Supporting Information

Comprehensive Interrogation on Acetylcholinesterase Inhibition by Ionic Liquids Using Machine Learning and Molecular Modeling

Jiachen Yan[†], Xiliang Yan[†, *], Song Hu[‡], Hao Zhu[§], Bing Yan[†, ‡, *]

[†]Institute of Environmental Research at Greater Bay Area, Key Laboratory for Water Quality and Conservation of the Pearl River Delta, Ministry of Education, Guangzhou University, Guangzhou 510006, China

[‡]School of Environmental Science and Engineering, Shandong University, Qingdao 266237, China

[§]The Rutgers Center for Computational and Integrative Biology, Camden, New Jersey 08102, USA

*Corresponding authors

Xiliang Yan, E-mail: yanxiliang1991@gzhu.edu.cn

Bing Yan, E-mail: drbingyan@yahoo.com

Number of pages: 9
Number of figures: 8
Number of tables: 8

Table of Contents

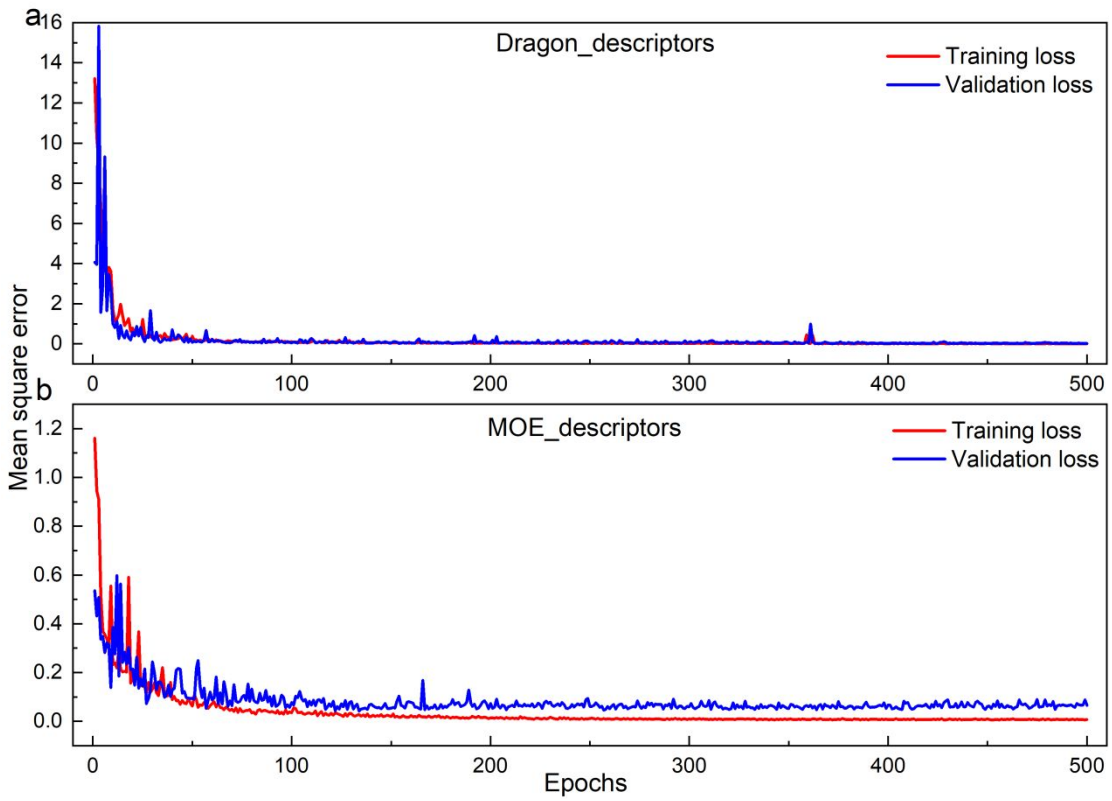

Figure S1. Training loss and validation loss against epochs. The mean square error of the predictions on the training set and the test set against iterations of the training procedure using dragon descriptors (a) and MOE descriptors (b).

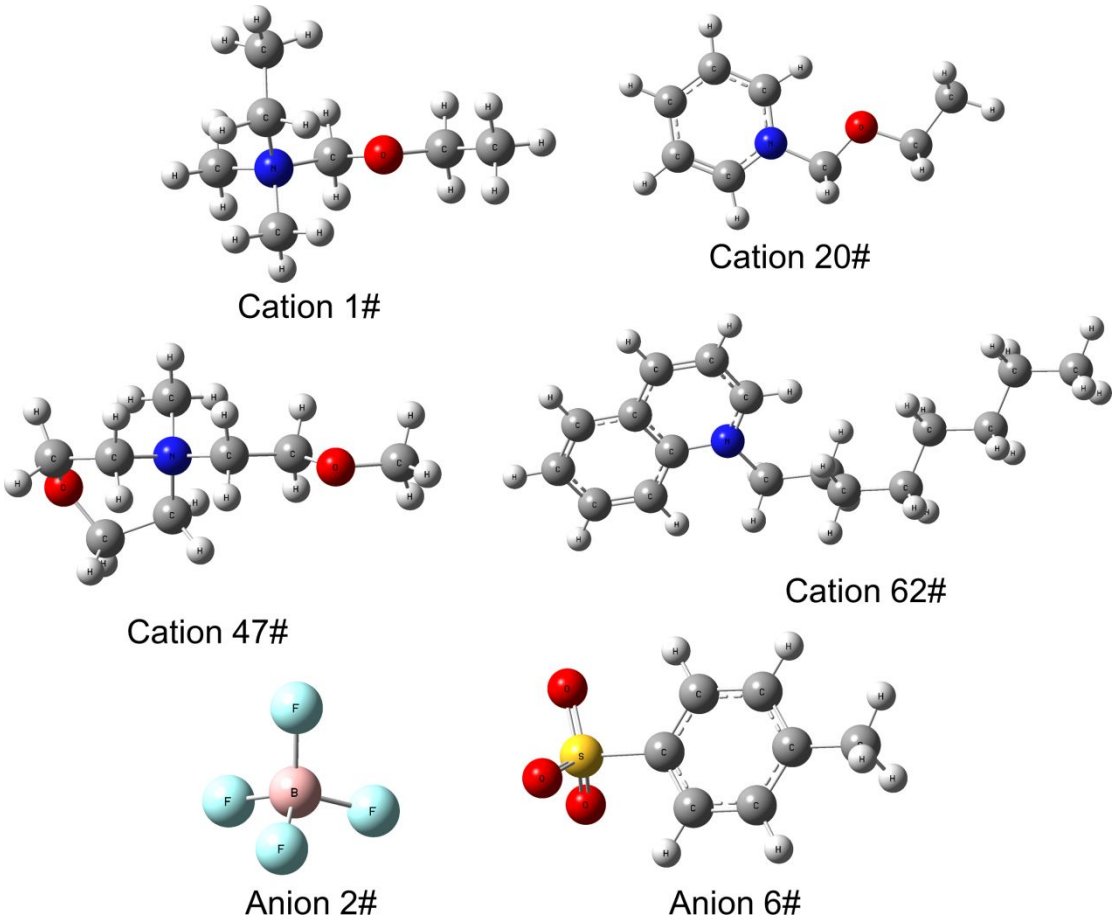

Figure S2. **Some representative optimized structures generated from DFT calculations.** All cations and anions were optimized using B3LYP (Becke-3 Parameter-Lee-Yang-Parr) hybrid functional in combination with 6-31+G(d,p) basis set.

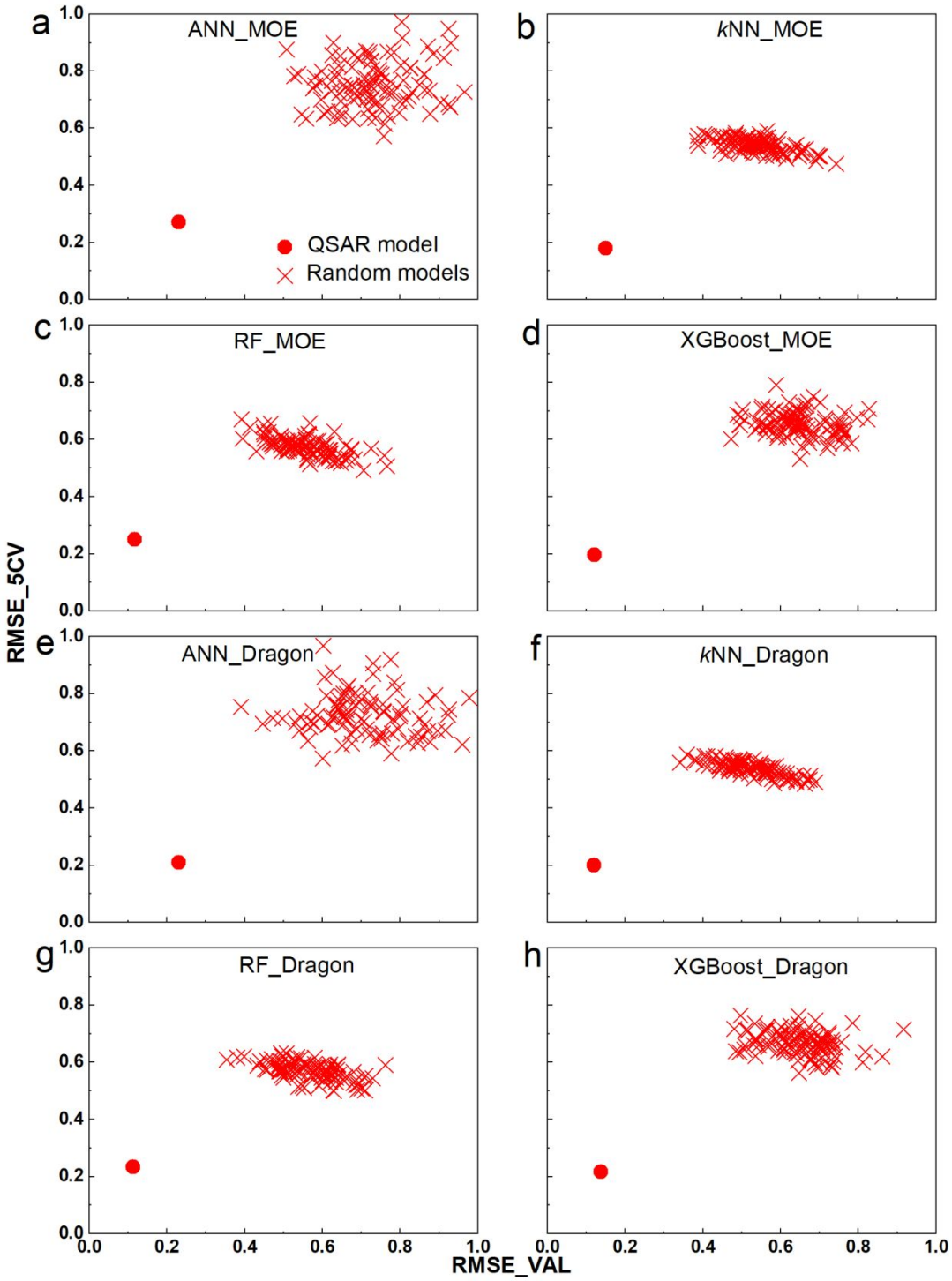

Figure S3. Model performances of y-scrambling permutation tests. The results of y-scrambling permutation tests using MOE descriptors (a-d) and Dragon descriptors (e-h). The red circles represent the results from QSAR models, while the red crosses represent the results from random models that molecular features remain the same but toxicity values undergo different permutations.

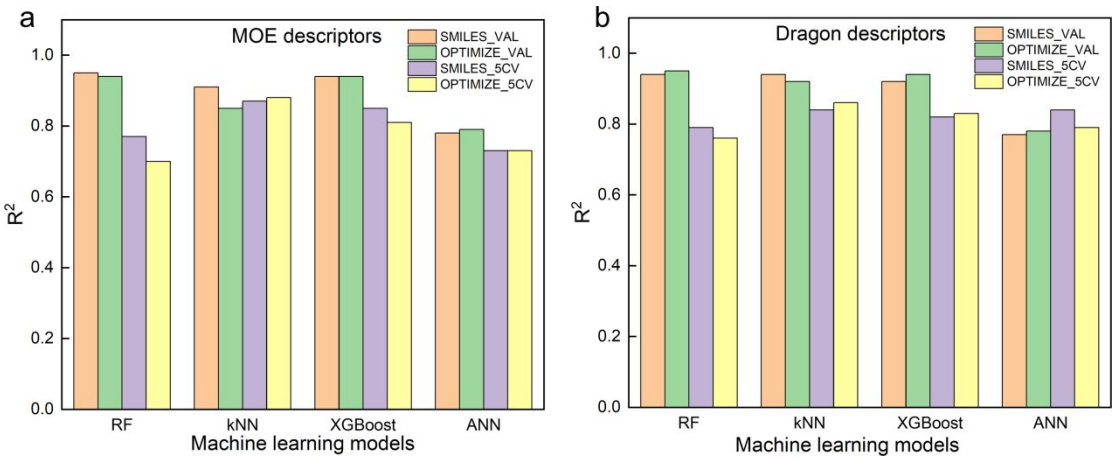

Figure S4. Comparison of model performances using SMILES-based descriptors and optimized structure descriptors. The MOE descriptors (a) and Dragon descriptors (b) were respectively calculated from SMILES representation and optimized structures, and then used for building machine learning models. The model performance was evaluated by determination coefficients (R^2) from five-fold cross validation (*i.e.*, SMILES_5CV and OPTIMIZE_5CV) and external validation (*i.e.*, SMILES_VAL and OPTIMIZE_VAL).

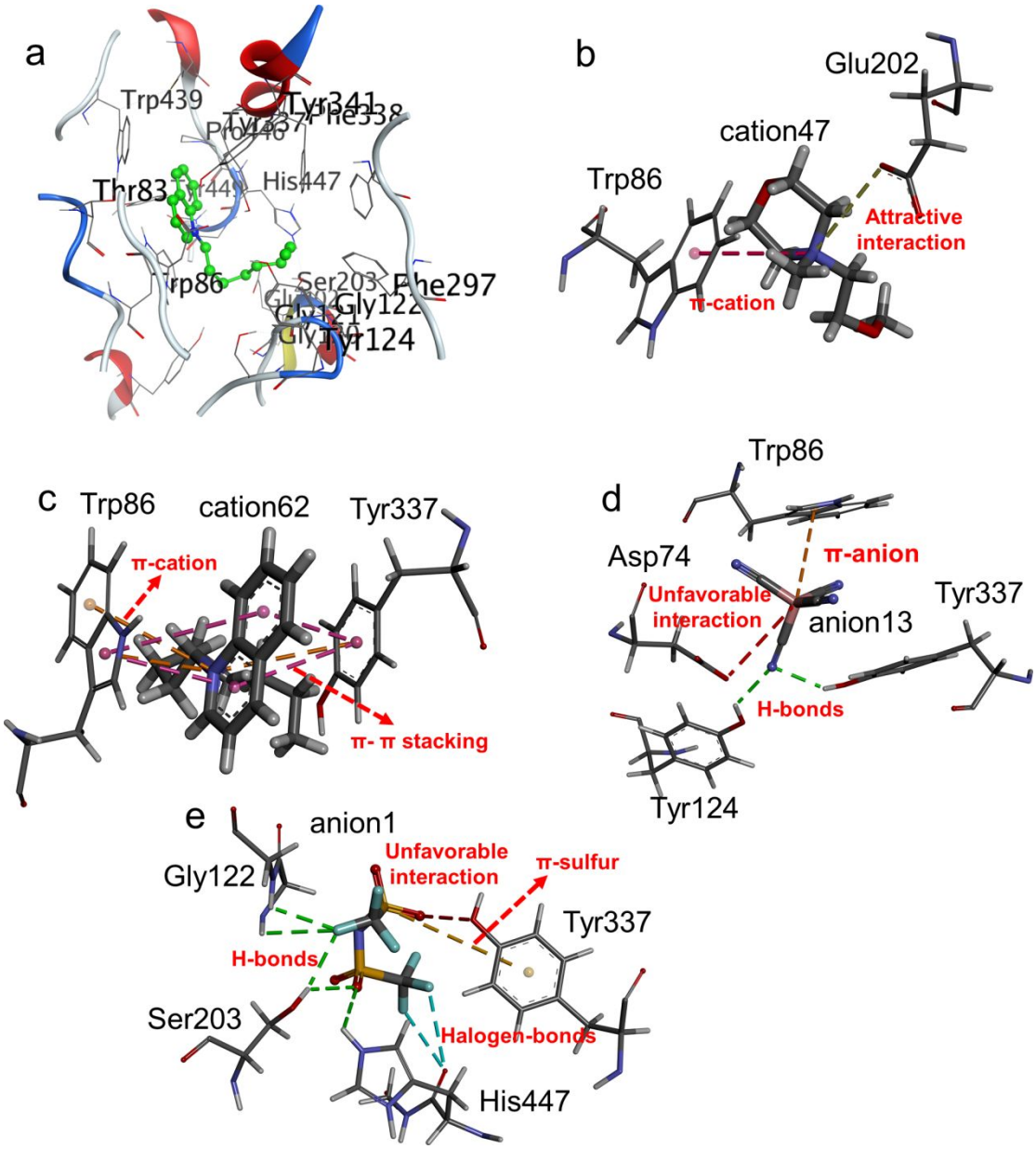

Figure S5. The pocket site and binding modes of some representative ionic liquids. (a) The pocket site was set according to the location of the co-crystal ligand. (b-e) Binding modes of some representative ILs (i.e., cation47, cation62, anion13, and anion1) in the binding pocket. Key residues for the interactions between ILs and AChE were shown in stick.

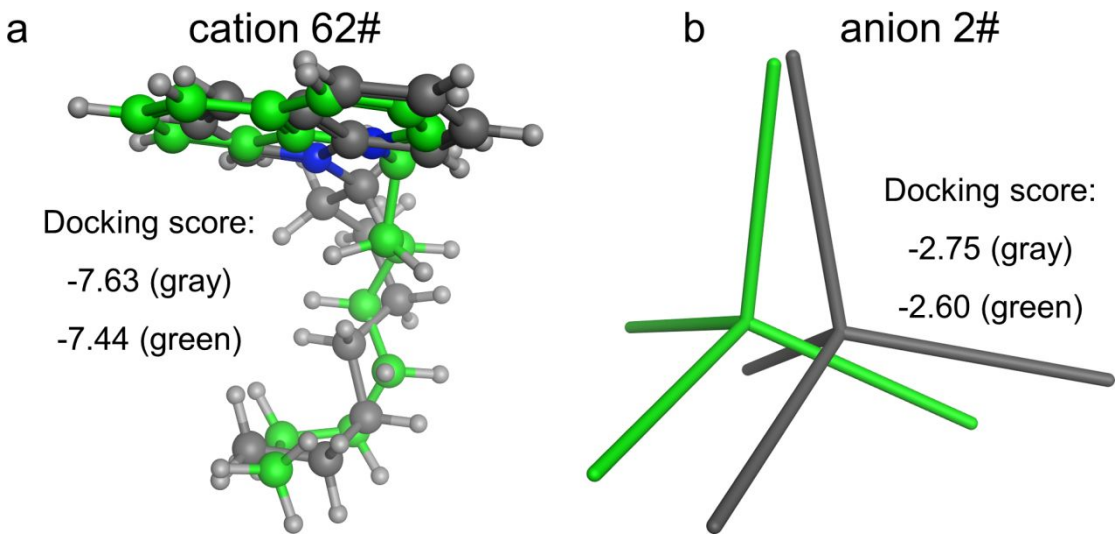

Figure S6. **Comparison of docking results using different atomic charges.** The conformation of IL153 (*i.e.*, cation62 and anion2) with the best docking score that was generated using the atomic charges assigned by MOE (gray) and CHELPG scheme (green). The more negative the docking value is, the stronger the molecule can be bound to the protein.

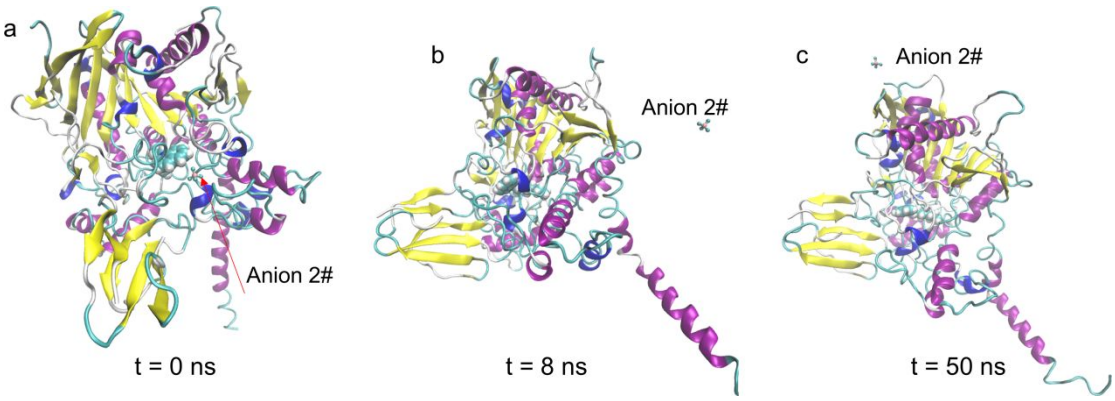

Figure S7. Snapshots of molecular dynamics trajectory of IL153-protein complex using CHELPG charges. Representative snapshots from the molecular dynamics trajectory of IL153-AChE complex at 0 ns (a), 5 ns (b), and 50 ns (c). The protein was shown in NewCartoon drawing method, cation62 were draw in VDW drawing method, anion2 was displayed in CPK drawing method. Water and ions were not shown for clarity. All images were rendered using the VMD software.

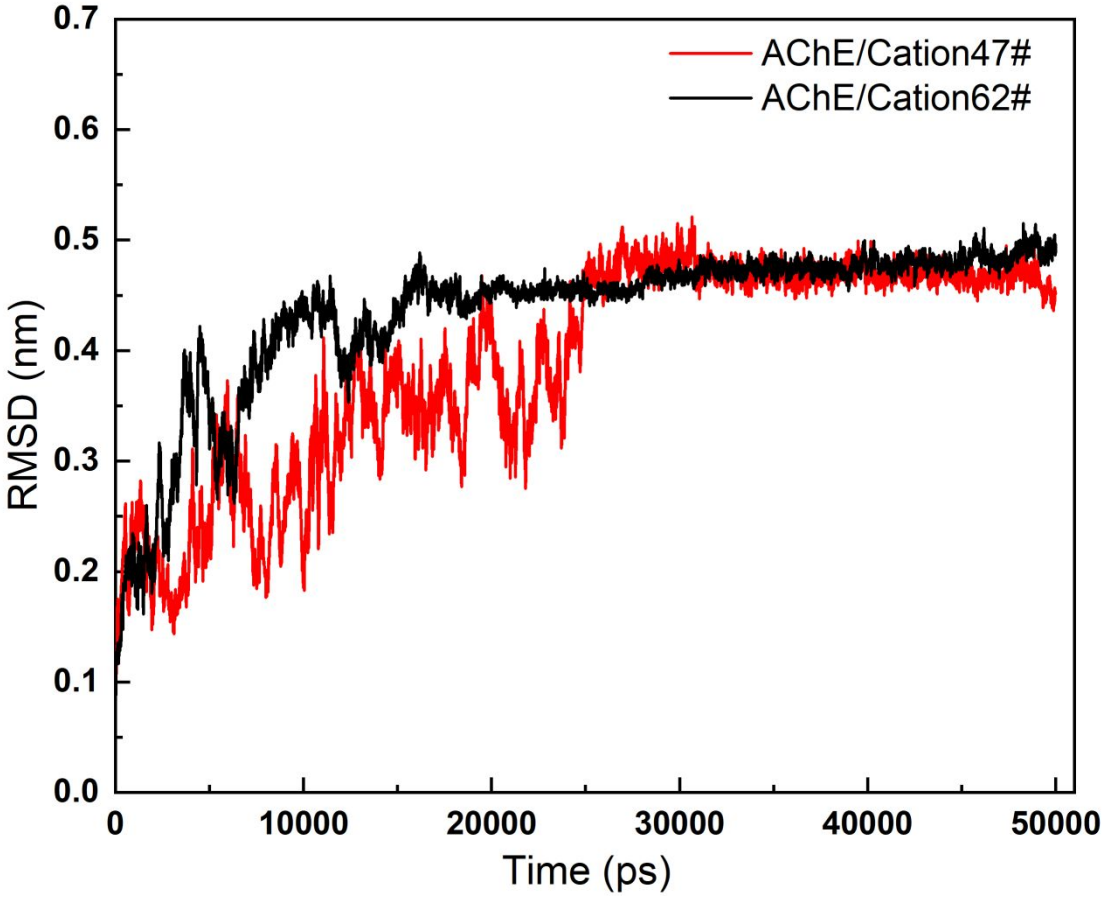

Figure S8. Backbone RMSD analysis during the MD simulations. Time-dependent RMSD values of the backbone atoms (Cα, N, C) of the AChE/Cation47 (red) and AChE/Cation62 (black) complexes.