SUPPORTING INFORMATION

Evaluating force field performance in thermodynamic calculations of cyclodextrin host-guest binding: water models, partial charges, and host force field parameters.

Niel M. Henriksen and Michael K. Gilson*

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, USA
*Email: mgilson@ucsd.edu

Supporting Table Captions

All Supporting Tables are provided in the associated spreadsheet file (SupportingTables.xlsx). The captions for each table are provided here:

Table S1. Comparison of the experimental and calculated binding free energies (kcal/mol) for all force field combinations. For each host-guest pair, a mean and SEM value is provided. Error metrics characterizing the deviation of the calculated values from the experimental values are given below.

Table S2. Comparison of the experimental and calculated binding enthalpies (kcal/mol) for all force field combinations. For each host-guest pair, a mean and SEM value is provided. Error metrics characterizing the deviation of the calculated values from the experimental values are given below.

Table S3. Error statistics between the computed and experimental binding free energy and binding enthalpy for each force field combination. Error statistics are provided for a variety of sub-groupings of the full host-guest data set, including: ammonium guests, alcohol guests, carboxylate guests, αCD hosts, βCD hosts, n-aliphatic guests, aliphatic carboxylates, and aromatic carboxylates. Details on the composition of each sub-group is provided below.

Table S4. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-TIP3P data set.

Table S5. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-TIP3P-sm1, -sm2, -sm3, and -shw data set.

Table S6. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-TIP4Pew data set.

Table S7. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-SPC data set.

Table S8. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-OPC data set.

Table S9. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-OPC-jc data set.

Table S10. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-OPC-phos data set.

Table S11. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4RG-OPC-jcphos data set.

Table S12. Complete experimental and calculated thermodynamic values (kcal/mol) for the Q4BG-TIP3P data set.

Table S13. Complete experimental and calculated thermodynamic values (kcal/mol) for the BGRG-TIP3P data set.

Table S14. Complete experimental and calculated thermodynamic values (kcal/mol) for the BGBG-TIP3P data set.

Table S15. Decomposition of the valence (bonded) enthalpic binding contribution into Bond, Angle, Dihedral, 1-4 LJ and 1-4 Electrostatic terms for selected host-guest pairs in the Q4RG-TIP3P simulation set (below) and the BGBG-TIP3P (at right). The mean values of each term are given for the bound and unbound simulation along with the binding difference (ΔE). All values in kcal/mol.

Table S16. Decomposition of the enthalpic binding contribution for a selection of host-guest pairs from the Q4RG-TIP3P simulation. Values are given for the full system (full), solute (sol), solvent (wat), and solute-solvent (w-s) components. These values are futher decomposed by Valence, Lennard-Jones (LJ), and Electrostatic (Elec) terms. The left column has ammonium guests, and the right column has carboxylate guests. The host-guest ID (see Table 1) is given along with a "-p" and "-s" suffix to indicate primary and secondary orientation of the guest, respectively. Total values differ slightly from Table S1 due to use of fewer data points from the simulation. All values in kcal/mol.

Supporting Figures

Supporting Figures are provided below.

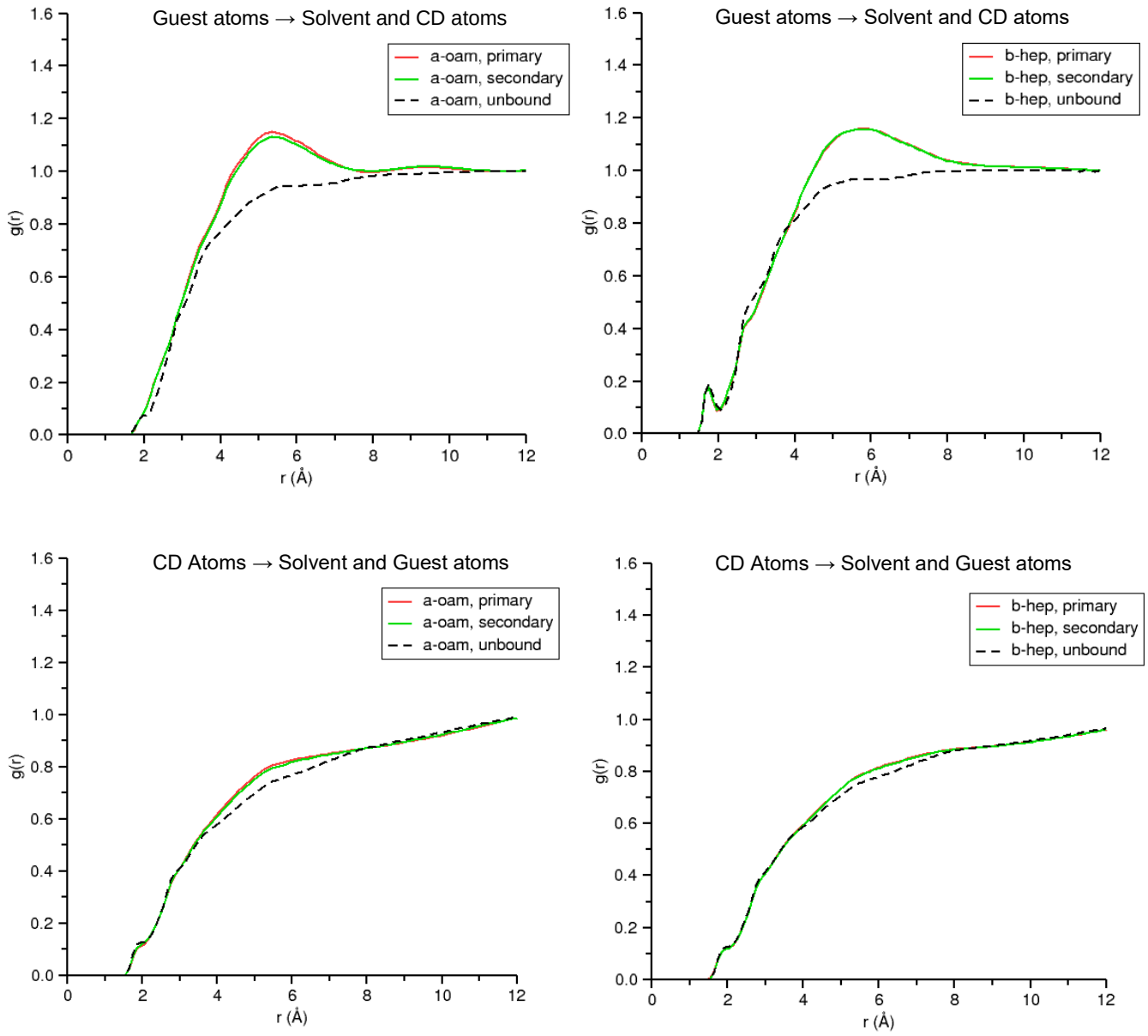

Figure S1. Radial distribution function (RDF) between guest atoms and the rest of the system atoms (*top row*) and between cyclodextrin atoms and all other atoms in the system (*bottom row*). In this approach, pairwise distances are computed between all atoms in the target (either guest or CD) and all other surrounding atoms. The density of the distances are then binned and normalized according to the volume of the shell defined by the bin. Two host-guest pairs are examined: αCD/octylammonium (a-oam) and βCD/heptanoate. RDFs are show for the guest bound in the primary orientation (red), secondary orientation (green), and unbound (black dashed).

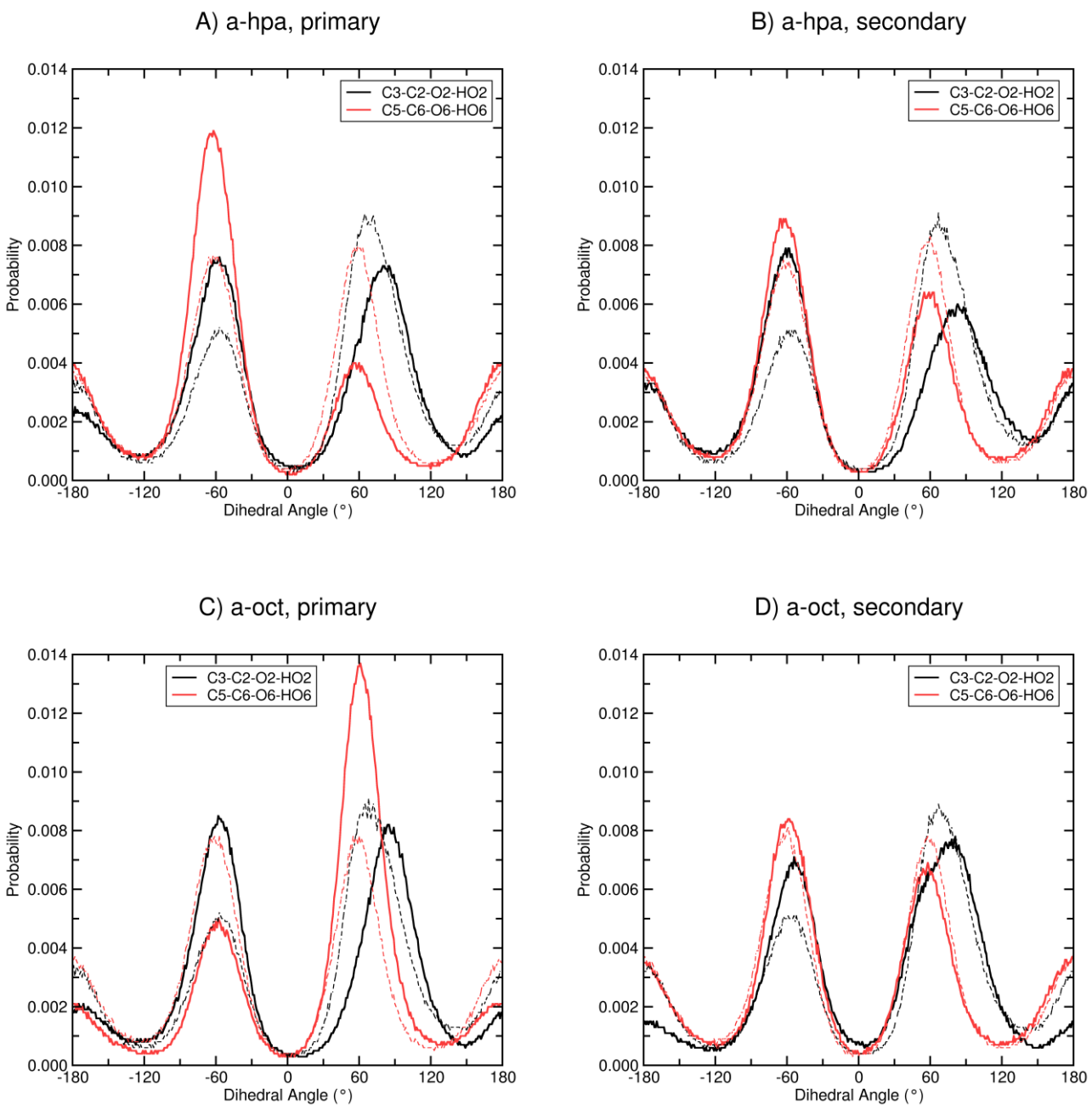

Figure S2. Dihedral probabilities for the O2 (black) and O6 (red) CD hydroxyls in the bound (solid) and unbound state (dashed). Data are shown for the a-hpa host-guest pair in the primary (A) and secondary orientation (B). Similarly, the a-oct host-guest pair in the primary (C) and secondary orientation (D).

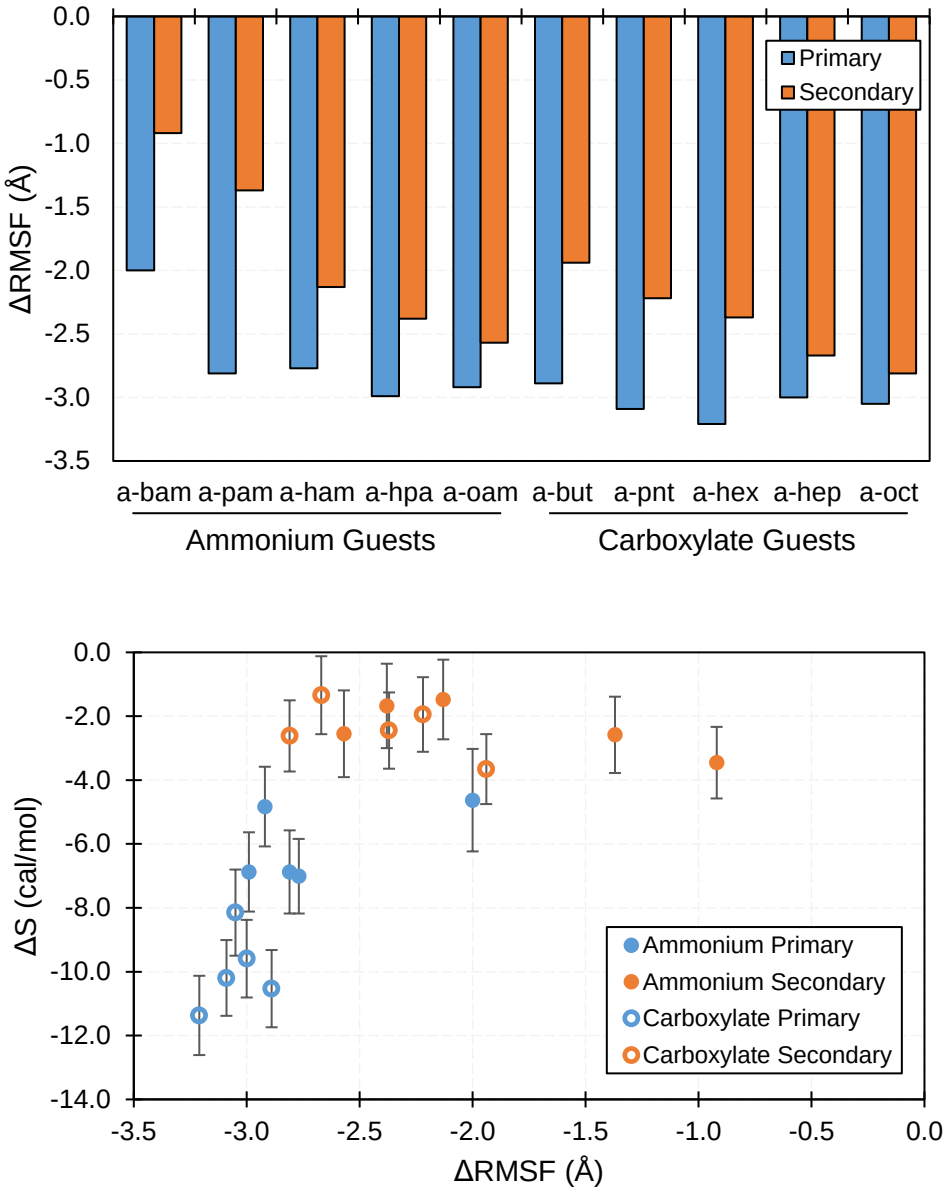

Figure S3. (Top) The summed difference in root mean squared fluctuations (ΔRMSF), i.e., the sum of the differences in oxygen atom RMSF between the bound and unbound state in αCD, for a selection of n-alkyl guests binding in the primary (blue) and secondary (orange) orientations. (Bottom) Correlation between the ΔRMSF and the computed binding entropy for a selection of n-alkyl guests. Ammonium guests are filled dots and carboxylate guests are open dots; binding orientation is indicated by blue for the primary side and orange for secondary side.

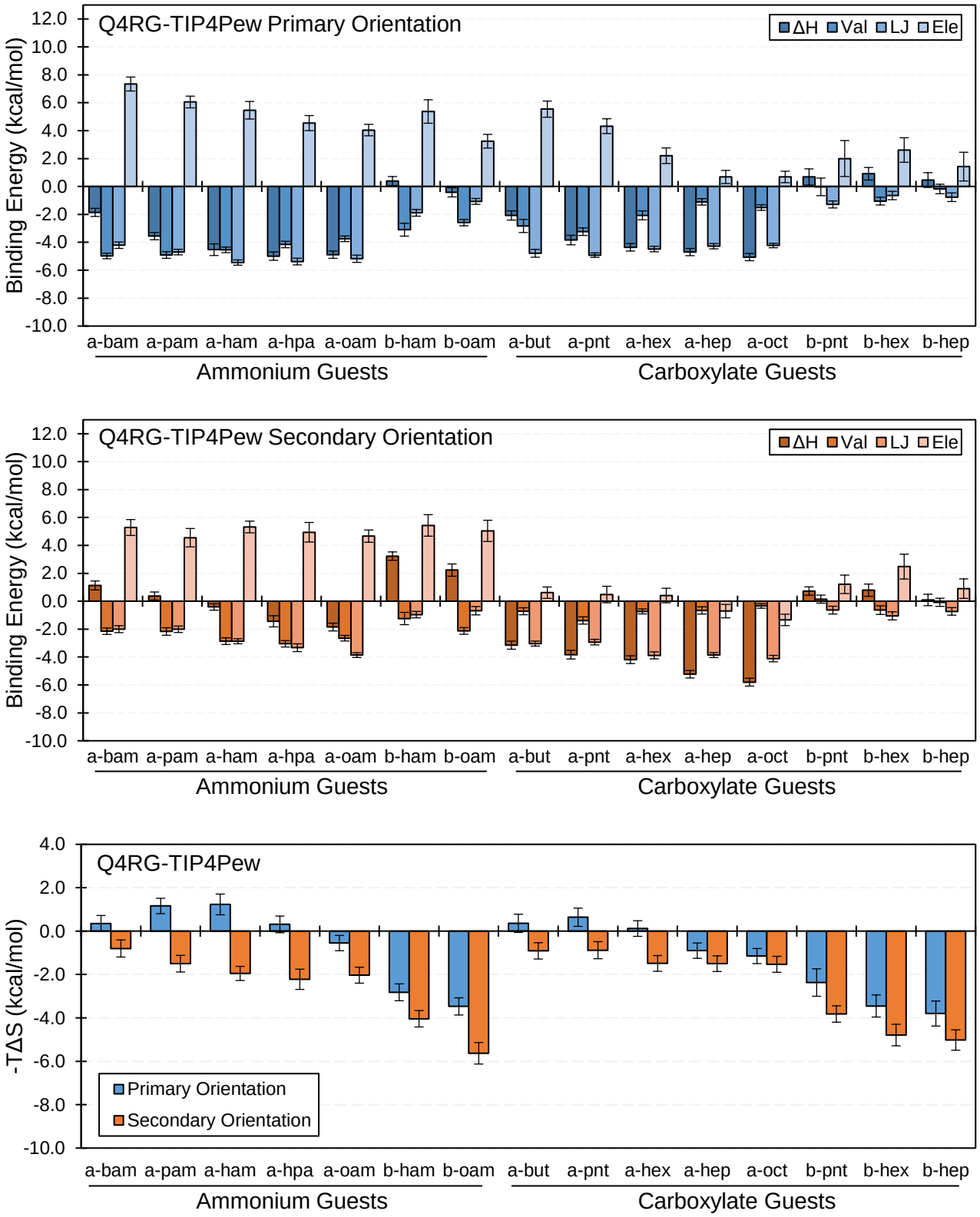

Figure S4. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the Q4RG-TIP4Pew simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

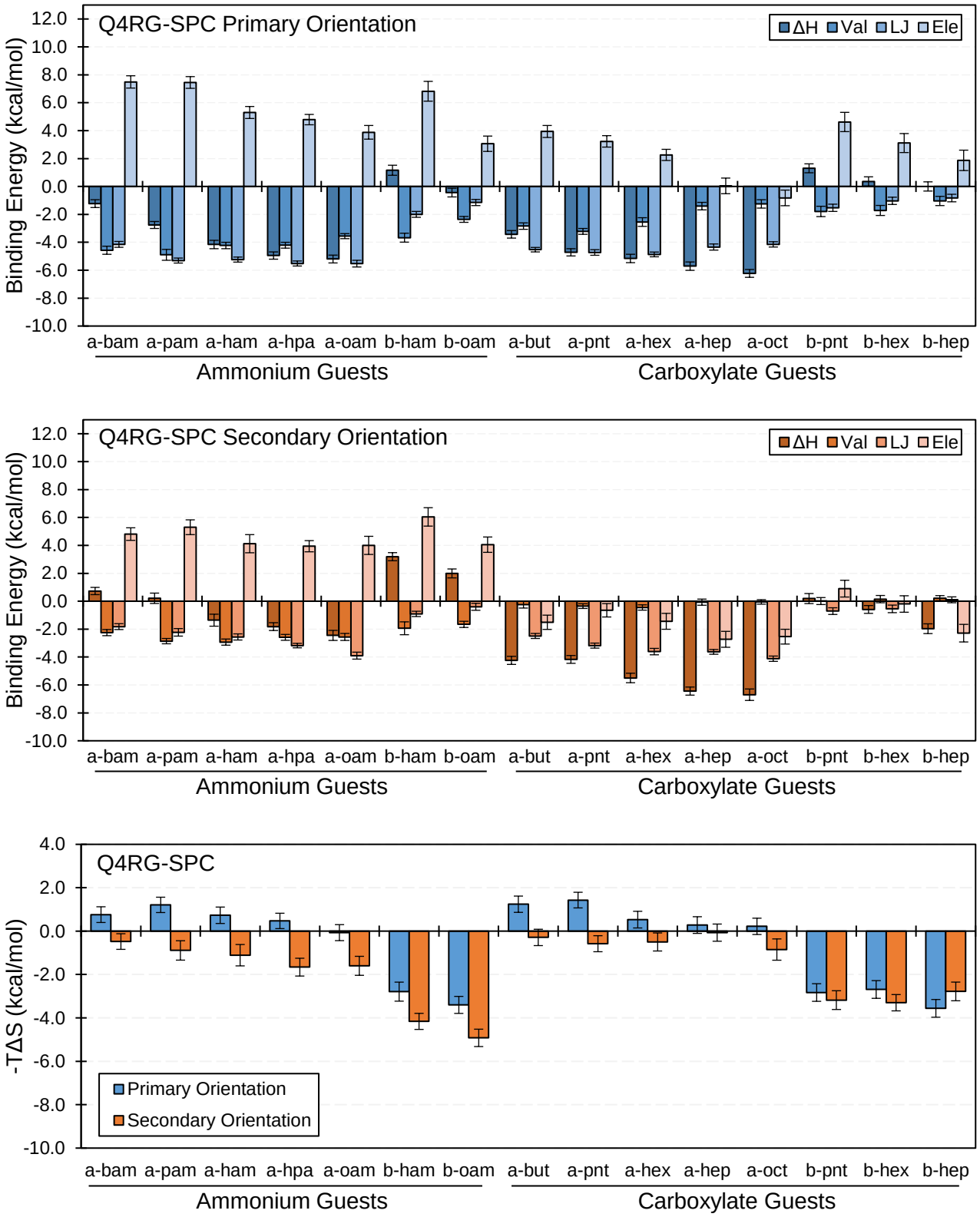

Figure S5. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the Q4RG-SPC simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

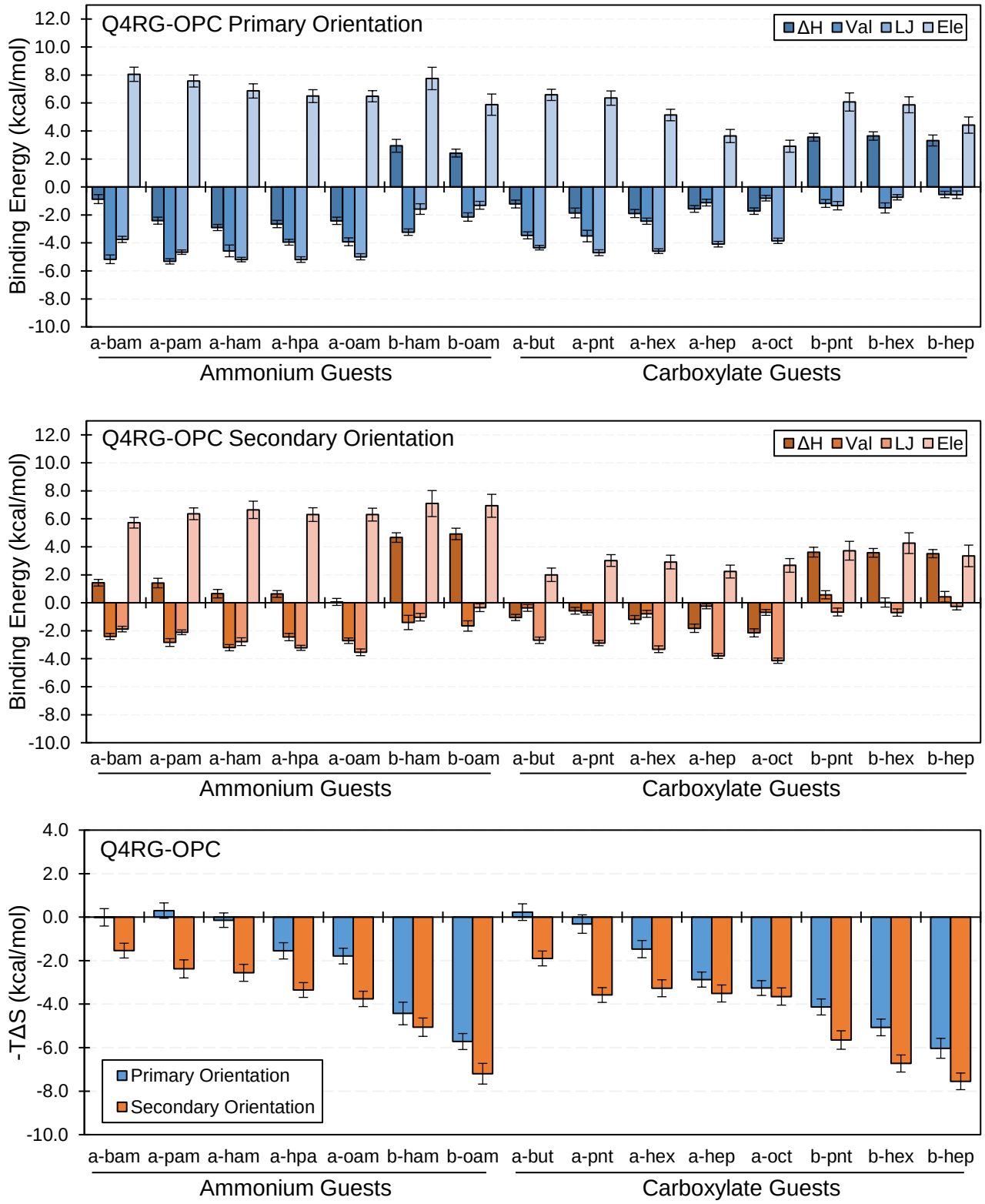

Figure S6. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the Q4RG-OPC simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

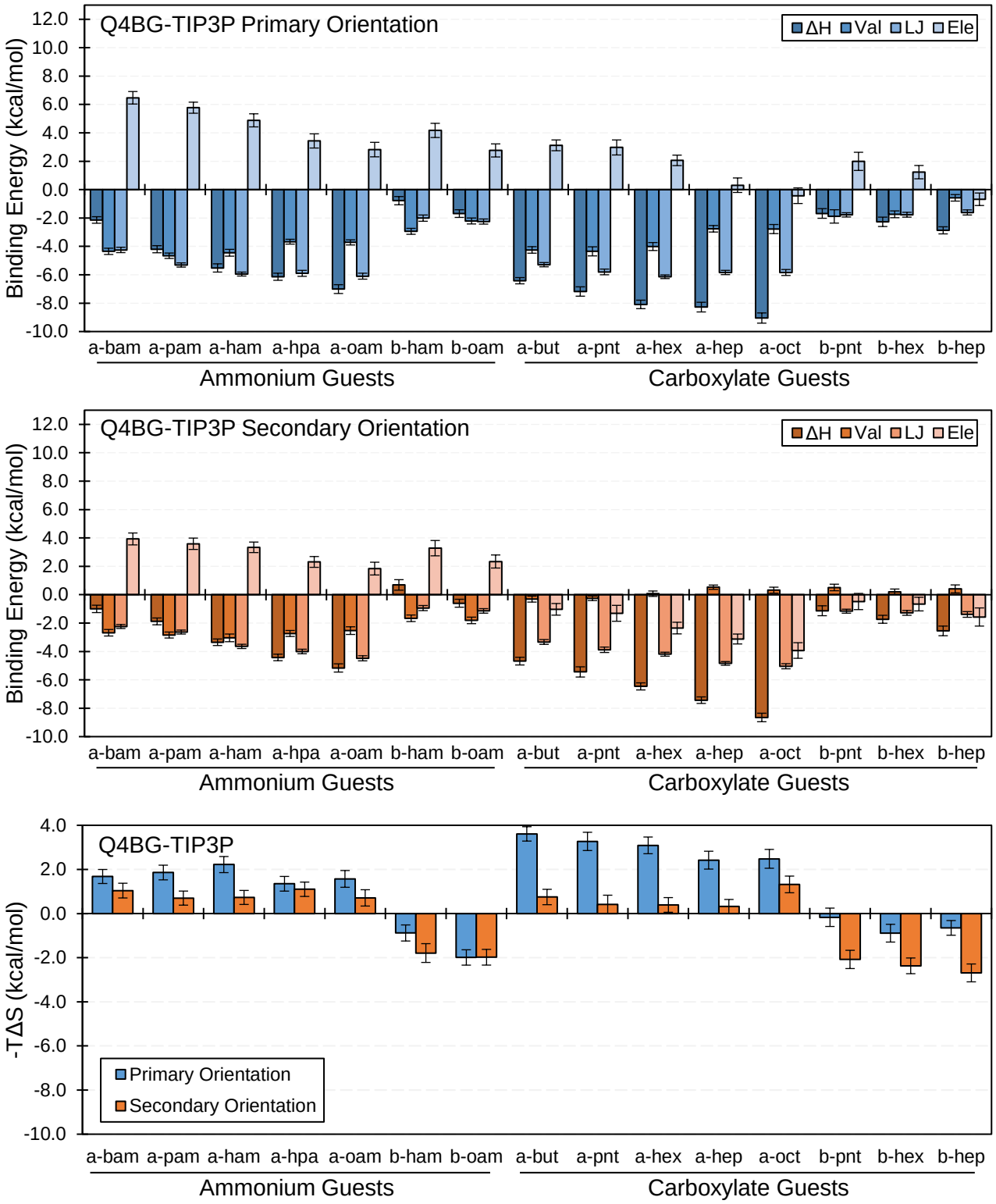

Figure S7. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the Q4BG-TIP3P simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

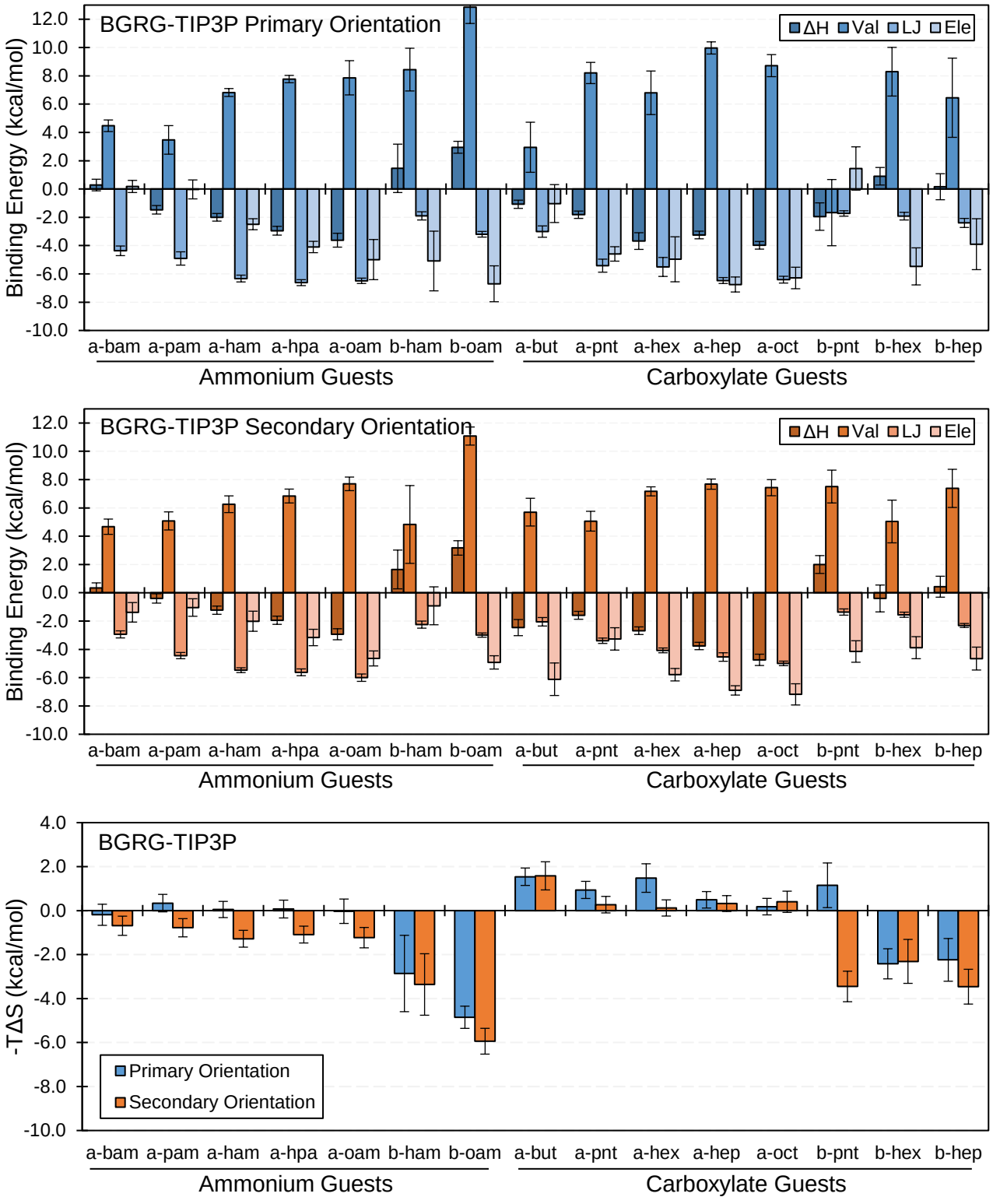

Figure S8. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the BGRG-TIP3P simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

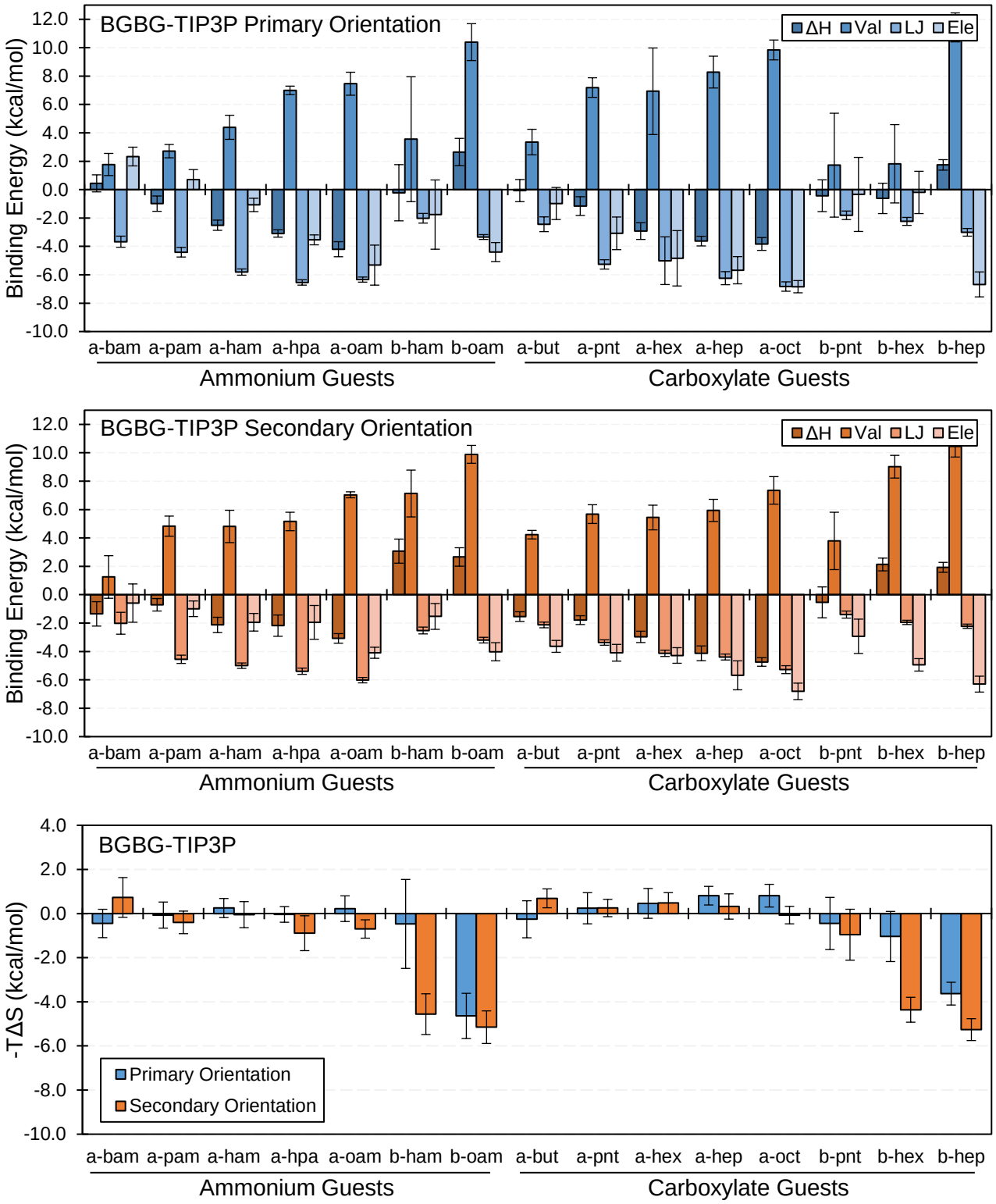

Figure S9. Binding enthalpy decompositions for the primary orientation (top/blue) and secondary orientation (center/orange), and entropic contribution to the free energy for both orientations (bottom) for selected guests from the BGBG-TIP3P simulation set. The number of guest carbon atoms increases from left to right for each guest class (see Table 1 for identification). Val = valence energy, LJ = Lennard-Jones energy, Ele = electrostatic energy.

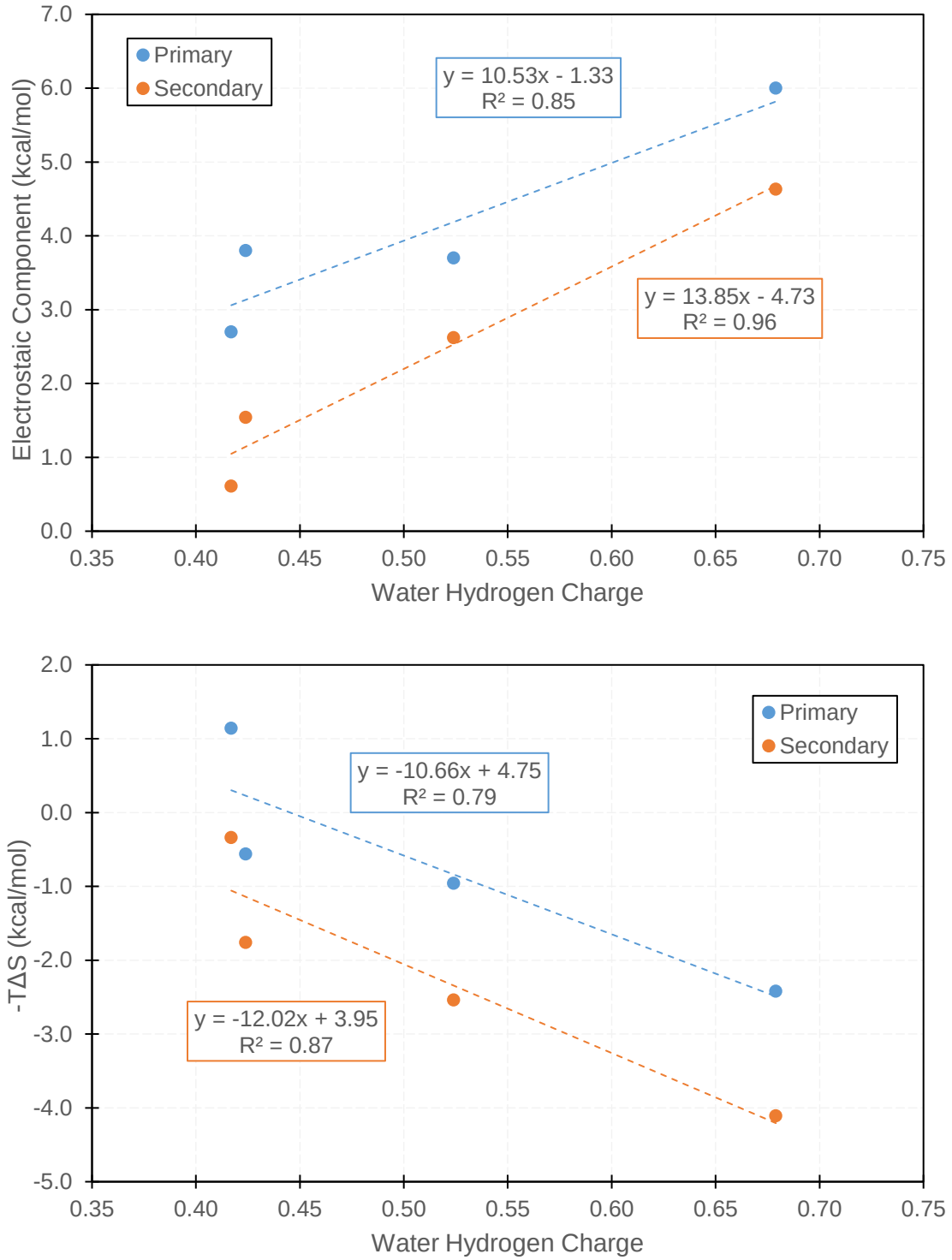

Figure S10. Correlation of the water oxygen charge magnitude with the average electrostatic contribution to the binding enthalpy (top) and with the average entropic contribution to the free energy (bottom). Correlation is shown for binding in both the primary and secondary orientations. The four points represent the four studied water models (simulation sets): TIP3P (Q4RG-TIP3P), TIP4Pew (Q4RG-TIP4Pew), SPC/E (Q4RG-SPC), and OPC (Q4RG-OPC). The average is performed over the n-alkyl ammonium and carboxylate guests (same as those in Figure S4).

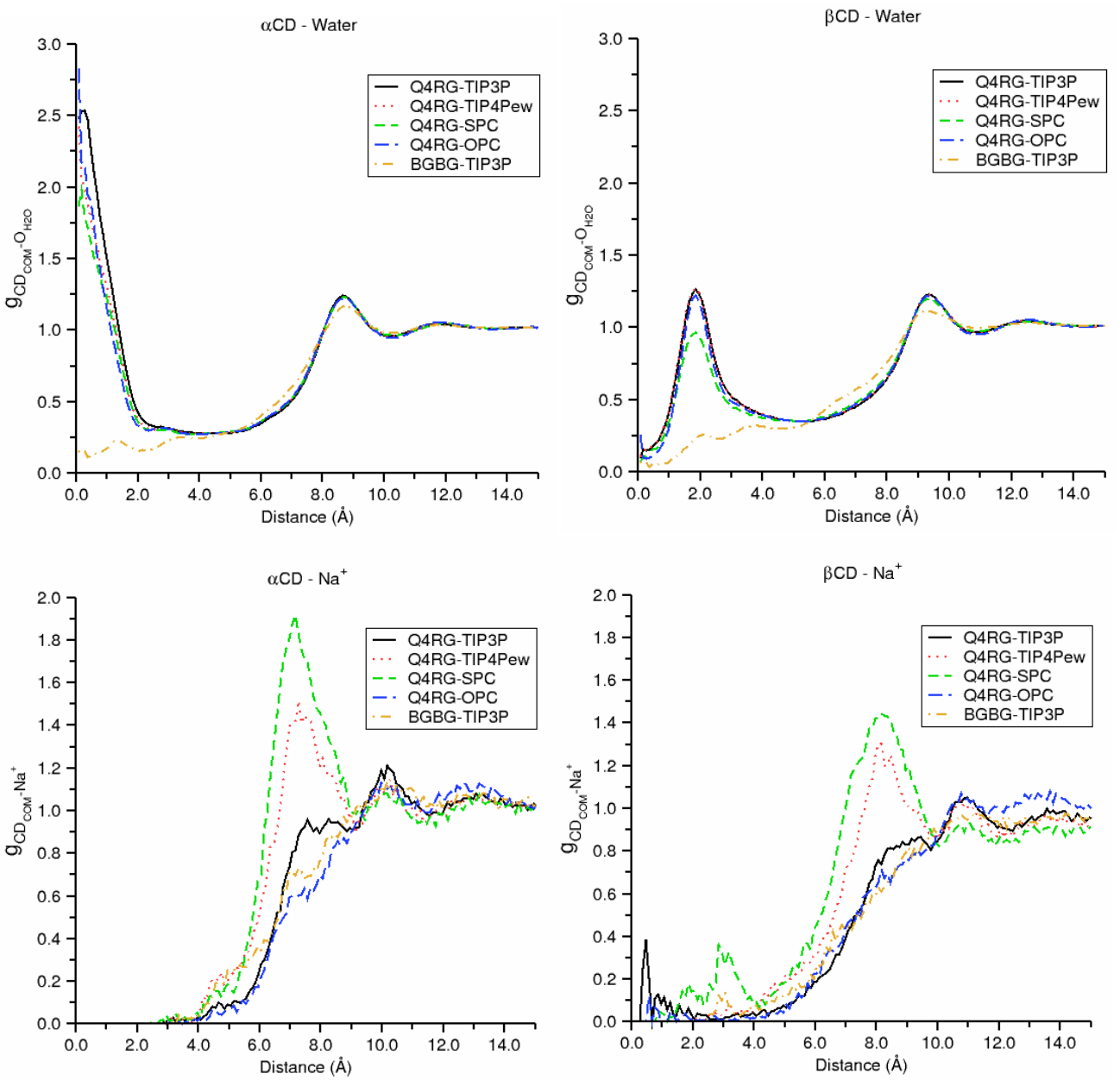

Figure S11. Radial distribution function plots between the CD center-of-mass and water (top) and between the CD center-of-mass and sodium ion (bottom) for the indicated simulation sets. Plots are shown for αCD (left) and βCD (right).

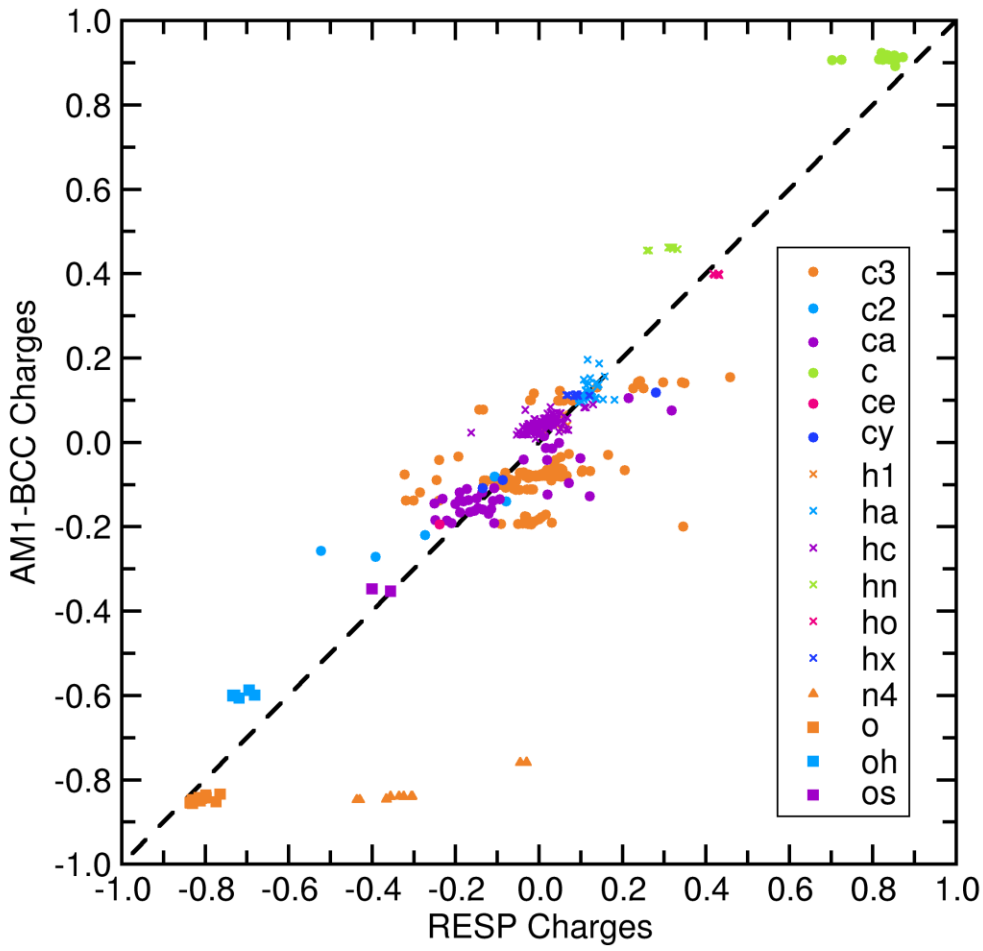

Figure S12. Comparison of guest partial charges assigned via the RESP and AM1-BCC methods. Carbon, hydrogen, nitrogen, and oxygen atoms are indicated by their initial letter and the identity of the GAFF atom type is given in the color coded legend.

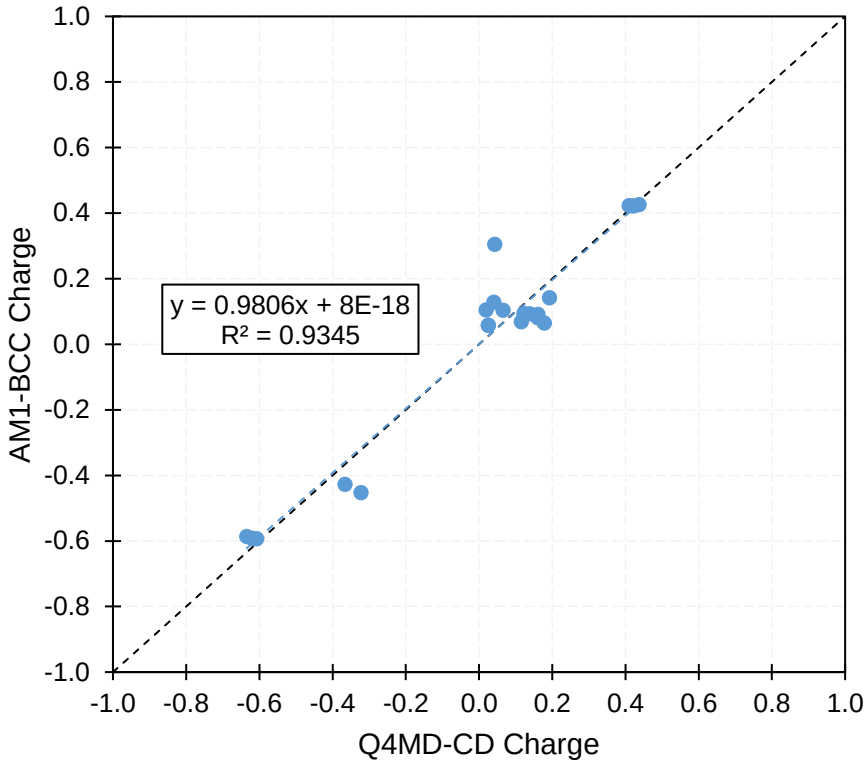

Figure S13. Comparison of the Q4MD-CD and AM1-BCC partial charge assignments for each unique atom in the CD host. The Q4MD-CD charges were derived by Cezard et al[1] using the RESP approach implemented in the R.E.D. Server tool.[2] The AM1-BCC charges were obtained via Antechamber calculation of a single glucose monomer with methoxy capped O1 and O4 termini, followed by adjustment of the partial charges to ensure neutrality following removal of the capping groups. The AM1-BCC charges were used in all simulations sets which begin with "BG" rather than "Q4".

1) Cézard, C.; Trivelli, X.; Aubry, F.; Djedaïni-Pilard, F.; Dupradeau, F.-Y. *Phys. Chem. Chem. Phys.* **2011**, *13* (33), 15103.
2) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. *Nucl. Acids Res.* **2011**, *39* (suppl 2), W511.

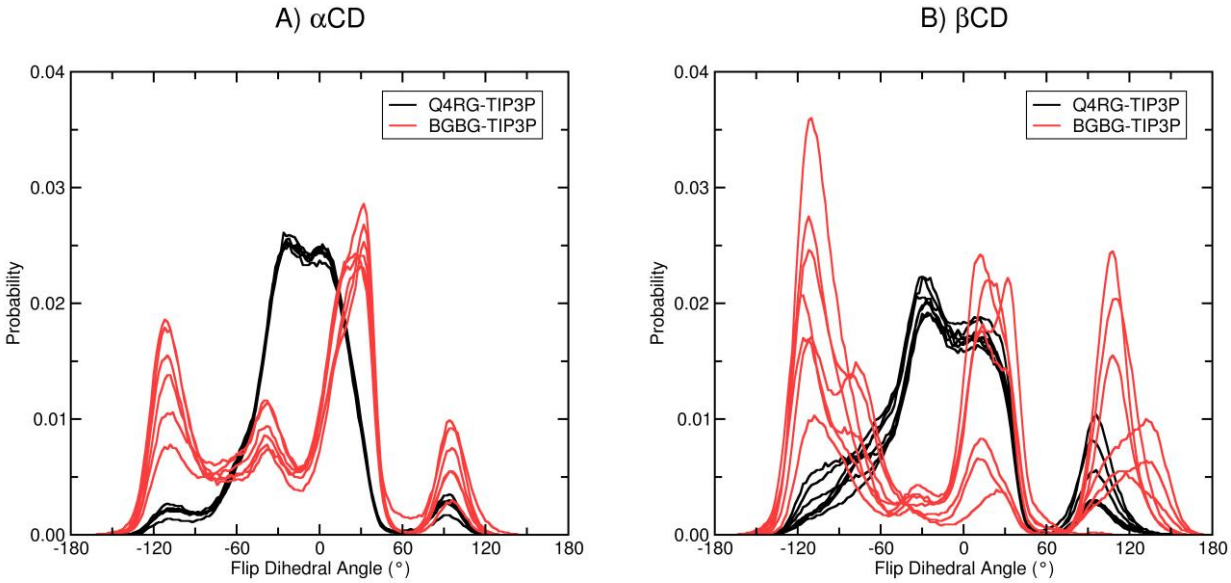

Figure S14. Probability distribution for the "flip" *pseudo*-dihedral angle, $O3_n$-$C4_n$-$C1_{n+1}$-$O2_{n+1}$, for unbound αCD (A) and βCD (B) from the Q4RG-TIP3P simulation set (black) and the BGBG-TIP3P simulation set (red). Six lines in A represent the six glucose monomers in αCD and likewise seven monomers in βCD.

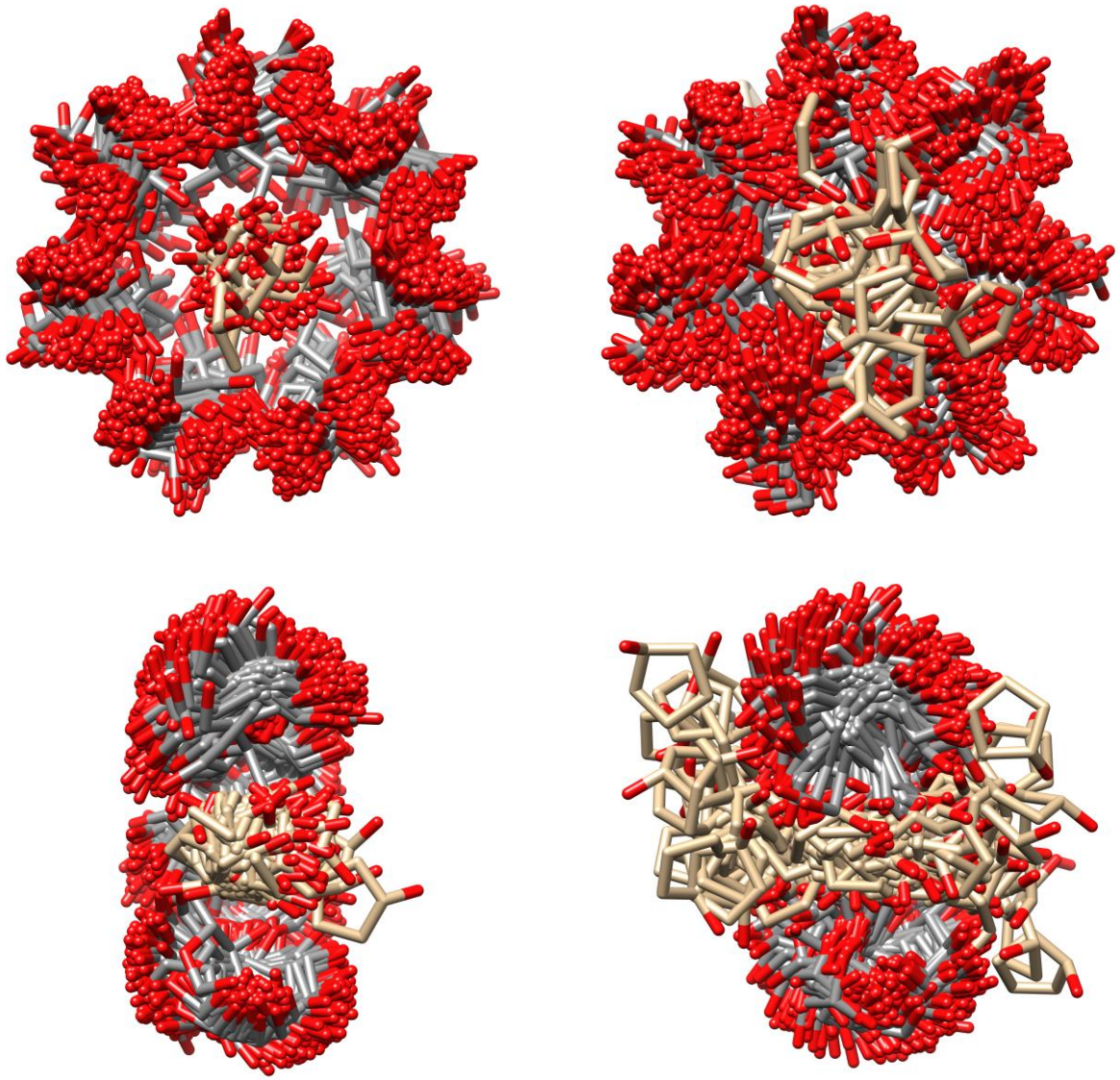

Figure S15. Comparison of the βCD:cyclopentanol secondary oriented bound state conformations generated by two different force field combinations. Two hundred evenly spaced snapshots from the simulation trajectories are overlaid via RMSD fit of the host atoms. The left column shows the ensemble generated by the Q4RG-TIP3P force field; at right is the ensemble generated by the BGBG-TIP3P force field. The top row presents a view looking towards secondary opening of the cyclodextrin. The bottom row presents a view from the side, with three of the glycosyl residues removed for easier viewing. The carbon atoms of the cyclodextrin are colored grey, whereas those of the cyclopentanol are colored tan.

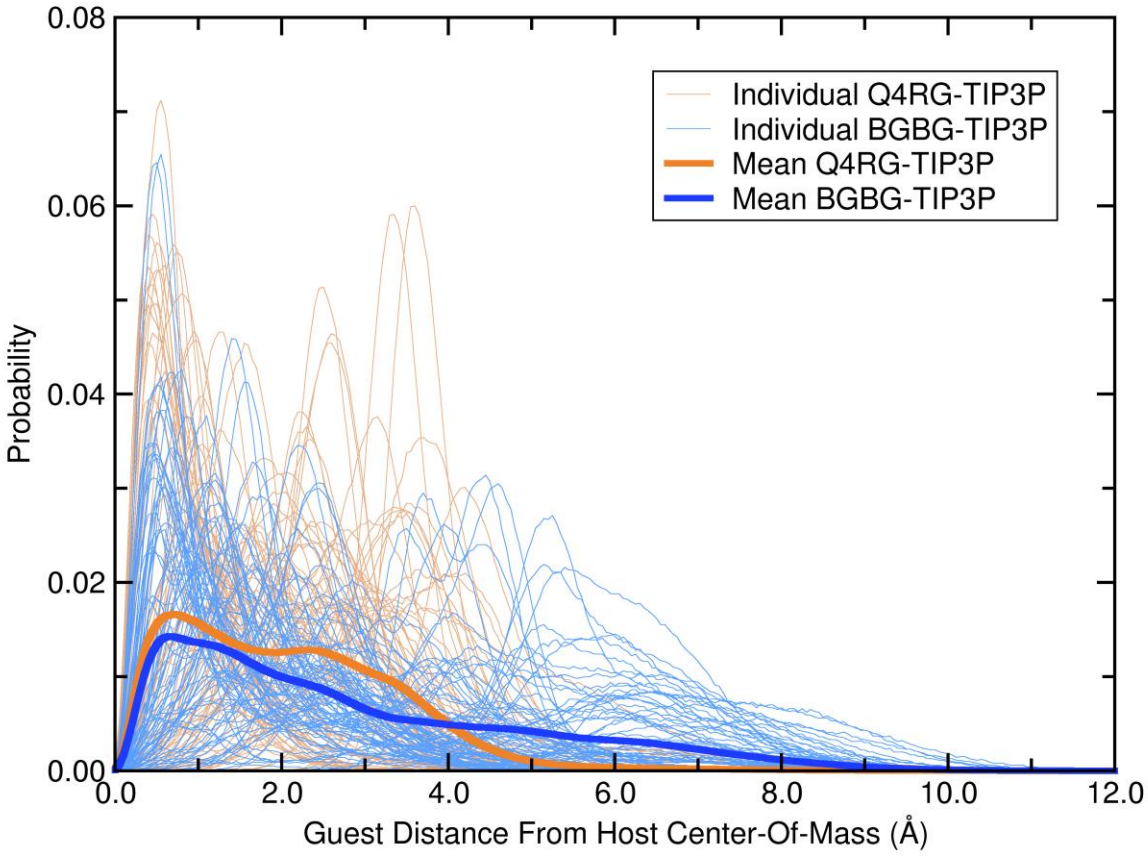

Figure S16. Probability distributions which demonstrate the fluctuation of all guests in the bound state for the Q4RG-TIP3P simulation set (blue) and the BGBG-TIP3P simulation set (orange). The distance is measured between the center of mass of the cyclodextrin and the center of mass of the guest. The thick lines represent the mean value at each histogram bin across all guests in both the primary and secondary orientation.

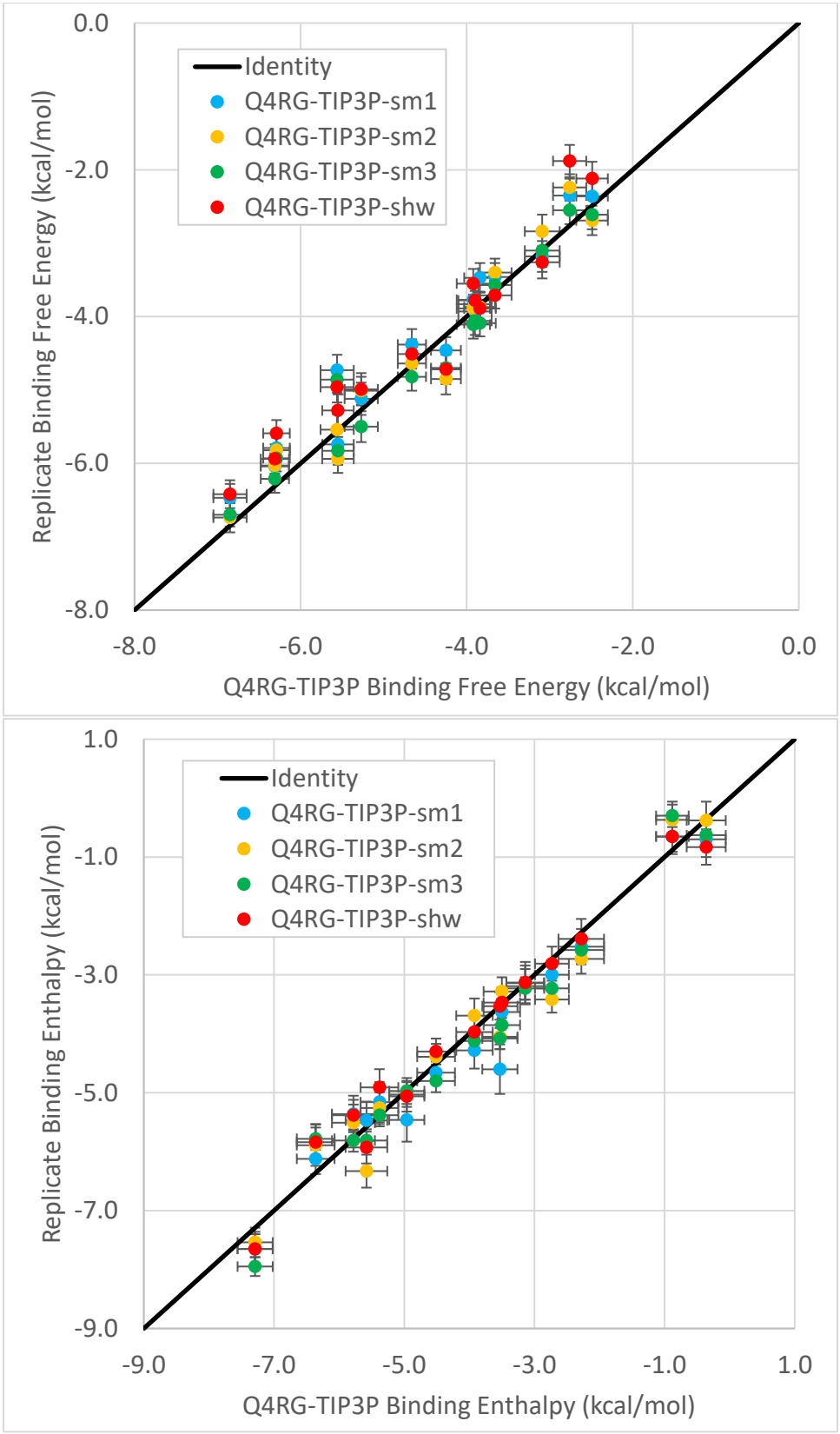

Figure S17. Scatter plot of binding free energies (top) and binding enthalpies (bottom) between a selection of 15 host-guest pairs from the Q4RG-TIP3P simulation set and a series of replicate simulation sets: Q4RG-TIP3P-sm1, Q4RG-TIP3P-sm2, and Q4RG-TIP3P-sm3. Additionally, a simulation set with the same selection of host-guest pairs but shorter "wall restraints" (see Methods) was evaluated: Q4RG-TIP3P-shw.

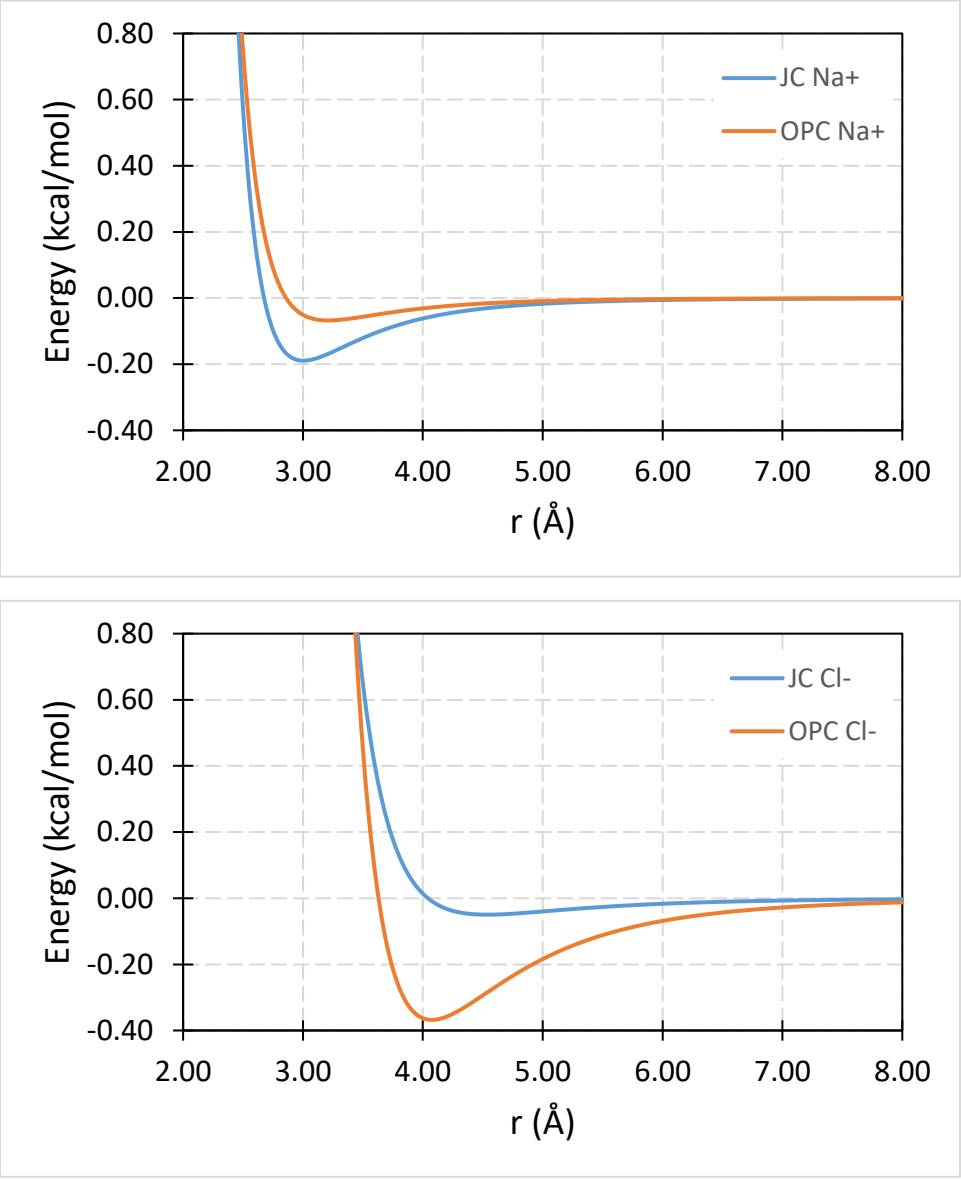

Figure S18. Comparison between the ion-water Lennard-Jones potential of the Joung-Cheatham parameters (blue) and the OPC specific parameters (orange). The potential for OPC water interacting with sodium (top) and chloride (bottom) are shown.