Supporting Information for “Role of Guanidinium-Carboxylate Ion Interaction in Enzyme Inhibition with Implication for Drug Design”

Aswathy N. Muttathukattil,[†] Sriraksha Srinivasan,[‡] Antarip Halder,[†] and Govardhan Reddy^{*,†}

[†]*Solid State and Structural Chemistry Unit, Indian Institute of Science, Bengaluru, Karnataka, India 560012*

[‡]*Department of Chemistry, St. Joseph's College, Bengaluru, Karnataka, India 560027*

E-mail: greddy@iisc.ac.in

Phone: +91-80-22933533. Fax: +91-80-23601310

Benchmarking of Various DFT Functionals: To select the most reliable DFT functional to study the interaction between guanidinium cation (Gdm^+) and carboxylate anion, we have obtained the optimized geometry (in gas phase) and interaction energy (in CPCM implicit solvent model) of the Gdm^+ -carboxylate complex using six different DFT functionals: (1) B3LYP,^{1,2} (2) B3LYP-D3(BJ),³ (3) M05-2X,⁴ (4) ω B97XD,⁵ (5) TPSS⁶ and (6) PBE0.⁷ The performance of these functionals is benchmarked against the results obtained using MP2⁸ level of theory. All the calculations are performed using the 6-31G+(d,p) basis set. The interaction energy of the Gdm^+ -carboxylate dimer calculated using the DFT-D3(BJ) functional is the closest to the value obtained from MP2 calculation (Figure S1). The long-range dispersion corrections, as implemented in Grimme's DFT-D/D2/D3 formalism, are important for these systems. They are perhaps more important than the middle range dispersion corrections since the hybrid meta-GGA functional M05-2X, which is parameterized to take care of the middle range dispersion corrections, shows relatively poor performance.

Figure S1: Interaction energy of Gdm^+ -carboxylate complex calculated using different levels of theory.

Interaction with carboxylate ions modifies the charge distribution in Gdm^+ : To understand the inter-molecular interactions between Gdm^+ and carboxylate ions in detail, we have obtained the ground state optimized geometries of these multi-ionic complexes observed in MD simulation using DFT (Figure 3 in main manuscript). DFT calculations show that all the three possible resonance canonical structures (or valence bond structures) of Gdm^+ (annotated as I, II and III in Figure 3A) contribute equally to the ground state of the monomer (Figure 3B) as all the three C-N bonds have equal length. This leads to a nearly sp^2 hybridization state for all the three nitrogen atoms. NBO analysis shows that all the natural bond orbitals that form the N-H σ -bonds have 30.5% s- and 69.5% p-character. However, contributions of these three structures do not remain homogeneous in the bidentate configuration of Gdm^+ -acetate pair (Figure 3C). Contribution of one resonance canonical structure decreases, since the amino group that does not participate in hydrogen bond formation shows a longer C-N bond compared to the other two amino groups. We observed a change in the hybridization state of the corresponding nitrogen from nearly sp^2 to sp^3 , which can be quantified by (a) increase in the pyramidalization of the amino group and (b) decrease in the percentage s-character of the natural bond orbitals that participate in the formation of N-H σ bonds (Figure S5). Similar changes are also observed in Gdm^+ -carboxylate- Gdm^+ trimer (Figure 3D) but to a smaller extent (Figure S5D). Interestingly, in the carboxylate- Gdm^+ -carboxylate bridging interaction, contribution of a resonance canonical structure of Gdm^+ dominates over the other two. Consequently, we observe that the nitrogen, which forms hydrogen bonds with both the carboxylate ions, remains nearly sp^2 hybridized with shorter C-N bond and planar amino group whereas, the other two amino nitrogens show a noticeable decrease in the percentage s-character in their corresponding natural bond orbitals. This results in elongated C-N bond and pyramidalized amino groups (Figure S5). Such typical charge redistribution results in stronger N-H $\cdots$ O bonds compared to the ones present in Gdm^+ -carboxylate- Gdm^+ trimer (Figure S5). Thus we observe that the carboxylate- Gdm^+ -carboxylate bridging interaction is significantly stronger ($E^{int} = -36.1$

kcal mol⁻¹) than the Gdm⁺-carboxylate-Gdm⁺ trimer ($E^{int} = -32.6$ kcal mol⁻¹), which has the same number of hydrogen bonds.

Figure S2: Optimized geometry of a carboxylate anion interacting with a sodium cation (Na⁺) obtained using B3LYP-D3(BJ)/6-311+G(d,p) level of theory. Important inter-atomic distances are reported in Å. The interaction energy of the complex is -12.9 kcal/mol.

S

Figure S3: Possible interaction patterns between Gdm⁺ and carboxylate ion (I - V). Intermolecular hydrogen bonds are shown as broken lines. Only conformation-I (having two intermolecular hydrogen bonds) converges to a stable cation-anion pair on optimization in gas phase using B3LYP-D3(BJ)/6-311+G(d,p) level of theory. All other conformations (having bifurcated or single hydrogen bond) lead to deprotonation of the cation and subsequent protonation of the anion.

Figure S4: Free energy change associated with the interaction of two propionate ions, and one acetate and one propionate ion in 0.5 M GdmCl solution. Both the dimer complexes are stable in the presence of Gdm⁺.

Figure S5: Gas phase optimized geometries of (A) isolated Gdm⁺, (B) isolated carboxylate anion, (C) bidentate configuration of Gdm⁺-carboxylate pair, (D) two Gdm⁺ interacting with a single carboxylate anion via bifurcated hydrogen bonds and (E) single Gdm⁺ bridging between two carboxylate anions. Broken straight lines indicate hydrogen bonds. In Gdm⁺, hybridization states of the amino nitrogens are quantified using (a) the average % s-character of the natural bond orbitals (NBO) that form the N-H σ bonds and (b) improper dihedral angle constituted by the central carbon, the nitrogen and the two hydrogen atoms (Ψ_N). Similarly, for the carboxylate ions the improper dihedral angle constituted by the carbon atom of the methyl group, the central carbon atom and the two oxygen atoms are reported (Ψ_C). Schematic representation of the definition of an improper dihedral angle is shown in (F). It is the angle between the planes defined by (i) r_{12} and r_{23} and (ii) r_{23} and r_{24} and hence quantifies the pyramidalization of the central atom (r_{ij} is the position vector connecting atom 1 and 2). All the dihedral angles are reported in degrees. For all the amino nitrogens that are involved in inter-ionic hydrogen bond formation, we have reported the red shift ($\Delta\nu$) in the N-H bond stretching frequency due to hydrogen bond formation in cm⁻¹. Larger red shift indicates stronger hydrogen bonds.

Figure S6: Structure of (A) T4 lysozyme (B) Hen egg white lysozyme (C) HIV protease (D) Pepsin and (E) Creatine kinase obtained in the simulation after a time period of 150 ns in 0.5 M GdmCl simulation (in blue) aligned with their respective PDB structure (in grey). The protein did not unfold within 150 ns. The active residues in the simulation and PDB structures are shown in red and green, respectively. Only small deviations are observed in the position of active residues.

Figure S7: Free energy profiles for the unbinding of Gdm^+ from HIV protease active site obtained from three independent ABF calculations. The free energy is projected onto the distance r between the center of mass of C_α active residues (Asp25 to Asp25') and center of mass of Gdm^+ . The free energy change is ≈ 15 kcal/mol indicating that Gdm^+ binds strongly in the protein active site.

Figure S8: Structures from the PDB show the binding of Gdm^+ part of the drug Zamavirin to the active site of (A) Influenza virus B/BEIJING/1/87 neuraminidase (B) N8 neuraminidase (C) Anhui H1N9 neuraminidases bridging two carboxylate residues. The protein structures are shown in cartoon representation and in grey. The specific interaction between Gdm^+ (blue) and carboxylate ions (red) is highlighted. The rest of the Zamavirin drug is shown in green.

Figure S9: Structures from the PDB where Gdm^+ bridging two carboxylate residues is observed. The protein structures are shown in cartoon representation and in grey. The specific interaction between Gdm^+ (blue) and carboxylate ions (red) is highlighted. (A) Chaperone protein ClpB nucleotide binding domain 1 (B) Murine adenosine deaminase (C) Chaperone protein ClpB nucleotide binding domain 2 (D) Topoisomerase V.

Figure S10: (A) Spatial density map of Gdm⁺ around T4 lysozyme when Asp11 is protonated. The density of Gdm⁺ near the active residues Asp11 and Glu20 decreases. The maximum occupancy of Gdm⁺ (> 0.8) shown in blue is now observed near the residues His31, Asp70 and Phe104. (B) Spatial density map of Gdm⁺ around Hen Egg White Lysozyme in 2 M GdmCl and it is exactly similar to that computed in 0.5 M GdmCl (Figure 5A in main manuscript). (C) Spatial density map of Gdm⁺ around Hen Egg White Lysozyme in 2 M GdmCl computed using OPLS force field for protein,^{9,10} SPC/E model for water¹¹ and GdmCl parameterized using the Kirkwood-Buff equations.¹² Gdm⁺ occupancy in the active site is similar to the occupancy in the density map obtained using the CHARMM force field. However, in the density map obtained using OPLS force field, occupancy of non-bridging Gdm⁺ is also high.

References

- (1) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (2) Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (3) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (4) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
- (5) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (6) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (7) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (8) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618–622.

- (9) Jorgensen, W.; Maxwell, D.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (10) Kaminski, G.; Friesner, R.; Tirado-Rives, J.; Jorgensen, W. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (11) Berendsen, H.; Grigera, J.; Straatsma, T. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (12) Weerasinghe, S.; Smith, P. A Kirkwood-Buff derived force field for the simulation of aqueous guanidinium chloride solutions. *J. Chem. Phys.* **2004**, *121*, 2180–2186.