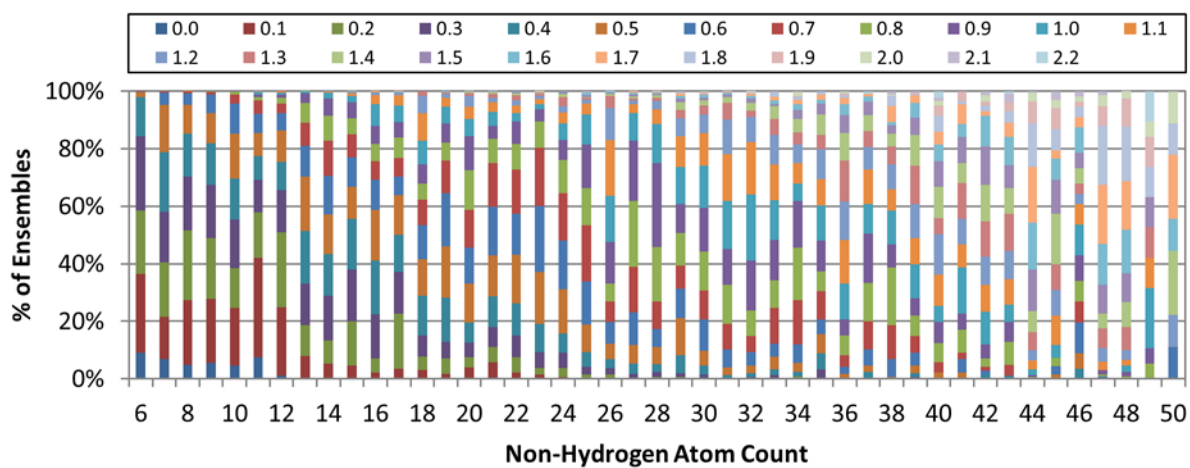
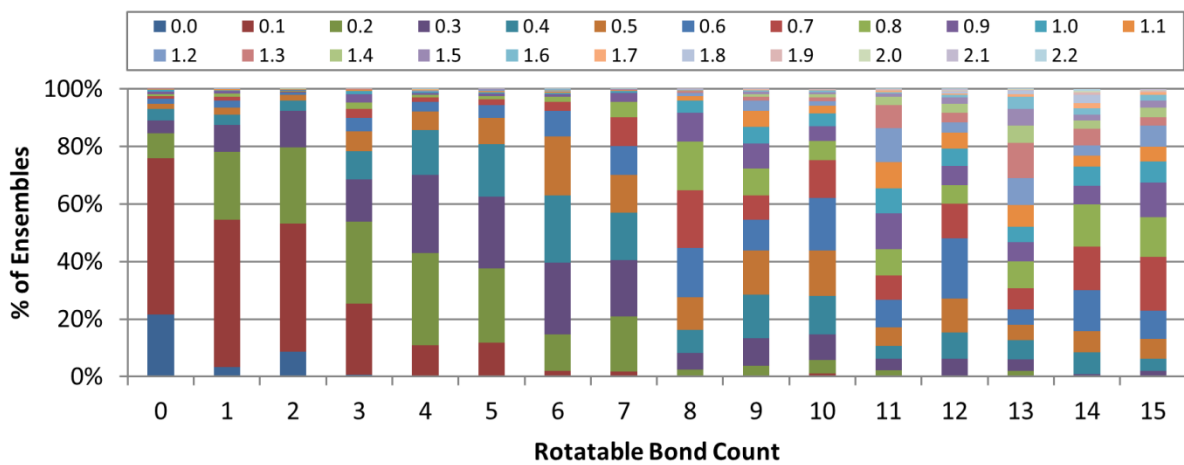


(a)

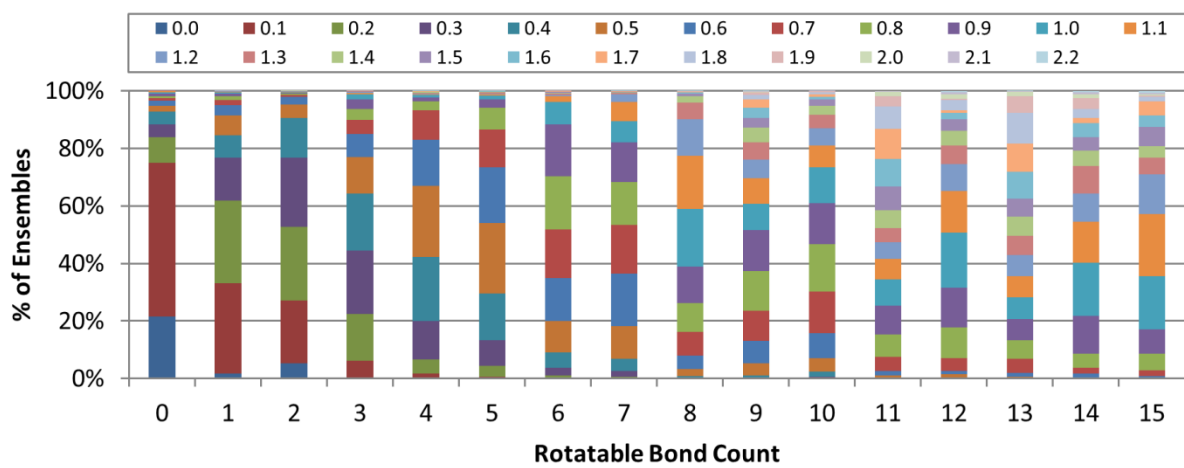


(b)

**Figure S1. Distribution of the RMSD accuracy of the conformer model vs. the non-hydrogen atom count.** The % distribution of the RMSD accuracy of the conformer model as a function of the non-hydrogen atom count: (a) before clustering and (b) after clustering.

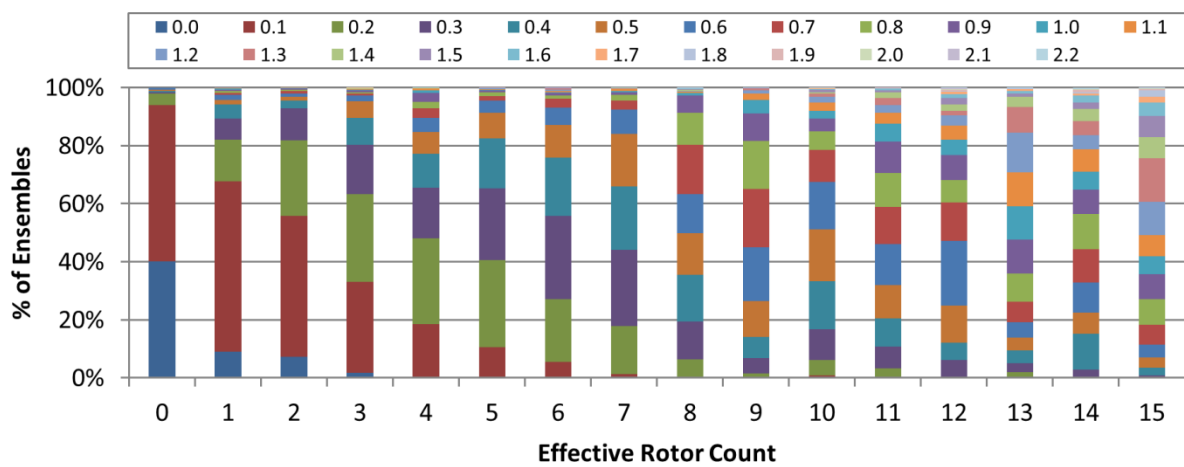


(a)

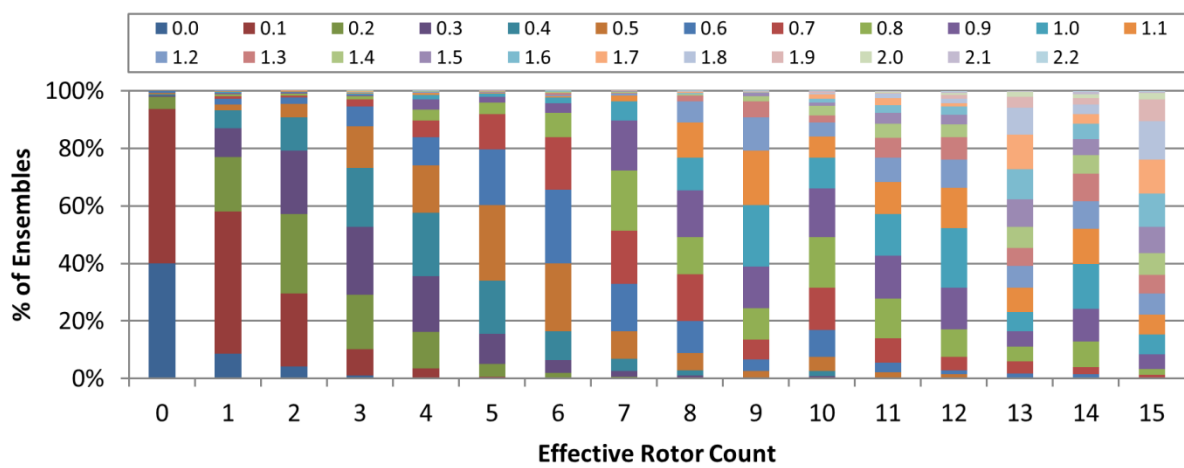


(b)

**Figure S2. Distribution of the RMSD accuracy of the conformer model vs. the rotatable bond count.** The % distribution of the RMSD accuracy of the conformer model as a function of the rotatable bond count: (a) before clustering and (b) after clustering.

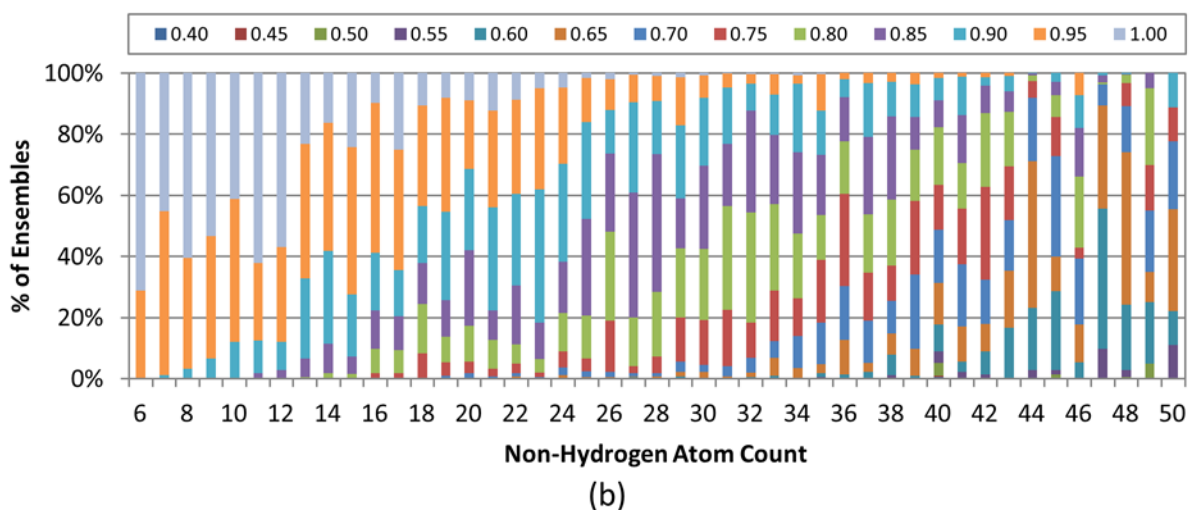
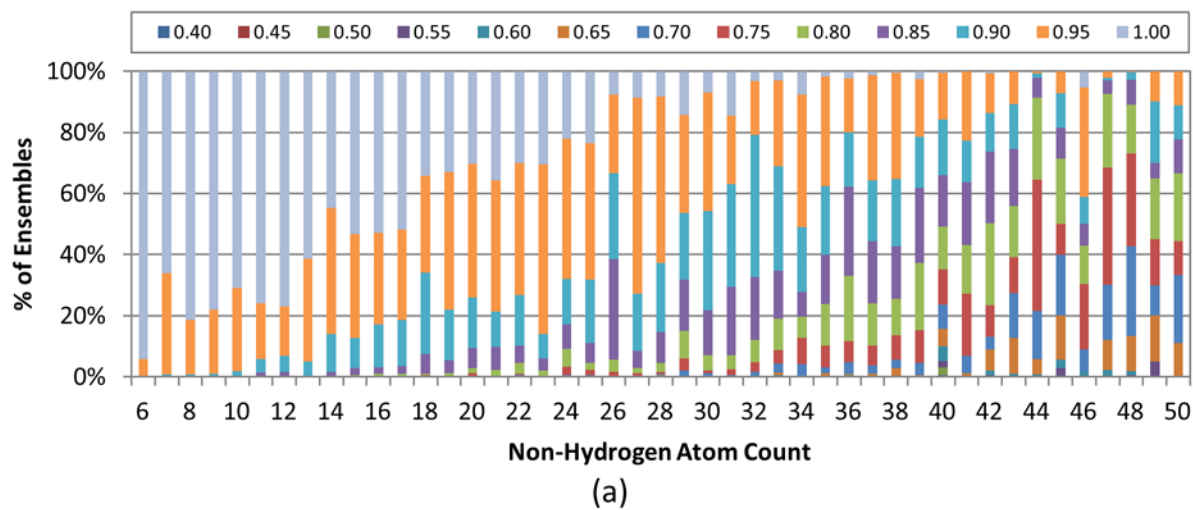


(a)

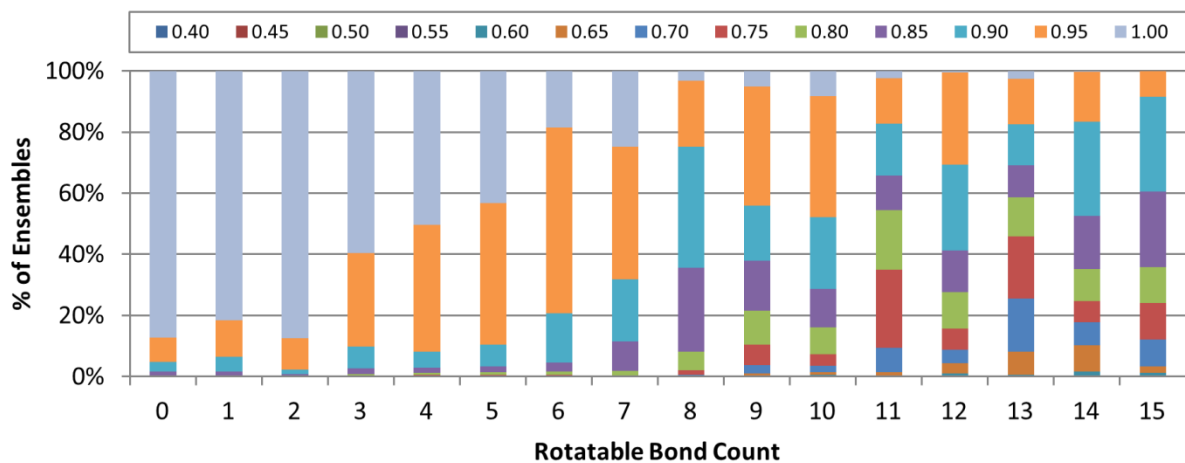


(b)

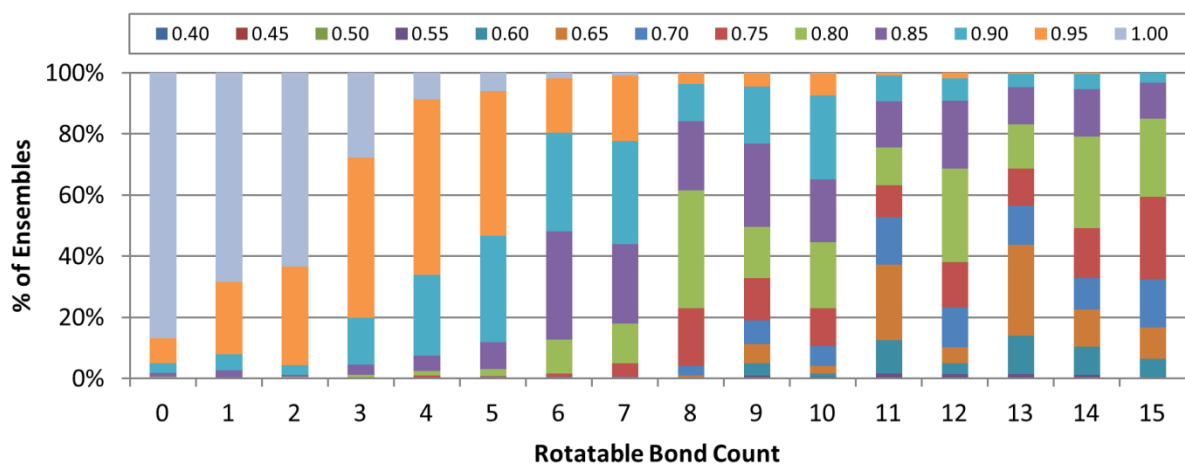
**Figure S3. Distribution of the RMSD accuracy of the conformer model vs. the effective rotor count.** The % distribution of the RMSD accuracy of the conformer model as a function of the effective rotor count: (a) before clustering and (b) after clustering.



**Figure S4. Distribution of the shape-optimized shape-Tanimoto ( $ST^{ST-opt}$ ) accuracy of the conformer model vs. the non-hydrogen atom count.** The % distribution of the  $ST^{ST-opt}$  accuracy of the conformer model as a function of the non-hydrogen atom count: (a) before clustering and (b) after clustering.

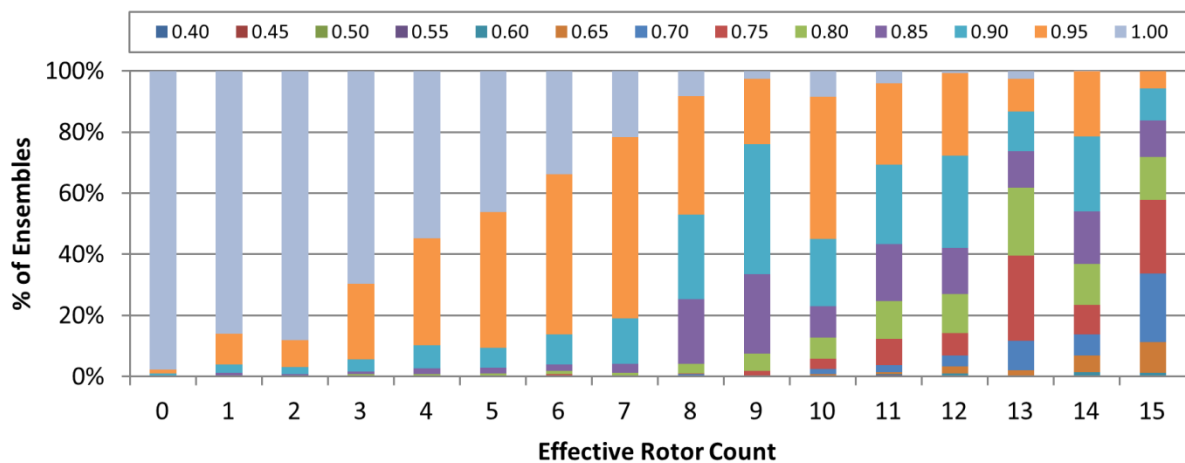


(a)

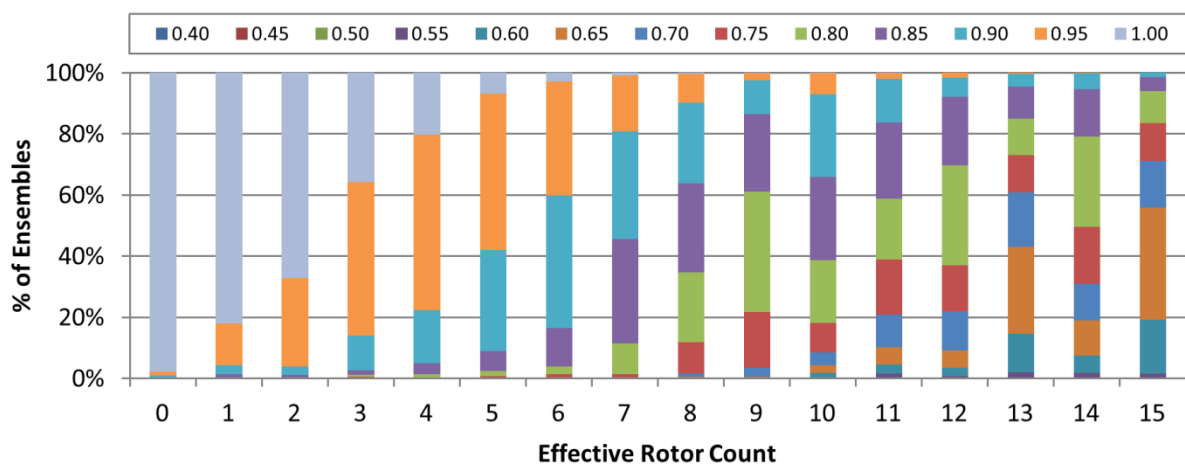


(b)

**Figure S5. Distribution of the shape-optimized shape-Tanimoto ( $ST^{ST-opt}$ ) accuracy of the conformer model vs. the rotatable bond count.** The % distribution of the  $ST^{ST-opt}$  accuracy of the conformer model as a function of the rotatable bond count: (a) before clustering and (b) after clustering.

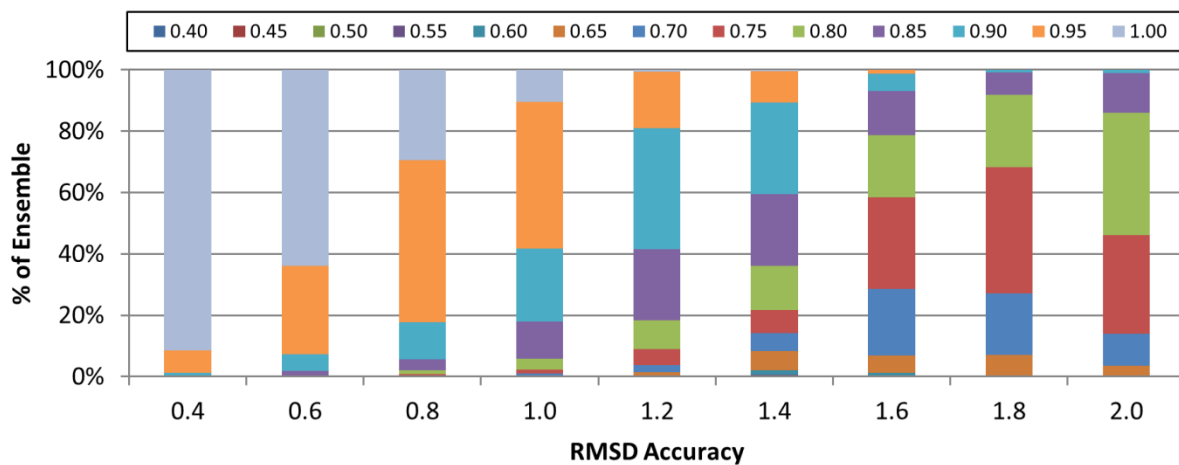


(a)

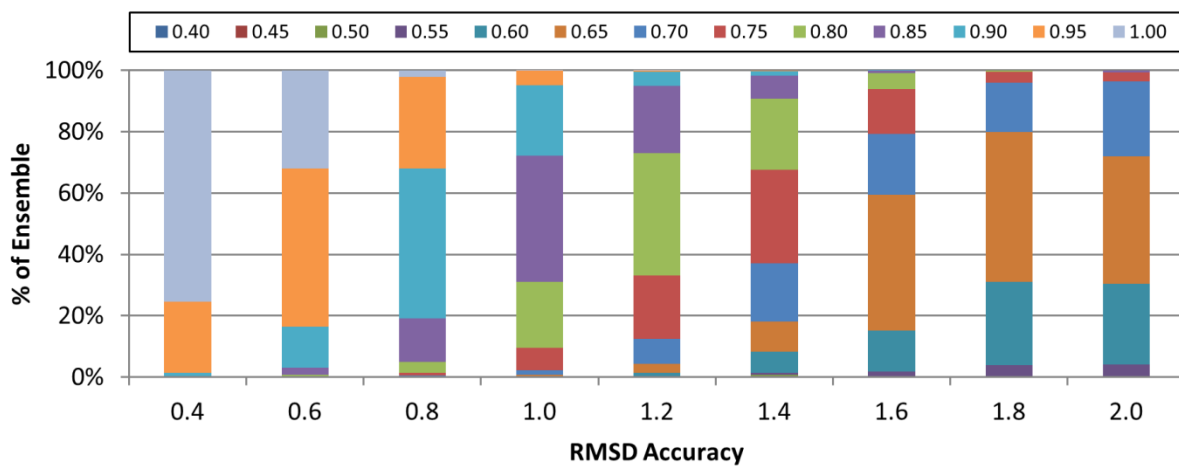


(b)

**Figure S6. Distribution of the shape-optimized shape-Tanimoto ( $ST^{ST-opt}$ ) accuracy of the conformer model vs. the effective rotor count.** The % distribution of the  $ST^{ST-opt}$  accuracy of the conformer model as a function of the effective rotor count: (a) before clustering and (b) after clustering.



(a)



(b)

**Figure S7. Correlation between the RMSD accuracy and shape-optimized shape-Tanimoto ( $ST^{ST-opt}$ ) accuracy of the conformer models.** The % distribution of the conformer models as a function of the  $ST^{ST-opt}$  accuracy for a given RMSD accuracy: (a) before cluster and (b) after clustering.