980) The perturbation of lipid bilayers by general anesthetics: a quantitative test of the disordered lipid hypothesis. *Mol Pharmacol* 18(1):84-90.
35. Kollman P (1993) Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. *Chemical Reviews* 93(7):2395-2417.
36. Deng Y & Roux B (2004) Hydration of Amino Acid Side Chains: Nonpolar and Electrostatic Contributions Calculated from Staged Molecular Dynamics Free Energy Simulations with Explicit Water Molecules. *J Phys Chem B* 108(42):16567-16576.
37. Natesan S*, et al.* (2013) Structural determinants of drug partitioning in n-hexadecane/water system. *J Chem Inf Model* 53(6):1424-1435.
38. Wohnsland F & Faller B (2001) High-Throughput Permeability pH Profile and High-Throughput Alkane/Water logP with Artificial Membranes. *J Med Chem* 44(6):923-930.
39. Crouzy S, Woolf TB, & Roux B (1994) A molecular dynamics study of gating in dioxolane-linked gramicidin A channels. *Biophys J* 67(4):1370-1386.
40. Allen TW, Andersen OS, & Roux B (2003) Structure of gramicidin a in a lipid bilayer environment determined using molecular dynamics simulations and solid-state NMR data. *J Am Chem Soc* 125(32):9868-9877.
41. Hummer G (2005) Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J Phys* 7(1):34.
42. Vorobyov I*, et al.* (2014) Ion-induced defect permeation of lipid membranes. *Biophys J* 106(3):586-597.
43. Bernardi RC*, et al.* (2009) Molecular dynamics study of biomembrane/local anesthetics interactions. *Molecular Physics* 107(14):1437-1443.
44. Bernardi RC & Pascutti PG (2012) Hybrid QM/MM Molecular Dynamics Study of Benzocaine in a Membrane Environment: How Does a Quantum Mechanical Treatment of Both Anesthetic and Lipids Affect Their Interaction. *J Chem Theo Comput* 8(7):2197-2203.
45. Vanommeslaeghe K & MacKerell Jr AD (2012) Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* 52(12):3144-3154.
46. Vanommeslaeghe K, Raman EP, & MacKerell Jr AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* 52(12):3155-3168.
47. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, & Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 52(7):1757-1768.
48. Perrin DD (1965) *Dissociation constants of organic bases in aqueous solution* (Butterworths, London).
49. Schwartz PA, Rhodes CT, & Cooper JW (1977) Solubility and ionization characteristics of phenytoin. *J Pharm Sci* 66(7):994-997.
50. Verdolino V, Cammi R, Munk BH, & Schlegel HB (2008) Calculation of pKa values of nucleobases and the guanine oxidation products guanidinohydantoin and spiroiminodihydantoin using density functional theory and a polarizable continuum model. *J Phys Chem B* 112(51):16860-16873.
51. Chovancova E*, et al.* (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* 8(10):e1002708.
52. Tomasi J, Mennucci B, & Cammi R (2005) Quantum mechanical continuum solvation models. *Chem Rev* 105(8):2999-3093.
53. Foresman JB, Keith TA, Wiberg KB, Snoonian J, & Frisch MJ (1996) Solvent effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation on ab initio reaction field calculations. *J Chem Phys* 100(40):16098-16104.
54. Singh UC & Kollman PA (1984) An Approach to Computing Electrostatic Charges for Molecules. *J Comput Chem* 5(2):129-145.
55. Besler BH, Merz KM, & Kollman PA (1990) Atomic Charges Derived from Semiempirical Methods. *J Comput Chem* 11(4):431-439.
56. Pulay P, Fogarasi G, Pang F, & Boggs JE (1979) Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *J Am Chem Soc* 101(10):2550-2560.

Appendix SA1

Table SA1. Relative conformation energies and dipole moments of PHT tautomers

	1 di-lactam	2 lactam/lactim	3 lactim/lactam	4 di-lactim
HF/6-31G(d)//HF/6-31G(d)				
μ(vac), D	3.17	3.38	5.72	3.36
Δ*E*(vacuum), kcal/mol	0.00	28.01	19.95	37.09
μ(cHex), D	3.35	3.68	6.16	3.66
ΔG_{sol}(cHex), kcal/mol	-2.82	-4.09	-3.73	-3.25
Δ*E*(cHex), kcal/mol	0.00	26.74	19.04	36.66
μ(water), D	3.71	4.59	6.92	4.63
ΔG_{sol}(water), kcal/mol	-10.89	-14.39	-13.90	-11.93
Δ*E*(water), kcal/mol	0.00	24.51	16.94	36.05
MP2/6-31G(d)//MP2/6-31G(d)				
μ(vac), D	2.67	3.19	4.93	3.18
Δ*E*(vacuum), kcal/mol	0.00	25.38	18.96	34.33
μ(cHex), D	3.39	3.60	6.31	3.54
$\Delta G_{sol,elec}$(cHex), kcal/mol	-4.64	-7.34	-5.20	-6.87
Δ*E*(cHex), kcal/mol	0.00	22.68	18.40	32.10
μ(water), D	4.69	5.34	8.95	5.01
$\Delta G_{sol,elec}$(wat), kcal/mol	-13.38	-28.45	-15.22	-26.62
Δ*E*(wat), kcal/mol	0.00	10.31	17.12	21.10

All relative conformational energies (Δ*E*) are with respect to di-lactam tautomers. HF/6-31G(d) solvation free energies (ΔG_{sol}) and relative conformational energies were obtained using single-point energy calculations with PCM continuum solvation model of Tomasi et al. (52). MP2/6-31G(d) solvation free energies (electrostatic component only, $\Delta G_{sol,elec}$) and relative conformational energies were obtained using single-point energy calculations with IPCM model (53). cHex – cyclohexane, wat – water.

Table SA2. Partial atomic charges for BZC

Atom name	Atom type	**ESP (MK)**	**CGENFF**	**new**
C1	CG331	**-0.327**	-0.269	-0.269
C2	CG321	**0.333**	0.060	0.060
O1	OG302	**-0.431**	-0.307	-0.307
C3	CG2O2	**0.717**	0.466	0.466
O2	OG2D1	**-0.526**	-0.494	-0.494
C4	CG2R61	**-0.090**	0.086	0.086
C5	CG2R61	**-0.198**	-0.107	-0.140
C6	CG2R61	**-0.149**	-0.112	-0.140
C7	CG2R61	**0.278**	0.058	0.168
C8	CG2R61	**-0.142**	-0.112	-0.140
C9	CG2R61	**-0.214**	-0.107	-0.140
N1	NG2S3	**-0.817**	-0.834	-0.742
H1	HGA3	**0.078**	0.090	0.090
H2	HGA3	**0.093**	0.090	0.090
H3	HGA3	**0.100**	0.090	0.090
H4	HGA2	**0.000**	0.090	0.090
H5	HGA2	**0.010**	0.090	0.090
H6	HGR61	**0.153**	0.115	0.115
H7	HGR61	**0.142**	0.115	0.115
H8	HGR61	**0.142**	0.115	0.115
H9	HGR61	**0.146**	0.115	0.115
H10	HGP4	**0.350**	0.381	0.341
H11	HGP4	**0.353**	0.381	0.341

ESP(MK) are partial atomic charges obtained from fitting to MP2/6-31G(d) electrostatic potential using the Merz-Singh-Kollman scheme (54, 55). **CGENFF** are partial atomic charges generated using CGENFF program (version 0.9.7 beta). **new** are optimized partial atomic charges used in this study. See Figure S4A for atom names.

Table SA3. Partial atomic charges for PHT

Atom name	Atom type	**ESP (MK)**	**CGENFF**	**new**
C1	CG2R61	**-0.125**	-0.115	-0.115
C2	CG2R61	**-0.126**	-0.110	-0.110
C3	CG2R61	**-0.156**	-0.130	-0.155
C4	CG2R61	**0.116**	-0.003	0.121
C5	CG2R61	**-0.213**	-0.130	-0.155
C6	CG2R61	**-0.090**	-0.110	-0.110
C7	CG3C50	**0.033**	0.797	0.160
C8	CG2R53	**0.577**	-0.116	0.513
O9	OG2D1	**-0.463**	-0.498	-0.470
N10	NG2R53	**-0.618**	-0.219	-0.500
C11	CG2R53	**0.709**	0.302	0.490
O12	OG2D1	**-0.506**	-0.435	-0.450
N13	NG2R53	**-0.614**	-0.501	-0.560
C14	CG2R61	**0.204**	-0.003	0.120
C15	CG2R61	**-0.151**	-0.130	-0.155
C16	CG2R61	**-0.159**	-0.110	-0.110
C17	CG2R61	**-0.116**	-0.115	-0.115
C18	CG2R61	**-0.120**	-0.110	-0.110
C19	CG2R61	**-0.211**	-0.130	-0.155
H20	HGR61	**0.124**	0.115	0.115
H21	HGR61	**0.131**	0.115	0.115
H22	HGR61	**0.116**	0.115	0.115
H23	HGR61	**0.148**	0.115	0.115
H24	HGR61	**0.125**	0.115	0.115
H25	HGP1	**0.390**	0.372	0.372
H26	HGP1	**0.360**	0.344	0.344
H27	HGR61	**0.111**	0.115	0.115
H28	HGR61	**0.137**	0.115	0.115
H29	HGR61	**0.121**	0.115	0.115
H30	HGR61	**0.133**	0.115	0.115
H31	HGR61	**0.134**	0.115	0.115

ESP(MK) are partial atomic charges obtained from fitting to MP2/6-31G(d) electrostatic potential using the Merz-Singh-Kollman scheme (54, 55). **CGENFF** are partial atomic charges generated using CGENFF program (version 0.9.7 beta). **new** are optimized partial atomic charges used in this study. See Figure S4B for atom names.

Table SA4. Dipole moment magnitude and orientation for BZC

	QM	CGENFF	new
μ_x	**-1.709**	-0.531	-1.408
μ_y	**-2.149**	-2.197	-2.468
μ_z	**1.250**	1.199	1.315
μ_{total}	**3.017**	2.559	3.131
μ_x/μ_y	**0.795**	0.242	0.571
μ_y/μ_z	**-1.719**	-1.832	-1.878
μ_x/μ_z	**-1.367**	-0.443	-1.071

QM are dipole component/magnitude values obtained from MP2/6-31G(d) calculations. **CGENFF** are dipole component/magnitude values obtained using CGENFF program (version 0.9.7 beta). **new** are dipole component/magnitude values obtained using optimized CHARMM parameters used in this study.

Table SA5. Dipole moment magnitude and orientation for PHT

	QM	CGENFF	new
μ_x	**-0.368**	0.451	-0.072
μ_y	**-2.520**	-3.902	-2.982
μ_z	**-0.803**	-4.125	-0.904
μ_{total}	**2.670**	5.696	3.117
μ_x/μ_y	**0.146**	-0.116	0.024
μ_y/μ_z	**3.137**	0.946	3.298
μ_x/μ_z	**0.458**	-0.109	0.080

QM are dipole component/magnitude values obtained from MP2/6-31G(d) calculations. **CGENFF** are dipole component/magnitude values obtained using CGENFF program (version 0.9.7 beta). **new** are dipole component/magnitude values obtained using optimized CHARMM parameters used in this study.

Table SA6. Interactions of BZC with water in vacuum

#	Orientation	QM			new				
		R	***θ***	**1.16 *IE***	*R*	*ΔR*	*θ*	*IE*	*ΔIE*
1	C=O…HOH linear	**2.05**		**-5.72**	1.77	*-0.28*		-6.00	*-0.28*
2	C=O…HOH 120°, tow. O2	**2.66**		**-3.47**	2.69	*0.03*		-2.93	*0.54*
3	C=O…HOH 120°, tow. C4	**3.12**		**-1.41**	2.99	*-0.13*		-1.93	*-0.52*
4	C-O…HOH 180°	**3.89**		**-0.35**	2.99	*-0.90*		0.38	*0.73*
5	C-O…HOH 90°	**2.43**		**-1.14**	2.02	*-0.41*		-1.94	*-0.80*
6	C-O…HOH var. angle	**2.25**	**61.63**	**-2.48**	1.89	*-0.36*	65	-2.67	*-0.19*
7	N-H10…HOH linear out of plane	**2.31**		**-3.04**	1.95	*-0.36*		-3.94	*-0.90*
8	N-H10…HOH linear in plane	**2.18**		**-4.98**	1.91	*-0.27*		-5.17	*-0.19*
9	N-H11…HOH linear out of plane	**2.30**		**-3.12**	1.95	*-0.35*		-4.03	*-0.91*
10	N-H11…HOH linear in plane	**2.17**		**-5.07**	1.91	*-0.26*		-5.24	*-0.17*
11	H2N…HOH var. angle	**2.19**	**110.59**	**-5.20**	1.99	*-0.20*	134	-5.08	*0.12*
11a	H2N…HOH 110.59° angle	**2.19**		**-5.20**	1.96	*-0.23*		-4.65	*0.55*
13	C5-H6…HOH linear in plane	**2.37**		**-0.87**	2.67	*0.30*		0.10	*0.97*
14	C6-H7…HOH linear in plane	**2.65**		**-1.95**	2.70	*0.05*		-1.33	*0.62*
15	C8-H8…HOH linear in plane	**2.66**		**-1.93**	2.70	*0.04*		-1.40	*0.53*
16	C9-H9…HOH linear in plane	**4.30**		**-0.23**	2.78	*-1.52*		-0.44	*-0.21*
								AE	***-0.01***
								RMSE	***0.57***

QM results are from HF/6-31G(d)//MP2/6-31G(d) calculations. **new** results are from calculations using optimized CHARMM force field parameters used in this study. *R* are interaction distances in Å, *IE* are interaction energies in kcal/mol. *θ* are interaction angles in degrees. **AE** is an average error in *IE*s, **RMSE** – root mean square error in *IE*s.

Table SA6 (continued). Interactions of BZC with water in vacuum

#	Orientation	QM			CGENFF				
		R	***θ***	**1.16 *IE***	*R*	*ΔR*	*θ*	*IE*	*ΔIE*
1	C=O…HOH linear	**2.05**		**-5.72**	1.77	*-0.28*		-5.86	*-0.14*
2	C=O…HOH 120°, tow. O2	**2.66**		**-3.47**	2.71	*0.05*		-2.78	*0.68*
3	C=O…HOH 120°, tow. C4	**3.12**		**-1.41**	2.99	*-0.13*		-1.56	*-0.14*
4	C-O…HOH 180°	**3.89**		**-0.35**	2.99	*-0.90*		0.15	*0.50*
5	C-O…HOH 90°	**2.43**		**-1.14**	2.02	*-0.41*		-1.78	*-0.64*
6	C-O…HOH var. angle	**2.25**	**61.63**	**-2.48**	1.89	*-0.36*	65	-2.46	*0.02*
7	N-H10…HOH linear out of plane	**2.31**		**-3.04**	1.98	*-0.33*		-3.63	*-0.59*
8	N-H10…HOH linear in plane	**2.18**		**-4.98**	1.95	*-0.23*		-4.62	*0.36*
9	N-H11…HOH linear out of plane	**2.30**		**-3.12**	1.97	*-0.33*		-3.70	*-0.58*
10	N-H11…HOH linear in plane	**2.17**		**-5.07**	1.94	*-0.23*		-4.70	*0.38*
11	H2N…HOH var. angle	**2.19**	**110.59**	**-5.20**	1.96	*-0.23*	133	-6.22	*-1.02*
11a	H2N…HOH 110.59° angle	**2.19**		**-5.20**	1.93	*-0.26*		-5.68	*-0.48*
13	C5-H6…HOH linear in plane	**2.37**		**-0.87**	2.65	*0.28*		-0.14	*0.73*
14	C6-H7…HOH linear in plane	**2.65**		**-1.95**	2.69	*0.04*		-1.33	*0.62*
15	C8-H8…HOH linear in plane	**2.66**		**-1.93**	2.69	*0.03*		-1.41	*0.52*
16	C9-H9…HOH linear in plane	**4.30**		**-0.23**	2.71	*-1.59*		-0.77	*-0.54*
								AE	***-0.02***
								RMSE	***0.53***

QM results are from HF/6-31G(d)//MP2/6-31G(d) calculations. **CGENFF** results are from calculations using CGENFF program (version 0.9.7 beta). *R* are interaction distances in Å, *IE* are interaction energies in kcal/mol. *θ* are interaction angles in degrees. **AE** is an average error in *IE*s, **RMSE** – root mean square error in *IE*s.

Table SA7. Interactions of PHT with water in vacuum

#	Orientation	QM		CGENFF				new			
		R	**1.16 *IE***	*R*	*dR*	*IE*	ΔIE	*R*	*dR*	*IE*	ΔIE
1	C8=O9…HOH linear	**2.09**	**-5.08**	1.73	*-0.36*	-7.97	*-2.88*	1.79	*-0.30*	-5.32	*-0.24*
2	C8=O9…HOH 120 deg. tow. C7	**4.31**	**-0.94**	3.48	*-0.83*	-2.46	*-1.52*	3.78	*-0.53*	-1.09	*-0.16*
3	C8=O9…HOH 120 deg. tow. N10	**2.02**	**-6.60**	1.71	*-0.31*	-9.60	*-3.00*	1.77	*-0.25*	-6.04	*0.56*
4	N10-H…OH2 in plane	**1.95**	**-7.84**	1.79	*-0.16*	-7.67	*0.17*	1.82	*-0.13*	-7.37	*0.47*
5	N10-H…OH2 90°	**1.96**	**-7.07**	1.8	*-0.16*	-6.99	*0.08*	1.83	*-0.13*	-6.91	*0.16*
6	N10…HOH 90°	**4.37**	**-0.54**	2.83	*-1.54*	-1.48	*-0.94*	4.06	*-0.31*	-0.83	*-0.29*
7	C11=O12…HOH linear	**2.08**	**-5.26**	1.8	*-0.28*	-5.27	*-0.01*	1.80	*-0.28*	-5.38	*-0.12*
8	C11=O12…HOH 120° tow. N10	**2.00**	**-7.14**	1.76	*-0.24*	-6.41	*0.73*	1.77	*-0.23*	-6.61	*0.52*
9	C11=O12…HOH 120° tow. N13	**2.01**	**-7.23**	1.76	*-0.25*	-7.67	*-0.43*	1.78	*-0.23*	-6.71	*0.53*
#	N13-H…OH2 in plane	**2.01**	**-6.60**	1.83	*-0.18*	-8.83	*-2.23*	1.85	*-0.16*	-6.63	*-0.02*
#	N13-H…OH2 90°	**2.03**	**-6.04**	1.83	*-0.20*	-8.82	*-2.78*	1.86	*-0.17*	-6.31	*-0.27*
#	H13…HOH 90°	**5.15**	**-1.19**	4.56	*-0.59*	-0.83	*0.36*	4.50	*-0.65*	-1.07	*0.12*
						AE	***-1.04***			***AE***	***0.11***
						RMSE	***1.65***			***RMSE***	***0.33***

QM results are from HF/6-31G(d)//MP2/6-31G(d) calculations. **CGENFF** results are from calculations using CGENFF program (version 0.9.7 beta). **new** results are from calculations using optimized CHARMM force field parameters used in this study. *R* are interaction distances in Å, *IE* are interaction energies in kcal/mol. θ are interaction angles in degrees. **AE** is an average error in *IE*s, **RMSE** – root mean square error in *IE*s.

Table SA8. BZC optimized geometry in vacuum

	QM	**CGENFF**		**new**	
Bond lengths			*diff.*		*diff.*
C1-C2	**1.512**	1.528	*0.016*	1.527	*0.015*
C2-O1	**1.447**	1.441	*-0.006*	1.438	*-0.009*
O1-C3	**1.359**	1.342	*-0.017*	1.342	*-0.017*
C3-O2	**1.224**	1.222	*-0.002*	1.222	*-0.001*
C3-C4	**1.482**	1.509	*0.027*	1.511	*0.029*
C4-C5	**1.400**	1.408	*0.008*	1.409	*0.009*
C5-C6	**1.389**	1.402	*0.013*	1.403	*0.014*
C6-C7	**1.404**	1.398	*-0.006*	1.398	*-0.006*
C7-C8	**1.404**	1.398	*-0.006*	1.398	*-0.006*
C8-C9	**1.390**	1.402	*0.012*	1.403	*0.012*
C9-C4	**1.401**	1.412	*0.011*	1.413	*0.013*
C7-N1	**1.399**	1.386	*-0.013*	1.386	*-0.014*
N1-H10	**1.015**	0.998	*-0.016*	1.014	*-0.001*
N1-H11	**1.015**	0.999	*-0.016*	1.014	*0.000*
Bond angles					
C1-C2-O1	**106.5**	108.5	*1.9*	107.6	*1.0*
C2-O1-C3	**114.5**	112.0	*-2.4*	113.7	*-0.8*
O1-C3-O2	**123.1**	123.9	*0.8*	123.3	*0.3*
O2-C3-C4	**124.8**	122.8	*-2.0*	124.0	*-0.8*
C3-C4-C5	**118.0**	118.8	*0.8*	118.7	*0.7*
C3-C4-C9	**122.6**	122.6	*0.1*	122.9	*0.3*
C4-C5-C6	**120.4**	120.7	*0.3*	120.8	*0.4*
C5-C6-C7	**120.5**	120.1	*-0.4*	120.1	*-0.4*
C6-C7-C8	**118.7**	119.9	*1.1*	119.9	*1.2*
C7-C8-C9	**120.8**	120.1	*-0.6*	120.1	*-0.6*
C8-C9-C4	**120.1**	120.6	*0.5*	120.6	*0.5*
C9-C4-C5	**119.4**	118.5	*-0.9*	118.4	*-1.0*
C6-C7-N1	**120.6**	120.1	*-0.5*	120.0	*-0.5*
C8-C7-N1	**120.5**	120.1	*-0.5*	120.1	*-0.5*
Dihedral angles					
H10-N1-C7	**114.0**	111.7	*-2.3*	113.2	*-0.8*
H11-N1-C7	**114.0**	111.7	*-2.3*	113.1	*-0.9*
H10-N1-H11	**110.6**	119.9	*9.3*	110.8	*0.2*
C1-C2-O1-C3	**-180.0**	-179.9	*0.0*	-179.9	*0.1*
C2-O1-C3-O2	**-0.1**	-0.2	*-0.1*	-0.2	*-0.1*
C2-O1-C3-C4	**179.9**	179.9	*0.0*	179.9	*0.0*
O1-C3-C4-C5	**179.9**	-179.6	*0.5*	-179.7	*0.4*
O1-C3-C4-C9	**0.9**	0.3	*-0.6*	0.3	*-0.6*
H10-N1-C7-C6	**-156.6**	-158.3	*-1.8*	-152.8	*3.8*
H11-N1-C7-C6	**-28.3**	-20.9	*7.4*	-25.8	*2.5*
H10-N1-C7-C8	**28.6**	21.5	*-7.1*	26.6	*-2.0*
H11-N1-C7-C8	**156.9**	158.9	*2.0*	153.6	*-3.3*

QM results are from MP2/6-31G(d) calculations. **CGENFF** results are from calculations using CGENFF program (version 0.9.7 beta). **new** results are from calculations using optimized CHARMM force field parameters used in this study. Bond lengths are in Å, bond and dihedral angles are in degrees.

Table SA9. PHT optimized geometry in vacuum

	QM	CGENFF		new	
Bond lengths			*diff.*		*diff.*
C4-C7	**1.521**	1.561	*0.040*	1.523	*0.001*
C7-C8	**1.543**	1.548	*0.005*	1.533	*-0.010*
C8-O9	**1.220**	1.235	*0.014*	1.232	*0.012*
C8-N10	**1.376**	1.373	*-0.003*	1.365	*-0.010*
N10-H25	**1.014**	1.012	*-0.003*	1.004	*-0.010*
N10-C11	**1.407**	1.375	*-0.032*	1.372	*-0.035*
C11-O12	**1.219**	1.223	*0.004*	1.221	*0.003*
C11-N13	**1.379**	1.374	*-0.006*	1.376	*-0.003*
N13-H26	**1.015**	1.006	*-0.009*	1.012	*-0.003*
N13-C7	**1.465**	1.454	*-0.010*	1.466	*0.001*
C7-C14	**1.519**	1.561	*0.042*	1.524	*0.005*
Bond angles					
C3-C4-C7	**117.9**	118.7	*0.8*	118.0	*0.1*
C4-C7-C8	**112.0**	111.0	*-1.0*	111.8	*-0.2*
C7-C8-O9	**126.8**	131.3	*4.6*	127.9	*1.2*
C7-C8-N10	**105.9**	101.8	*-4.0*	104.6	*-1.3*
C8-N10-H25	**124.0**	121.8	*-2.2*	122.7	*-1.3*
C8-N10-C11	**113.6**	117.4	*3.7*	114.6	*1.0*
N10-C11-O12	**126.2**	128.2	*2.0*	126.6	*0.4*
N10-C11-N13	**105.3**	105.1	*-0.2*	106.6	*1.3*
C11-N13-H26	**116.9**	124.6	*7.7*	118.0	*1.1*
C11-N13-C7	**112.9**	111.8	*-1.1*	110.9	*-2.0*
N13-C7-C14	**112.8**	116.0	*3.1*	116.3	*3.4*
N13-C7-C4	**110.2**	109.7	*-0.6*	109.8	*-0.5*
C14-C7-C4	**111.6**	110.9	*-0.7*	110.2	*-1.4*
C7-C14-C15	**121.6**	122.4	*0.8*	122.2	*0.6*
C7-N13-H26	**120.1**	123.6	*3.4*	121.3	*1.1*
C8-C7-N13	**101.4**	103.9	*2.5*	102.7	*1.3*
Dihedral angles					
C2-C3-C4-C7	**-175.5**	-178.2	*-2.7*	-177.5	*-2.0*
C3-C4-C7-C8	**-170.6**	-167.6	*3.0*	-173.2	*-2.5*
C4-C7-C8-N10	**-113.6**	-119.0	*-5.4*	-120.0	*-6.5*
C7-C8-N10-C11	**2.0**	2.4	*0.4*	6.7	*4.7*
C8-N10-C11-N13	**-7.5**	-2.6	*4.9*	-8.3	*-0.8*
N10-C11-N13-C7	**10.3**	1.5	*-8.7*	6.2	*-4.1*
C11-N13-C7-C14	**-124.6**	-115.0	*9.7*	-117.4	*7.2*
N13-C7-C14-C15	**2.8**	3.4	*0.6*	-3.3	*-6.1*
C7-C14-C15-C16	**179.7**	-178.4	*1.9*	-178.5	*1.8*
O9-C8-N10-H25	**0.2**	0.2	*0.0*	-2.6	*-2.8*
H25-N10-C11-O12	**-4.6**	1.1	*5.7*	0.9	*5.5*
O12-C11-N13-H26	**-24.8**	-1.8	*23.0*	-26.5	*-1.7*
C11-N13-C7-C8	**-8.9**	-0.2	*8.7*	-2.3	*6.6*
N13-C7-C8-N10	**3.9**	-1.2	*-5.1*	-2.4	*-6.4*

See legend for Table S8.

Table SA10. BZC vibrational frequencies along with MOLVIB assignments

	Scaled MP2/6-31G(d)										new													
#	Freq	PED									Freq	PED												
1	**35.3**	torC3-C4	70	torC2-O1	24						**40.3**	torC3-C4	68	torC2-O1	19									
2	**72.5**	torC2-O1	68	torC3-C4	29						**88.4**	torO1-C3	32	wC4H	29	ta6RNG1a	15							
3	**76.3**	wC4H	40	torO1-C3	25	ta6RNG1a	18				**93.8**	torC2-O1	78	torC3-C4	21									
4	**98.8**	dO1C3C4	35	dC4H	27	iO1	24				**110.0**	iO1	31	dO1C3C4	29	dC4H	22							
5	**137.9**	torO1-C3	54	ta6RNG1a	28						**149.0**	torO1-C3	45	ta6RNG1a	26									
6	**238.3**	dC4H	36	iO1	20	dC2CO	19				**237.1**	dC4H	31											
7	**254.5**	torC1-C2	64								**262.1**	torC1-C2	85											
8	**273.3**	torN1-C7	99								**270.6**	torN1-C7	86											
9	**283.1**	torC1-C2	27	wC4H	26	tp6RNG1	20				**294.4**	iO1	33	sC4-C3	18									
10	**296.4**	sC4-C3	22	da6RNG1a	21	iO1	20				**308.3**	wC7H	42	wC4H	23									
11	**359.7**	dC7H	36	dC3O	21						**373.9**	dC7H	25	dC3O	19									
12	**371.0**	ta6RNG1	79	ta6RNG1a	29						**388.8**	dC7H	34	dC2CO	28									
13	**372.8**	dC2CO	32	dC7H	27						**423.3**	ta6RNG1	85	ta6RNG1a	30									
14	**415.4**	tp6RNG1	80								**477.3**	dO1C3C4	27	dC2CO	24									
15	**479.2**	dO1C3C4	34	dC4H	16						**484.3**	ta6RNG1a	34	wC7H	23									
16	**481.0**	wC7H	75								**574.9**	sC7-N1	20	dt6RNG1	17									
17	**592.8**	da6RNG1a	24	dC3O	18						**642.8**	tp6RNG1	89											
18	**617.8**	da6RNG1	60	da6RNG1a	16						**687.9**	da6RNG1	60	da6RNG1a	18									
19	**648.0**	wN1C7	78								**713.9**	wN1C7	65											
20	**677.6**	wC3O2	75								**794.1**	wC3O2	50											
21	**742.1**	wC8H	44	wC6H	22						**821.3**													
22	**759.5**	wC6H	56	wC8H	27						**854.1**	wC8H	34	wC9H	32	wC6H	16							
23	**785.2**	dt6RNG1	19								**856.4**	wC6H	26	wC5H	21	rC2H	16							
24	**790.8**	rC1H'	43	rC2H	42						**869.6**	sO1-C3	22											
25	**818.2**	wC9H	82	wC8H	27						**873.8**	rC2H	38	rC1H'	25									
26	**828.9**										**926.9**	sO1-C3	34	rC1H	34									
27	**835.2**	wC5H	89	wC6H	19						**959.7**	wC9H	27	wC5H	25	wC8H	24	wC6H						
28	**869.2**	sC2-O1	38	rC1H	25	sC1-C2	15				**962.7**	dt6RNG1	36											
29	**984.4**	dt6RNG1	48								**994.0**	wC5H	19	wC6H	18	wC9H	17	wC8H						
30	**1020.5**	sC1-C2	40	sC2-O1	25						**995.6**													
31	**1055.8**	dN1C7	46								**1014.4**	dN1C7	54											
32	**1097.4**	sO1-C3	25	sC2-O1	24						**1048.4**	sC1-C2	60											
33	**1107.0**	rC1H	40	sC1-C2	25	dC2CO	16				**1066.8**	rC1H'	59	rC2H	35									
34	**1116.1**	dN1C7	21	dC6H	19	dC8H	18				**1110.6**	sC2-O1	34	rC1H	26									
35	**1151.2**	rC2H	55	rC1H'	30						**1126.1**	sC8-C9	34	sC5-C6	17									
36	**1161.3**	dC8H	24	dC9H	21	dC5H	17				**1176.1**	sC5-C6	33	sC8-C9	21									
37	**1250.3**	iC2H	78	rC1H'	19						**1269.8**	sC4-C3	22											
38	**1262.0**	sC4-C3	26	sO1-C3	21						**1284.0**	sC7-N1	21	dC8H	15									
39	**1266.5**	sC7-N1	48								**1315.3**	iC2H	93											
40	**1284.6**	dC9H	27	dC5H	25	dC6H	16				**1362.4**	dC5H	22											
41	**1355.2**	wC2H	58	dsC1H	25						**1397.2**													
42	**1384.4**	sC9-C4	20	sC4-C5	19	sC6-C7	15				**1406.9**	dsC1H	94											
43	**1396.7**	dsC1H	69	wC2H	17						**1418.2**	cC2-H	55	wC2H	27									
44	**1423.4**	sC5-C6	32	sC8-C9	30						**1429.0**	daC1H	83											

new results are from calculations using optimized CHARMM force field parameters used in this study All frequencies (freq) are in cm^{-1} and MOLVIB assignments using potential energy decomposition (PED) are in %. Definition of various independent internal coordinates can be found in (56).

Table SA10 (continued). BZC vibrational frequencies along with MOLVIB assignments

	Scaled MP2/6-31G(d)							new					
#	Freq	PED						Freq	PED				
45	**1463.2**	daC1H'	93					**1429.5**	daC1H'	91			
46	**1473.4**	daC1H	72	cC2-H	22			**1446.6**	sC8-C9	16	sC5-C6	16	
47	**1489.8**	cC2-H	69	daC1H	16			**1464.2**	dN1H10	65			
48	**1496.0**							**1478.9**	dC9H	18			
49	**1566.2**	sC6-C7	20	sC4-C5	18	sC9-C4	17	**1517.8**	dC5H	27	dC9H	24	
50	**1598.9**	dN1H10	29	sC5-C6	16	sC8-C9	15	**1557.7**	wC2H	45	cC2-H	38	
51	**1620.6**	dN1H10	67					**1583.0**	sC7-N1	27			
52	**1702.3**	sC3-O2	83					**1757.2**	sC3-O2	83			
53	**2941.3**	sC1-H	100					**2854.9**	sC2-H	98			
54	**2951.6**	sC2-H	100					**2892.0**	sC2-H	99			
55	**3003.5**	sC2-H	92					**2902.3**	sC1-H	98			
56	**3026.7**	sC1-H	100					**2959.4**	sC1-H	100			
57	**3032.6**	sC8-H8	97					**2959.8**	sC1-H	99			
58	**3033.2**	sC6-H7	95					**3054.0**	sC9-H9	58	sC5-H6	34	
59	**3039.1**	sC1-H	93					**3055.9**	sC5-H6	49	sC9-H9	31	
60	**3066.1**	sC5-H6	95					**3057.2**	sC6-H7	49	sC8-H8	41	
61	**3074.8**	sC9-H9	97					**3060.2**	sC6-H7	40	sC8-H8	40	
62	**3367.4**	sN1-H	100					**3459.5**	sN1-H	100			
63	**3472.3**	sN1-H	100					**3541.2**	sN1-H	100			

new results are from calculations using optimized CHARMM force field parameters used in this study All frequencies (freq) are in cm^{-1} and MOLVIB assignments using potential energy decomposition (PED) are in %. Definition of various independent internal coordinates can be found in (56).

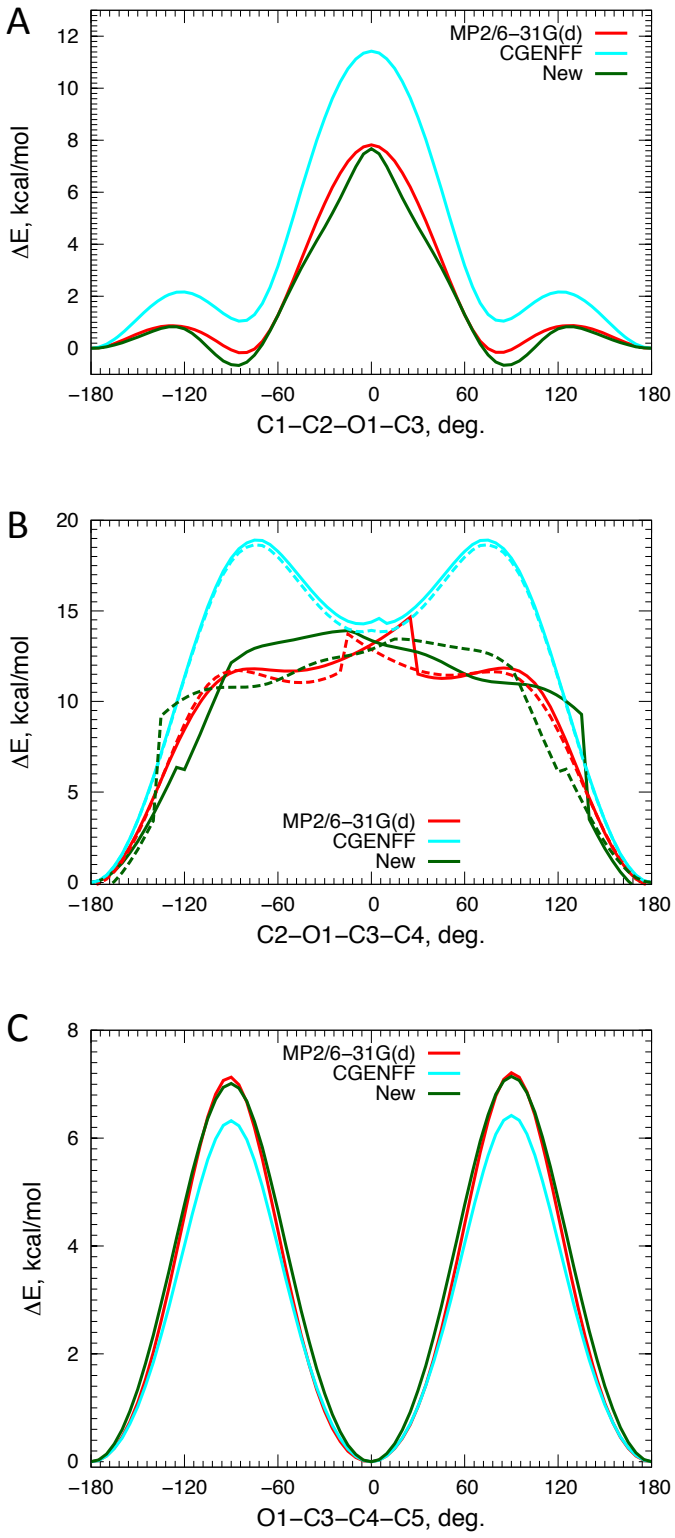

Figure SA1. BZC relaxed dihedral potential energy scans for (A) C1–C2–O1–C3, (B) C2–O1–C3–C4, (C) O1–C3–C4–C5 dihedral angles. Other internal degrees of freedom were allowed to relax. MP2/6-31G(d) scans are shown by red lines, CHARMM scans using parameters generated by the CGENFF program (version 0.9.7 beta) as cyan lines, and CHARMM scans using optimized parameters (new) by dark-green lines. Scans in the forward direction (from −180° to 180°) are shown as solid lines, whereas those in the backward direction (from 180° to −180°) are shown as dashed lines.

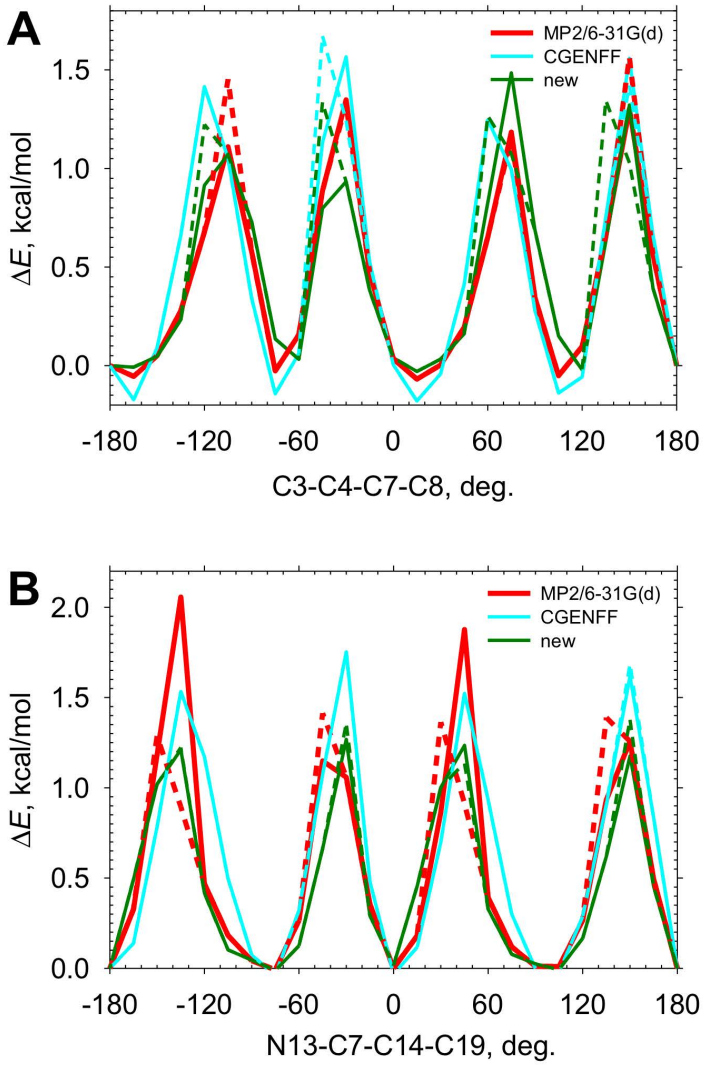

Figure SA2. PHT relaxed dihedral potential energy scans for (A) C3–C4–C7–C8, (B) N13–C7–C14–C19 dihedral angles. Other internal degrees of freedom were allowed to relax. MP2/6-31G(d) scans are shown as red lines, CHARMM scans using parameters generated by the CGENFF program (version 0.9.7 beta) as cyan lines, and CHARMM scans using optimized parameters (new) as dark-green lines. Scans in the forward direction (from −180° to 180°) are shown as solid lines, whereas those in the backward direction (from 180° to −180°) are shown as dashed lines.

Appendix SA2. Topology and parameters for BZC

```
* Toppar stream file generated by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
* and modified by Igor Vorobyov, Sept. 2013
*

read rtf card ! append
* Topologies generated by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
* and modified by Igor Vorobyov, Sept. 2013
*
36 1

MASS   257 HGA2     1.00800  ! alphatic proton, CH2
MASS   258 HGA3     1.00800  ! alphatic proton, CH3
MASS   269 HGP4     1.00800  ! polar H, neutral conjugated -NH2 group (NA bases)
MASS   277 HGR6    1.00800  ! aromatic H
MASS   294 CGCA   12.01100  ! carbonyl C: esters, [neutral] carboxylic acids
MASS   304 CGR6  12.01100  ! 6-mem aromatic C
MASS   318 CG32   12.01100  ! aliphatic C for CH2
MASS   322 CG33   12.01100  ! aliphatic C for methyl group (-CH3)
MASS   341 NGAM   14.00700  ! external amine ring nitrogen (planar/aniline), phosphoramidate
MASS   375 OGES   15.99940  ! ester -O-
MASS    365 OGCA    15.99940   ! carbonyl O: amides, esters, [neutral] carboxylic acids, aldehydes,
uera

RESI BNZC       0.000 !
GROUP            ! CHARGE
ATOM C1     CG33  -0.269 !
ATOM C2     CG32   0.060 !
ATOM O1     OGES  -0.307 !
ATOM C3     CGCA   0.466 !
ATOM O2     OGCA  -0.494 !
ATOM C4     CGR6  0.086 !
ATOM C5     CGR6 -0.140 !
ATOM C6     CGR6 -0.140 !
ATOM C7     CGR6  0.168 !
ATOM C8     CGR6 -0.140 !
ATOM C9     CGR6 -0.140 !
ATOM N1     NGAM  -0.742 !
ATOM H1     HGA3    0.090 !
ATOM H2     HGA3    0.090 !
ATOM H3     HGA3    0.090 !
ATOM H4     HGA2    0.090 !
ATOM H5     HGA2    0.090 !
ATOM H6     HGR6   0.115 !
ATOM H7     HGR6   0.115 !
ATOM H8     HGR6   0.115 !
ATOM H9     HGR6   0.115 !
ATOM H10    HGP4    0.341 !
ATOM H11    HGP4    0.341 !

BOND C1   C2
BOND C1   H1
BOND C1   H2
BOND C1   H3
BOND C2   O1
BOND C2   H4
BOND C2   H5
BOND O1   C3
BOND C3   O2
BOND C3   C4
BOND C4   C9
BOND C4   C5
BOND C5   C6
BOND C5   H6
BOND C6   C7
BOND C6   H7
BOND C7   C8
BOND C7   N1
BOND C8   C9
BOND C8   H8
BOND C9   H9
```

```
BOND N1   H10
BOND N1   H11
IMPR C3     C4     O2     O1
IMPR N1     H11    H10    C7

auto angle dihe

END

read param card flex  ! append
* Parameters generated by analogy by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
* and modified by Igor Vorobyov, September 2013
*

ATOMS
MASS   257 HGA2     1.00800  ! alphatic proton, CH2
MASS   258 HGA3     1.00800  ! alphatic proton, CH3
MASS   269 HGP4     1.00800  ! polar H, neutral conjugated -NH2 group (NA bases)
MASS   277 HGR6    1.00800  ! aromatic H
MASS   294 CGCA   12.01100  ! carbonyl C: esters, [neutral] carboxylic acids
MASS   304 CGR6  12.01100  ! 6-mem aromatic C
MASS   318 CG32   12.01100  ! aliphatic C for CH2
MASS   322 CG33   12.01100  ! aliphatic C for methyl group (-CH3)
MASS   341 NGAM   14.00700  ! external amine ring nitrogen (planar/aniline), phosphoramidate
MASS   375 OGES   15.99940  ! ester -O-
MASS    365 OGCA    15.99940   ! carbonyl O: amides, esters, [neutral] carboxylic acids, aldehydes,
uera

BONDS
CGCA  CGR6  254.00     1.4800 ! ZOIC, benzoic acid, MBOA, methylbenzoate, jal
CGCA   OGCA    750.00       1.2200  ! PROT adm jr. 5/02/91, acetic acid pure solvent; LIPID methyl
acetate
CGCA  OGES   150.00     1.3340 ! LIPID methyl acetate
CGR6 CGR6  305.00     1.3750 ! PROT benzene, JES 8/25/89
CGR6 NGAM   400.00     1.3900 ! PYRIDINE aminopyridine, adm jr., 7/94
CGR6 HGR6   340.00     1.0800 ! PROT phe,tyr JES 8/25/89
CG32  CG33   222.50     1.5280 ! PROT alkane update, adm jr., 3/2/92
CG32  OGES   320.00     1.4400 ! PROTNA serine/threonine phosphate
CG32  HGA2    309.00     1.1110 ! PROT alkane update, adm jr., 3/2/92
CG33  HGA3    322.00     1.1110 ! PROT alkane update, adm jr., 3/2/92
NGAM  HGP4    488.00     1.0160 ! viv 09/13


ANGLES
CGR6 CGCA  OGCA    70.00    125.10   20.00   2.44200 ! viv 09/13
CGR6 CGCA  OGES    50.00    111.00   20.00   2.36000 ! MBOA, methylbenzoate, jal
OGCA  CGCA  OGES    90.00    125.90  160.00   2.25760 ! LIPID acetic acid
CGCA  CGR6 CGR6   45.00    120.00 ! ZOIC, benzoic acid, MBOA, methylbenzoate, jal
CGR6 CGR6 CGR6   40.00    120.00   35.00   2.41620 ! PROT JES 8/25/89
CGR6 CGR6 NGAM    45.00    121.00 ! viv 09/13
CGR6 CGR6 HGR6    30.00    120.00   22.00   2.15250 ! PROT JES 8/25/89 benzene
CG33  CG32  OGES    75.70    108.70 ! viv 09/13
CG33  CG32  HGA2     34.60    110.10   22.53   2.17900 ! PROT alkane update, adm jr., 3/2/92
OGES  CG32  HGA2     60.00    109.50 ! PROT adm jr. 4/05/91, methyl acetate
HGA2   CG32  HGA2     35.50    109.00    5.40   1.80200 ! PROT alkane update, adm jr., 3/2/92
CG32  CG33  HGA3     34.60    110.10   22.53   2.17900 ! PROT alkane update, adm jr., 3/2/92
HGA3   CG33  HGA3     35.50    108.40    5.40   1.80200 ! PROT alkane update, adm jr., 3/2/92
CGR6 NGAM  HGP4     42.00    113.60 ! viv 09/13
HGP4   NGAM  HGP4     31.00    107.50 ! viv 09/13
CGCA  OGES  CG32    40.00    113.00   30.00   2.26510 ! viv 09/13

DIHEDRALS
OGCA  CGCA  CGR6 CGR6     1.1500  2   180.00 ! viv 09/13
OGES  CGCA  CGR6 CGR6     0.9500  2   180.00 ! viv 09/13
CGR6 CGCA  OGES  CG32      1.2500  1   180.00 ! ZINC12 , from CGR6 CGCA OGES CG33, PENALTY= 0.9
CGR6 CGCA  OGES  CG32      1.5000  2   180.00 ! ZINC12 , from CGR6 CGCA OGES CG33, PENALTY= 0.9
CGR6 CGCA  OGES  CG32      0.0500  6   180.00 ! ZINC12 , from CGR6 CGCA OGES CG33, PENALTY= 0.9
OGCA  CGCA  OGES  CG32      0.9650  1   180.00 ! LIPID methyl acetate
OGCA  CGCA  OGES  CG32      3.8500  2   180.00 ! LIPID methyl acetate
CGCA   CGR6 CGR6 CGR6      3.1000   2    180.00 ! ZOIC, benzoic acid, MBOA, methylbenzoate; default
parameter; kevo & jal
CGCA   CGR6 CGR6 HGR6       2.4000   2    180.00 ! ZOIC, benzoic acid, MBOA, methylbenzoate; default
parameter; kevo & jal
CGR6 CGR6 CGR6 CGR6     3.1000  2   180.00 ! PROT JES 8/25/89
```

```
CGR6 CGR6 CGR6 NGAM      5.0000  2   180.00 ! PYRIDINE aminopyridine, yin
CGR6 CGR6 CGR6 HGR6      4.2000  2   180.00 ! PROT JES 8/25/89 benzene
NGAM  CGR6 CGR6 HGR6      2.4000  2   180.00 ! PYRIDINE aminopyridine Kenno: 4.2 -> 2.4
HGR6  CGR6 CGR6 HGR6      2.4000  2   180.00 ! PROT JES 8/25/89 benzene
CGR6 CGR6 NGAM  HGP4       1.1700  2   180.00 ! viv 09/13
OGES  CG32  CG33  HGA3       0.1950  3     0.00 ! PROT alkane update, adm jr., 3/2/92
HGA2   CG32  CG33  HGA3       0.1600  3     0.00 ! PROT rotation barrier in Ethane (SF)
CG33  CG32  OGES  CGCA      0.0014  1      0.00 ! viv 09/13  MP2/6-31G(d) fit
CG33  CG32  OGES  CGCA      3.0430  2    180.00 ! viv 09/13  MP2/6-31G(d) fit
CG33  CG32  OGES  CGCA      1.8598  3    180.00 ! viv 09/13  MP2/6-31G(d) fit
CG33  CG32  OGES  CGCA      0.7461  4    180.00 ! viv 09/13  MP2/6-31G(d) fit
HGA2   CG32  OGES  CGCA      1.0865  1      0.00 ! viv 09/13 MP2/6-31G(d) fit
HGA2   CG32  OGES  CGCA      2.9983  2    180.00 ! viv 09/13 MP2/6-31G(d) fit
HGA2   CG32  OGES  CGCA      0.8517  3      0.00 ! viv 09/13 MP2/6-31G(d) fit
HGA2   CG32  OGES  CGCA      0.8641  4    180.00 ! viv 09/13 MP2/6-31G(d) fit

IMPROPERS
CGCA  CGR6 OGCA  OGES     72.0000  0     0.00 ! MBOA, methyl benzoate; MOLVIB looks good; jal
NGAM   HGP4    HGP4    CGR6     -2.5000   0      0.00 ! -2.0 PYRIDINE aminopyridine 11/10 kevo: sic!
Compensates for in-plane force from CGR6 CGR6 NGAM HGP4

NONBONDED nbxmod  5 atom cdiel fshift vatom vdistance vfswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
!see mass list above for better description of atom types
HGA2     0.0       -0.0350     1.3400 ! alkane, igor, 6/05
HGA3     0.0       -0.0240     1.3400 ! alkane, yin and mackerell, 4/98
HGR6    0.0       -0.0300     1.3582 ! benzene
HGP4     0.0       -0.0460     0.2245 ! polar H, conjugated amines (NA bases)
CGCA    0.0       -0.0980     1.7000 ! methyl acetate update viv 12/29/06
CGR6   0.0       -0.0700     1.9924 ! INDO/TRP
CG32    0.0       -0.0560     2.0100   0.0 -0.01 1.9 ! alkane (CT2), 4/98, yin, adm jr, also used by
viv
CG33     0.0         -0.0780      2.0500    0.0 -0.01 1.9 ! alkane (CT3), 4/98, yin, adm jr; Rmin/2
modified from 2.04 to 2.05
NGAM    0.0       -0.2000     1.8500 ! PROT
OGES    0.0       -0.1000     1.6500 ! ester; LJ from THP, sng 1/06
OGCA     0.0          -0.1200       1.7000    0.0 -0.12 1.40 ! carbonyl. Also consistent with adm,
acetaldehyde, 11/08
END
```

Appendix SA3. Topology and parameters for PHT

```
* Toppar stream file generated by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
* and modified by I. Vorobyov, September 2013
*

read rtf card flex !  append
* Topologies generated by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
*
36 1

MASS   266 HGP1     1.00800  ! polar H
MASS   277 HGR61    1.00800  ! aromatic H
MASS    302 CG2R53   12.01100   ! 5-mem ring, double bound to N and adjacent to another heteroatom,
purine C8, his CE1 (0,+1), 2PDO, kevo
MASS   304 CG2R61  12.01100  ! 6-mem aromatic C
MASS   329 CG3C50  12.01100  ! 5-mem ring aliphatic quaternary C (cholesterol, bile acids)
MASS    365 OG2D1    15.99940   ! carbonyl O: amides, esters, [neutral] carboxylic acids, aldehydes,
urea
MASS   349 NG2R53  14.00700  ! amide in 5-memebered NON-SP2 ring (slightly pyramidized), 2PDO, kevo

auto angle dihe

RESI PHT1       0.000
GROUP            ! CHARGE
ATOM C1     CG2R61 -0.115
ATOM C2     CG2R61 -0.110 !
ATOM C3     CG2R61 -0.155 !
ATOM C4     CG2R61  0.121 !
ATOM C5     CG2R61 -0.155 !
ATOM C6     CG2R61 -0.110 !
ATOM C7     CG3C50  0.160 !
ATOM C8     CG2R53  0.513 !
ATOM O9     OG2D1  -0.470 !
ATOM N10    NG2R53 -0.500 !
ATOM C11    CG2R53  0.490 !
ATOM O12    OG2D1  -0.450 !
ATOM N13    NG2R53 -0.560 !
ATOM C14    CG2R61  0.120 !
ATOM C15    CG2R61 -0.155 !
ATOM C16    CG2R61 -0.110 !
ATOM C17    CG2R61 -0.115 !
ATOM C18    CG2R61 -0.110 !
ATOM C19    CG2R61 -0.155 !
ATOM H20    HGR61   0.115 !
ATOM H21    HGR61   0.115 !
ATOM H22    HGR61   0.115 !
ATOM H23    HGR61   0.115 !
ATOM H24    HGR61   0.115 !
ATOM H25    HGP1    0.372 !
ATOM H26    HGP1    0.344 !
ATOM H27    HGR61   0.115 !
ATOM H28    HGR61   0.115 !
ATOM H29    HGR61   0.115 !
ATOM H30    HGR61   0.115 !
ATOM H31    HGR61   0.115 !

BOND C1   C6
BOND C1   C2
BOND C1   H20
BOND C2   C3
BOND C2   H21
BOND C3   C4
BOND C3   H22
BOND C4   C5
BOND C4   C7
BOND C5   C6
BOND C5   H23
BOND C6   H24
BOND C7   N13
```

```
BOND C7   C8
BOND C7   C14
BOND C8   O9
BOND C8   N10
BOND N10  C11
BOND N10  H25
BOND C11  O12
BOND C11  N13
BOND N13  H26
BOND C14  C19
BOND C14  C15
BOND C15  C16
BOND C15  H27
BOND C16  C17
BOND C16  H28
BOND C17  C18
BOND C17  H29
BOND C18  C19
BOND C18  H30
BOND C19  H31
IMPR C8     C7     N10    O9
IMPR C11    N10    N13    O12

END

read param card flex ! append
* Parameters generated by analogy by
* CHARMM General Force Field (CGenFF) program version 0.9.7 beta
* and modified by I. Vorobyov, September 2013
*

ATOMS
MASS   266 HGP1     1.00800  ! polar H
MASS   277 HGR61    1.00800  ! aromatic H
MASS    302 CG2R53  12.01100  ! 5-mem ring, double bound to N and adjacent to another heteroatom,
purine C8, his CE1 (0,+1), 2PDO, kevo
MASS   304 CG2R61  12.01100  ! 6-mem aromatic C
MASS   329 CG3C50  12.01100  ! 5-mem ring aliphatic quaternary C (cholesterol, bile acids)
MASS    365 OG2D1    15.99940  ! carbonyl O: amides, esters, [neutral] carboxylic acids, aldehydes,
urea
MASS   349 NG2R53  14.00700  ! amide in 5-memebered NON-SP2 ring (slightly pyramidized), 2PDO, kevo
validation/optimization.

BONDS
CG2R53 CG3C50  300.00     1.5300 ! Molecu , from CG2R53 CG3C52, PENALTY= 10 ! viv 09/13 ok
CG2R53 NG2R53  460.00     1.3800 !460 370 *NEW* 2PDO, 2-pyrrolidinone, kevo
CG2R53 OG2D1   570.00     1.2350 !560 620 *NEW* 2PDO, 2-pyrrolidinone, kevo
CG2R61 CG2R61  305.00     1.3750 ! PROT benzene, JES 8/25/89
CG2R61 CG3C50  230.00     1.4500 ! viv 09/13
CG2R61 HGR61   340.00     1.0800 ! PROT phe,tyr JES 8/25/89
CG3C50 NG2R53  370.00     1.4500 ! Molecu , from CG3C52 NG2R53, PENALTY= 10 ! viv 09/13 ok
NG2R53 HGP1    470.00     1.0150 !470 440 *NEW* 2PDO, 2-pyrrolidinone, kevo

ANGLES
CG3C50 CG2R53 NG2R53  95.00    109.80 ! viv 09/13
CG3C50 CG2R53 OG2D1    65.00    122.30 ! viv 09/13
NG2R53 CG2R53 NG2R53   75.00    104.40 ! MHYO, 5-methylenehydantoin, xxwy
NG2R53 CG2R53 OG2D1    65.00    127.80 ! 2PDO, 2-pyrrolidinone, kevo
CG2R61 CG2R61 CG2R61   40.00    120.00   35.00   2.41620 ! PROT JES 8/25/89
CG2R61 CG2R61 CG3C50   45.80    120.00 ! Molecu , from CG2R61 CG2R61 CG321, PENALTY= 10
CG2R61 CG2R61 HGR61    30.00    120.00   22.00   2.15250 ! PROT JES 8/25/89 benzene
CG2R53 CG3C50 CG2R61   62.00    104.20 ! viv 09/13
CG2R53 CG3C50 NG2R53  105.00    110.30 ! viv 09/13
CG2R61 CG3C50 CG2R61   51.80    106.00 ! viv 09/13
CG2R61 CG3C50 NG2R53   52.00     108.30 ! viv 09/13
CG2R53 NG2R53 CG2R53   55.00    113.50 ! MRDN, methylidene rhodanine, kevo & xxwy
CG2R53 NG2R53 CG3C50   55.00    113.50 ! viv 09/13
CG2R53 NG2R53 HGP1     38.00    119.50 ! 2PDO, 2-pyrrolidinone (H1-N1-C2), kevo
CG3C50 NG2R53 HGP1     38.00    116.00 ! Molecu , from CG3C52 NG2R53 HGP1, PENALTY= 1.2

DIHEDRALS
NG2R53 CG2R53 CG3C50 CG2R61      3.5000   3    180.00 ! Molecu , from NG2R50 CG2R52 CG3C52 CG2RC0,
PENALTY= 79.5
NG2R53 CG2R53 CG3C50 NG2R53      1.0500   3    180.00 ! Molecu , from NG2R53 CG2R53 CG3C52 CG3C52,
PENALTY= 82
```

```
OG2D1   CG2R53 CG3C50 CG2R61      0.0800   3      0.00 ! Molecu , from OG2D1 CG2R53 CG3C52 CG3C52,
PENALTY= 104
OG2D1   CG2R53 CG3C50 NG2R53      0.0800   3      0.00 ! Molecu , from OG2D1 CG2R53 CG3C52 CG3C52,
PENALTY= 82
CG3C50 CG2R53 NG2R53 CG2R53      0.5000   2    180.00 ! Molecu , from NG2R53 CG2R53 NG2R53 CG2R53,
PENALTY= 87.5
CG3C50 CG2R53 NG2R53 HGP1       3.5000  2   180.00 ! viv 09/13
NG2R53 CG2R53 NG2R53 CG2R53     0.5000  2   180.00 ! MHYO, 5-methylenehydantoin, xxwy
NG2R53 CG2R53 NG2R53 CG3C50      0.5000   2    180.00 ! Molecu , from NG2R53 CG2R53 NG2R53 CG311,
PENALTY= 30.6
NG2R53 CG2R53 NG2R53 HGP1       0.8000  2   180.00 ! MHYO, 5-methylenehydantoin, xxwy
OG2D1  CG2R53 NG2R53 CG2R53     1.1000  2   180.00 ! MRDN, methylidene rhodanine, kevo & xxwy
OG2D1   CG2R53 NG2R53 CG3C50      2.5900   2    180.00 ! Molecu , from OG2D1 CG2R53 NG2R53 CG3C52,
PENALTY= 1.2
OG2D1  CG2R53 NG2R53 HGP1       0.8600  2   180.00 ! 2PDO, 2-pyrrolidinone, kevo
CG2R61 CG2R61 CG2R61 CG2R61     3.1000  2   180.00 ! PROT JES 8/25/89
CG2R61 CG2R61 CG2R61 CG3C50      3.1000   2    180.00 ! Molecu , from CG2R61 CG2R61 CG2R61 CG321,
PENALTY= 10
CG2R61 CG2R61 CG2R61 HGR61      4.2000  2   180.00 ! PROT JES 8/25/89 benzene
CG3C50 CG2R61 CG2R61 HGR61       2.4000   2    180.00 ! Molecu , from CG321 CG2R61 CG2R61 HGR61,
PENALTY= 10
HGR61  CG2R61 CG2R61 HGR61      2.4000  2   180.00 ! PROT JES 8/25/89 benzene
CG2R61 CG2R61 CG3C50 CG2R61     1.1493  2   180.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 CG2R61     3.4435  3   180.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 CG2R61     0.1241  4     0.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 CG2R53     1.8520  2    180.0 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 CG2R53     2.8876  3     0.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 CG2R53     0.2424  4     0.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 NG2R53     1.4433  2   180.0 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 NG2R53     0.1841  3    0.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R61 CG2R61 CG3C50 NG2R53     0.1644  4    0.00 ! viv 09/13 MP2/6-31G(d) scan fit
CG2R53 CG3C50 NG2R53 CG2R53      2.3100   3    180.00 ! Molecu , from CG3C52 CG3C52 NG2R53 CG2R53,
PENALTY= 84
CG2R53 CG3C50 NG2R53 HGP1       0.0000   3     0.00 ! viv 09/13
CG2R61 CG3C50 NG2R53 CG2R53     3.5000  3   180.00 ! viv 09/13
CG2R61 CG3C50 NG2R53 HGP1       0.1000  3      0.00 ! viv 09/13 introduced ~26 deg. puckering based
on QM

IMPROPERS
CG2R53 CG3C50 NG2R53 OG2D1      90.0000   0      0.00 ! Molecu , from CG2R53 CG3C52 NG2R53 OG2D1,
PENALTY= 1
CG2R53 NG2R53 NG2R53 OG2D1      90.0000   0      0.00 ! MHYO, 5-methylenehydantoin, xxwy from 2PDO
WILDCARD

NONBONDED nbxmod  5 atom cdiel fshift vatom vdistance vfswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

!see mass list above for better description of atom types
HGP1     0.0       -0.0460     0.2245 ! polar H
HGR61    0.0       -0.0300     1.3582 ! benzene
CG2R53   0.0        -0.0200     2.2000 ! IMIA, imidazole; bulk solvent of 5 maybridge cmpds (kevo);
consistent with CG2R64
CG2R61   0.0       -0.0700     1.9924 ! INDO/TRP
CG3C50    0.0        -0.0360      2.0100    0.0 -0.01 1.9 ! extrapolation based on CG301, CG321 and
CG3C52, kevo
OG2D1     0.0         -0.1200       1.7000    0.0 -0.12 1.40 ! carbonyl. Also consistent with adm,
acetaldehyde, 11/08
NG2R53   0.0       -0.2000     1.8500 ! amide in 5-memebered ring (slightly pyramidized), 2PDO, kevo

END
```