TUPÃ: Electric field analyses for molecular simulations

Marcelo D. Polêto*[1] and Justin A. Lemkul†[1,2]

[1]Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061, United States
[2]Center for Drug Discovery, Virginia Tech, Blacksburg, VA 24061, United States

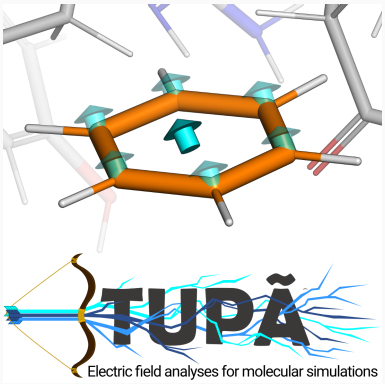
April 1, 2022

Abstract

We introduce TUPÃ, a Python-based algorithm to calculate and analyze electric fields in molecular simulations. To demonstrate the features in TUPÃ, we present three test cases in which the orientation and magnitude of the electric field exerted by biomolecules help explain biological phenomena or observed kinetics. As part of TUPÃ, we also provide a PyMOL plugin to help researchers visualize how electric fields are organized within the simulation system. The code is freely available and can be obtained at `https://mdpoleto.github.io/tupa/`.

Keywords: Electric field, electrostatics, molecular dynamics, force fields, molecular mechanics

*mdpoleto@vt.edu
†jalemkul@vt.edu



Here we present TUPÃ, a Python-based algorithm to analyze electric fields in molecular simulations. The software was designed to be flexible and allow users to investigate electric fields orientation and magnitude on atoms, bonds, or even in absolute coordinates in a simulation box throughout the trajectory. A PyMOL plugin to visualize the electric fields alongside molecules is provided. We expect TUPÃ to aid in efforts to analyze electrostatics-driven events in molecular simulations.

INTRODUCTION

Electrostatic interactions dominate many aspects of macromolecular dynamics, frequently giving rise to important molecular phenomena, such as ion transport[1,2] or enzyme catalysis[3,4]. The molecular mechanisms behind such processes are subject of multiple investigations in both experimental and computational fields. Molecular dynamics (MD) simulations have been commonly used to study molecular interactions and structural features of biomolecules that are required to trigger electrostatic-driven events.

By taking advantage of the classical molecular mechanics (MM) description of particles and their assigned partial charges, MD engines compute short-range electrostatic forces between N atoms in a system using Coulomb's Law. Analogously, the electric field ($\vec{E}$) exerted by N atoms at a given point in the simulation box can be computed according to Equation 1:

$$\vec{E}_{(x,y,z)} = \sum_{i=1}^{N} \frac{1}{4\pi\varepsilon_0} \frac{Q_i}{r^2} \cdot \hat{r}_{(x,y,z)} \tag{1}$$

in which ε_0 is the electric permittivity, Q_i is the partial charge of atom i, r is the distance between atom i and the point in space and $\hat{r}$ is the unit vector of the distance. Thus, by defining a point in space and iterating over N atoms in each trajectory frame, one can obtain the magnitude and directionality of electric fields in a simulation box.

Understanding how biomolecules create suitable electrostatic environments for many biological purposes can be a powerful tool for computational biophysicists. To date, only a few studies have calculated the electric field generated by biological systems or investigated how electric field are oriented and correlated with biological processes of interest[5–15]. More recently, Nash et al.[16] developed an algorithm that

calculates the electric field exerted on atoms involved in a covalent bond by using the Tinker[17] MD package in the context of the AMOEBA polarizable force field[18–20]. Aiming to facilitate such studies while using different force fields and MD packages, we present *TUPÃ*, a Python-based algorithm to calculate and analyze electric fields in molecular simulations. Importantly, TUPÃ does not make any assumption about the underlying force field used for the MD simulation; as long as discrete particles are present to which charges are assigned, electric fields can be calculated.

ALGORITHM STRUCTURE

In TUPÃ, we apply the concept of atom-centered partial charges commonly used in MD simulations to calculate the electric field that a set of atoms (from now referred as *environment set*) exert on a given Cartesian point (from now referred as *probe set*) in the simulation box.

Our code is based on Python3[21] syntax and it employs the MDAnalysis[22,23] package to handle MD trajectory and topology files. Therefore, all file formats supported by MDAnalysis should be compatible with our workflow. Also, selection of atoms to establish the environment and probe sets is based on the selection syntax in MDAnalysis.

For simplicity, TUPÃ can be executed by simply providing a configuration file, a topology file, and a trajectory file via the command-line, as shown below:

```
$ python TUPA.py -top topology_file -traj trajectory_file -config
    configuration_file -outdir Results
```

In the configuration file, the user defines the calculation mode (see below), the set of atoms composing the environment and the atoms/coordinates to be used as

probes. A sample configuration file is described in Supplementary Materials, but users can conveniently create one via the command-line:

```
$ python TUPA.py -template config_sample.dat
```

Note that any atoms in the system can be included in the environment set, including proteins, lipids, carbohydrates, nucleic acids, etc. For highly dynamic molecules, such as water and ions, users can choose to include their contribution in the calculation set using a radial cutoff around the environment. Doing so employs the *around* selection function of MDAnalysis, allowing the selection of atoms within the cutoff, including image atoms that are present as a consequence of periodic boundary conditions (PBC).

TUPÃ has three calculation modes: I) **ATOM MODE**, in which the position of a selected atom is tracked and used as probe to calculate $\vec{E}$; II) **BOND MODE**, in which the midpoint of 2 selected atoms is tracked and is used as probe to calculate $\vec{E}$ and its alignment to the defined bond axis; and III) **COORDINATE MODE**, in which a fixed (X,Y,Z) coordinate is used as probe throughout the entire trajectory.

The general workflow of TUPÃ is presented in Figure 1. First, the configuration file is parsed and parameters are stored, including the calculation mode. After, topology and trajectory files are read to create an MDAnalysis universe. Iterating over each trajectory frame, the coordinates of the environment and probe atoms are updated, making it possible to iterate over each atom in the environment set to calculate their X, Y and Z contributions for $\vec{E}$ and store them as arrays. After, the total vector contributions are calculated and used to obtain the vector $\vec{E}$ and its magnitude. Exclusively for BOND MODE calculations, TUPÃ also computes the projection of $\vec{E}$ onto the defined bond axis via Equation 2:

$$\vec{E}_{proj} = \frac{\vec{E}_{total} \cdot \vec{r}_{bond}}{|\vec{r}_{bond}|} \tag{2}$$

where $\vec{r}_{bond}$ is the bond axis vector. By doing so, it becomes possible to calculate how much of $\vec{E}$ is actually aligned to the bond axis by simply evaluating:

$$\%aligned = \frac{|\vec{E}_{proj}| \times 100}{|\vec{E}_{total}|} \tag{3}$$

During the iteration loop over the environment atoms, TUPÃ keeps track of each atom's residue number, which allows it to calculate and store the partial contribution of each residue to the total $\vec{E}$ for each trajectory frame. Hence, TUPÃ also calculates and outputs the average contribution of each residue to $\vec{E}$ and their respective alignment with it. If BOND MODE is set, the alignment is calculated between each residue's contribution to $\vec{E}$ and the defined bond axis.

TUPÃ calculations were finished in a matter of minutes for the systems presented here (see Table S1) due to the computational efficiency of MDAnalysis and NumPy[24]. In addition, it is important to stress that even though our approach is compatible with any all-atom additive and polarizable force fields relying in a partial charge description, the accuracy of $\vec{E}$ is entirely dependent on the force field itself.

Visualizing electric fields

Visualizing the electric field vectors at a given position within the simulation box can be a useful tool to understand the configuration of the electrostatic forces and their impact on biological phenomena. Thus, we developed pyTUPÃmol, a PyMOL plugin that can be used to plot $\vec{E}$ calculated previously as 3D arrow objects and render them together with visualizations of the molecular systems. We note that

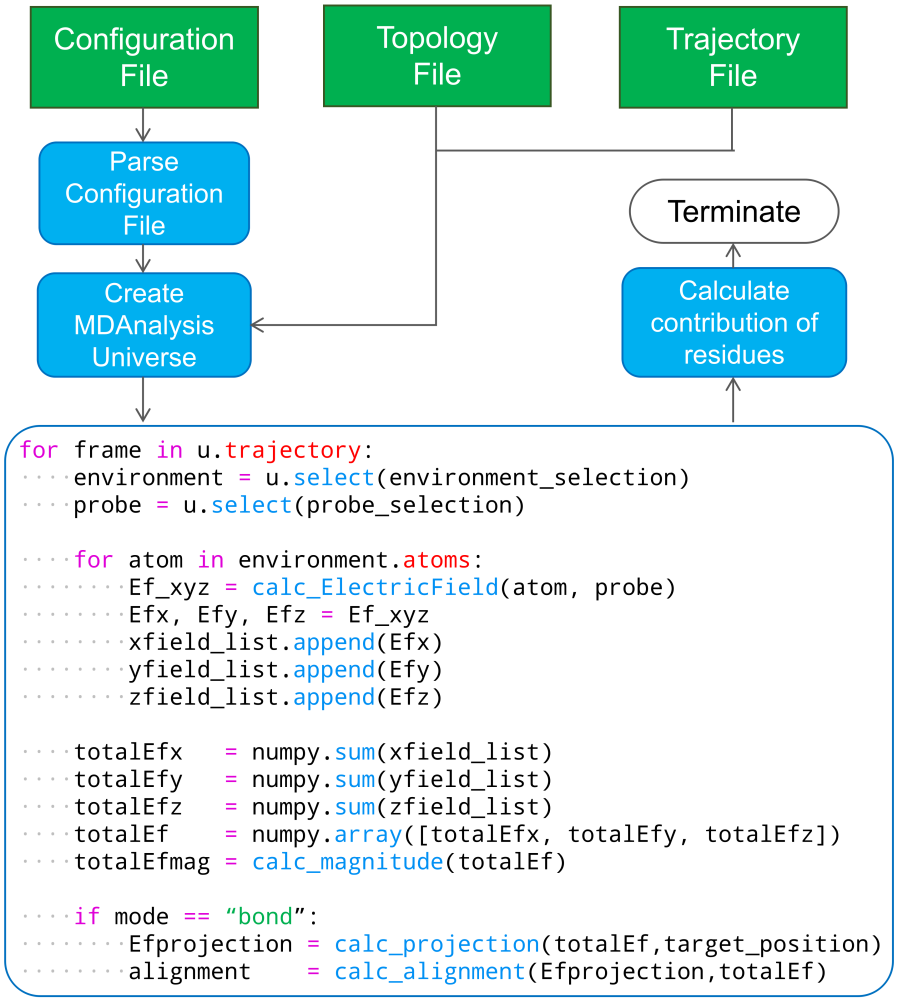


Figure 1: General workflow of TUPÃ and a pseudocode for its basic functions. After the configuration file is parsed, an MDAnalysis universe is built and used to extract atomic positions and charges to calculate $\vec{E}$ exerted on the probe. The contributions of each residue to $\vec{E}$ are also calculated.

all figures depicting the electric field orientation in this article were rendered with pyTUPÃmol and PyMOL. The plugin is freely available and it is distributed as part of TUPÃ code.

The plugin has three functions that can be called from the PyMOL command-line:

- *efield_point*: generates a 3D vector at the probe coordinate based on an array [$\vec{E}_X$, $\vec{E}_Y$, $\vec{E}_Z$]. Users can either provide a PyMOL selection for the probe using PyMOL syntax or an absolute (X,Y,Z) coordinate.

- *efield_bond*: generates a 3D vector similar to *efield_point*, but the probe position is defined halfway between two selected atoms comprising the bond axis.

- *draw_bond_axis*: generates a 3D vector between two selected atoms to visually define the bond axis used in TUPÃ. Users can either provide two PyMOL selections using PyMOL syntax or absolute (X,Y,Z) coordinates. By default, the direction is defined as from Selection1 towards Selection2.

TEST CASES

To demonstrate possible applications of TUPÃ in biological systems, we evaluated how the electric field organization plays an important role in three test cases. First, we characterized the electric field sensed by a benzene ligand bound to the T4 bacteriophage lysozyme (PDB 227L[25]) using the Drude-2019 polarizable force field[26]. Second, we calculated the well-studied electric field preorganization at the active site of the ketosteroid isomerase (PDB 5KP4[13]). We used the CHARMM36m force field[27] to study the electric field projected onto the C=O bond of 19-nortestosterone (19NT), a substrate-like molecule. Our third and last study case was the K^+-selective

NaK ion channel (PDB 3OUF[28]), also known as NaK2K, in which we investigated the electric field organization throughout the ion channel core and the electric force sensed by the coordinated K^+ ions. Detailed methods describing how systems were built can be found in Supplementary Material.

It is important to note that TUPÃ assumes that all trajectories have been recentered in a simulation box and that molecules in the environment set are not split across the box. As mentioned above, highly dynamic molecules can be selected taking the PBC into account. For the cases below, trajectories were recentered using the IMAGE facilities within the CHARMM software[29].

T4 lysozyme

As early studies of electric field effects on enzyme catalysis focused on enzymes such as egg-hen-white lysozyme and T4 bacteriophage lysozyme[30], we investigated the electric field exerted by T4 lysozyme on a benzene ligand bound to its active site (Figure 2A) throughout 100 ns of MD simulations. Here, we used the the polarizable Drude-2019 force field to demonstrate that polarizable force fields are suitable to our approach as long as they are based on the classic MM convention (partial charges assigned to atoms).

In our TUPÃ calculations, we included all lysozyme atoms and the SWM4-NDP water molecules[31] within a 10 Å cutoff radius from the benzene ligand in the environment set used, thus accounting for their contributions to the total electric field $\vec{E}$ (see Section 3.1 in the Supplementary Material). It is important to note that if the probe set has multiple atoms, TUPÃ uses the center of geometry (COG) of the selection for the electric field calculation. As shown in Figure 2B, the time series of $\vec{E}$ magnitude fluctuates around an average of 44 $\pm$ 19 MV/cm at the benzene

COG. Although a stable field was observed, it is important to mention that benzene's dynamics within the binding site are fast and multiple rotation events were observed. Regardless, the ligand stayed within the crystallographic binding site and the orientation of the electric field was maintained during the entire simulation. In fact, when we used the same environment set (lysozyme and waters within 10 Å of benzene) to calculate $\vec{E}$ sensed by each benzene carbon atom, we observe the same field orientation (Figure 2C).

One of the main features implemented in TUPÃ is that it can calculate the contribution of each residue ($\vec{E}_{res}$) to the total $\vec{E}$, together with the alignment between $\vec{E}_{res}$ and $\vec{E}$. This feature allows for further investigations into the effectiveness of each contribution and might prove useful when designing point mutations. In Figure 2D, we show $\vec{E}_{res}$ for T4 lysozyme, highlighting that residues Glu11, Asp70, Gln105 and Arg145 were identified as the major contributions to the total $\vec{E}$ with magnitudes of 49 $\pm$ 7 MV/cm, 38 $\pm$ 6 MV/cm, 15 $\pm$ 6 MV/cm, and 28 $\pm$ 5 MV/cm, respectively. Note that these contributions may have different directionalities and, therefore, do not add up to the $\vec{E}$ average value. To study the directionality of each residue contribution, the user may compose the environment set with only the residue of interest.

Ketosteroid Isomerase (KSI)

The electric field model of catalysis, in which a preorganized electric field exerted by the enzyme facilitates catalysis, has been studied for many years[32–34]. Particularly, ketosteroid isomerase in complex with 19-nortestosterone has been the subject of many studies in this regard due to its low degree of conformational rearrangement necessary for catalysis. The Boxer group demonstrated that KSI exerts a large

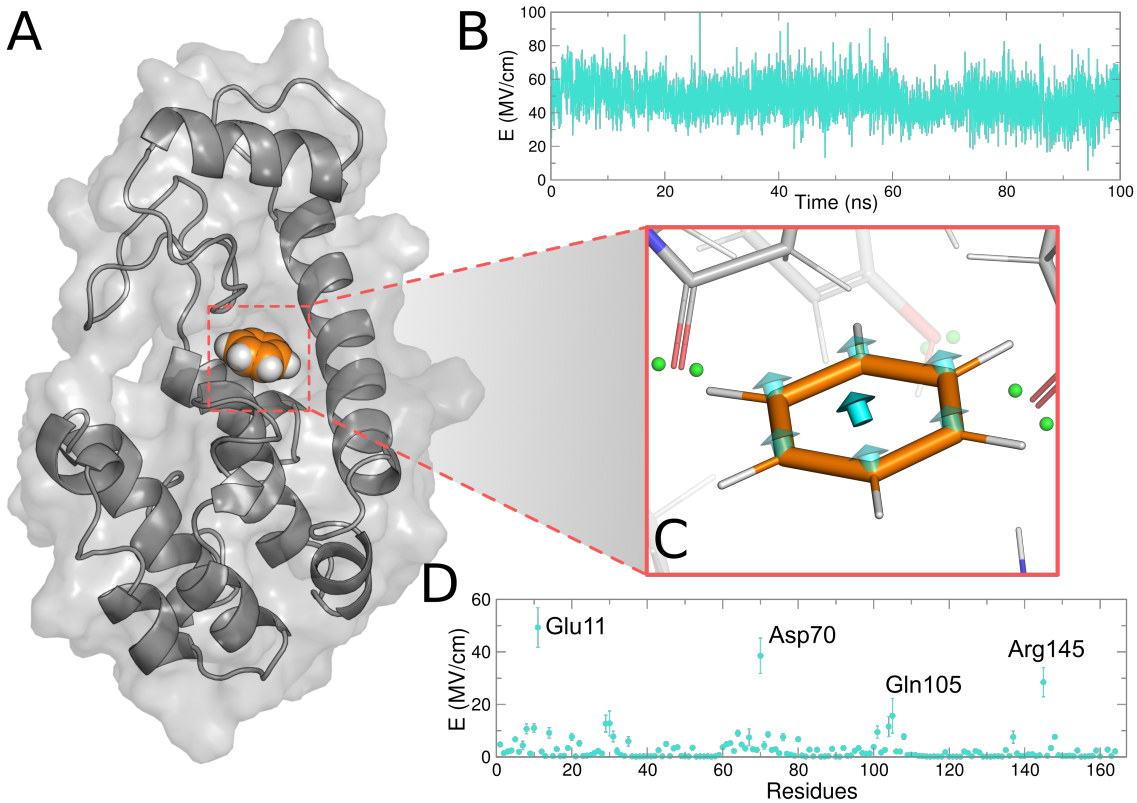


Figure 2: Electric field properties exerted by T4 lysozyme. A) Benzene ligand (carbon in orange) bound to T4 lysozyme. B) Time series of $\vec{E}$ sensed by benzene at its COM. C) Direction of $\vec{E}$ at the benzene COM (solid cyan vector) and at each benzene carbon atom (translucent cyan vectors). Oxygen lone pairs are shown in green spheres. D) Contribution of each T4 lysozyme residue to the total $\vec{E}$.

electric field on the C=O bond of 19NT, stabilizing the intermediate state[12,13,15]. Such calculations were supported by Stark spectroscopy results[12] and quantum mechanics calculations[15,35], suggesting the potential of understanding the electric field organization at enzyme active sites.

In this work, we reproduced the results obtained by Fried et al.[12] using TUPÃ. By using the BOND MODE and selecting the carbonyl atoms as probes (see Section 3.2 in the Supplementary Material), we were able to calculate the electric field exerted by the enzyme onto the C=O bond axis (Figure 3). For this calculation, we defined the bond axis direction as the same as the bond dipole (C $\rightarrow$ O). We observed an electric field magnitude of 136 $\pm$ 16 MV/cm over 10 ns (Figure 3A), consistent with the large field observed by Fried et al.[12]. Also supported by the findings of Wu et al.[15], the total $\vec{E}$ was mostly aligned with the bond axis (95% on average), although pointing in the opposite direction (Figure 4B).

As extensively described by Fried et al.[12], Wu and Boxer[13] and Wu et al.[15], the residues that contribute the most to $\vec{E}$ and its projection onto the bond axis are the ones involved in catalysis, namely Tyr16, Asp40, and Asp103. Tyr16 and Asp103 interact with the carbonyl group (Figure 4A) and exert the greater contribution to $\vec{E}$, while Asp40 also plays an important role in generating the total field (Figure 3B). Although the direction of $\vec{E}$ has been mathematically described before, here we used pyTUPÃ to explicitly draw $\vec{E}$, $\vec{E}_{proj}$ and their orientation in relation to the bond axis as shown in Figure 4B.

Overall, our results obtained by using TUPÃ agree with previously reported data[12], reflecting the robustness and flexibility of our approach.

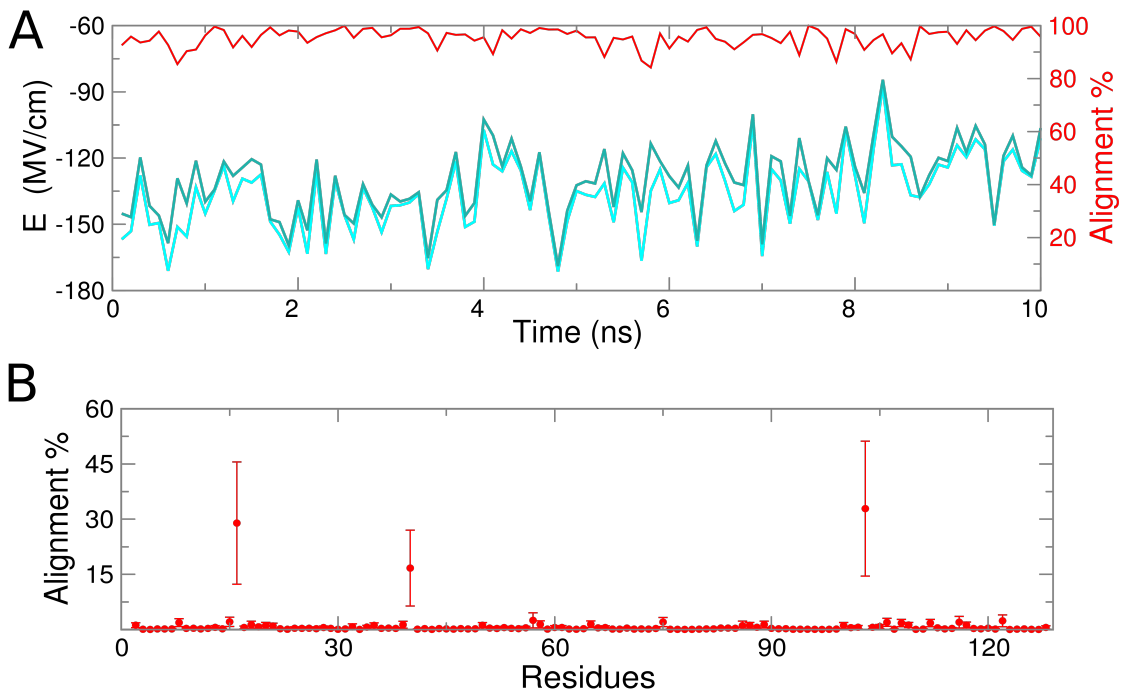


Figure 3: Electric field properties of KSI electric field exerted at 19NT carbonyl bond. A) Time series of $\vec{E}$ (cyan), $\vec{E}_{proj}$ (dark turquoise) and their alignment percentage (red). B) Alignment percentage of electric fields exerted by each KSI residue and the 19NT carbonyl bond axis. Tyr16, Asp40, and Asp103 participate in KSI catalysis.

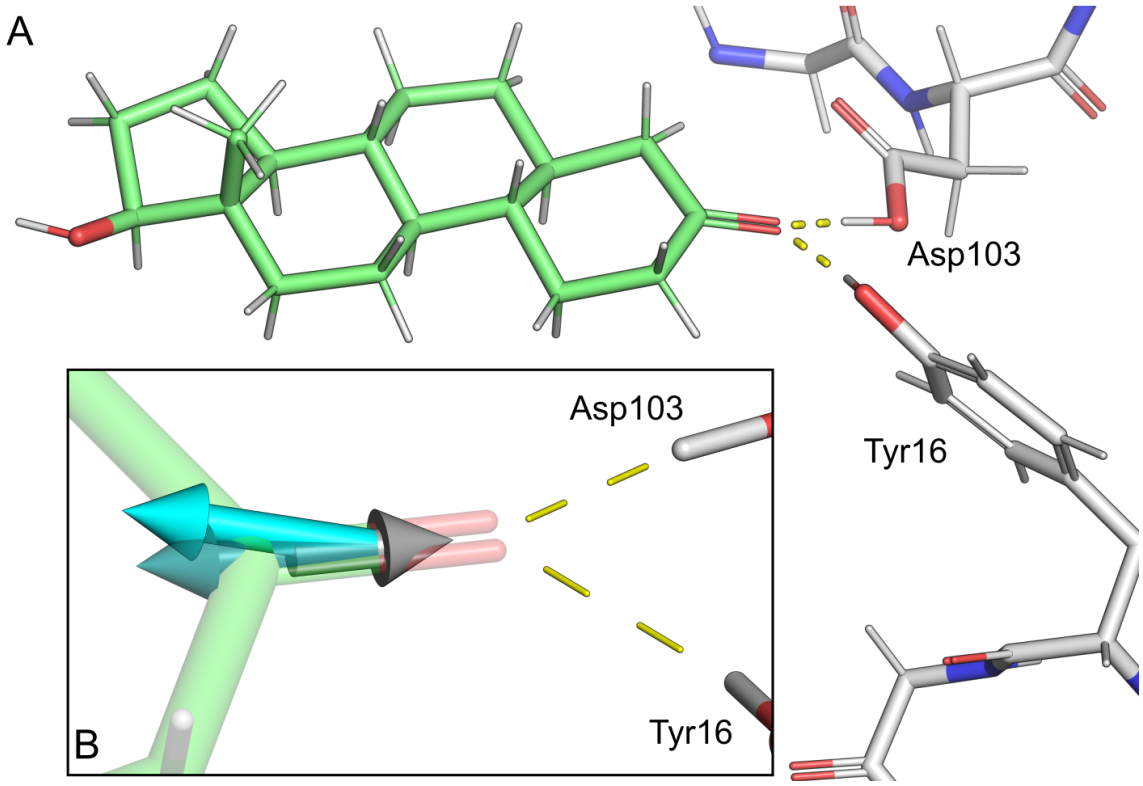


Figure 4: Binding mode of 19NT in KSI active site. A) Tyr16 and Asp103 interact with the carbonyl oxygen of 19NT, positioning the substrate-like inhibitor within the active site. B) Electric field exerted by KSI on the carbonyl bond. The C=O bond axis is shown in translucent gray, $\vec{E}$ in solid cyan, and $\vec{E}_{proj}$ in translucent dark turquoise.

NaK2K ion channel

Ion channels play crucial roles in biological processes and have been thoroughly studied via MD simulations to better comprehend the molecular forces driving ion transport, flux, and selectivity[36–41]. Here, we studied the double mutant (D66Y and N68D) NaK2K reported by Derebe et al.[28] and derived from the NaK channel family to demonstrate the electric field organization throughout the channel core. Since we wanted to conserve the main crystallographic features of NaK2K structure, our MD protocol included a position restraint force of 50 kJ/(mol nm^2) on backbone atoms.

We calculated the electric field exerted by the protein and membrane atoms ($\vec{E}_{PM}$) throughout the ion channel core by running TUPÃ in COORDINATE MODE multiple times (see Section 3.3 in the Supplementary Material). Note that this calculation does not consider the field exerted by the coordinated K^+ ions (K1 to K4) on each other and we are solely focusing on the electric field generated by the ion channel and the membrane. We took advantage of the fact that our trajectories were centered in the X-Y plane and the channel core was aligned to the Z-axis to scan $\vec{E}_{PM}$ at coordinates (0,0,26) through (0,0,72). As shown in Figure 5A and B, $\vec{E}_{PM}$ is oriented toward the channel core, which would facilitate the transport of K^+ ions both inward and outward. Figure 5A shows how the magnitude of the $\vec{E}_{PM}$ Z-axis component ($\vec{E}_{PMz}$) varies throughout the channel core. Starting from Z = 72 Å, $\vec{E}_{PMz}$ increases and peaks close to the K1 binding site (Z = 51 Å), with a magnitude of 169 $\pm$ 7 MV/cm. The $\vec{E}_{PMz}$ decreases through K2 and K3 binding sites (Z = 44 Å and Z = 40 Å, respectively), reaching values of $\approx$ 30 MV/cm and 0 MV/cm, respectively. $\vec{E}_{PMz}$ peaks again at K4 binding site (192 $\pm$ 18 MV/cm) with an opposite direction and the decreases towards Z = 26 Å.

Although understanding how the ion channel organizes the electric field in its core,

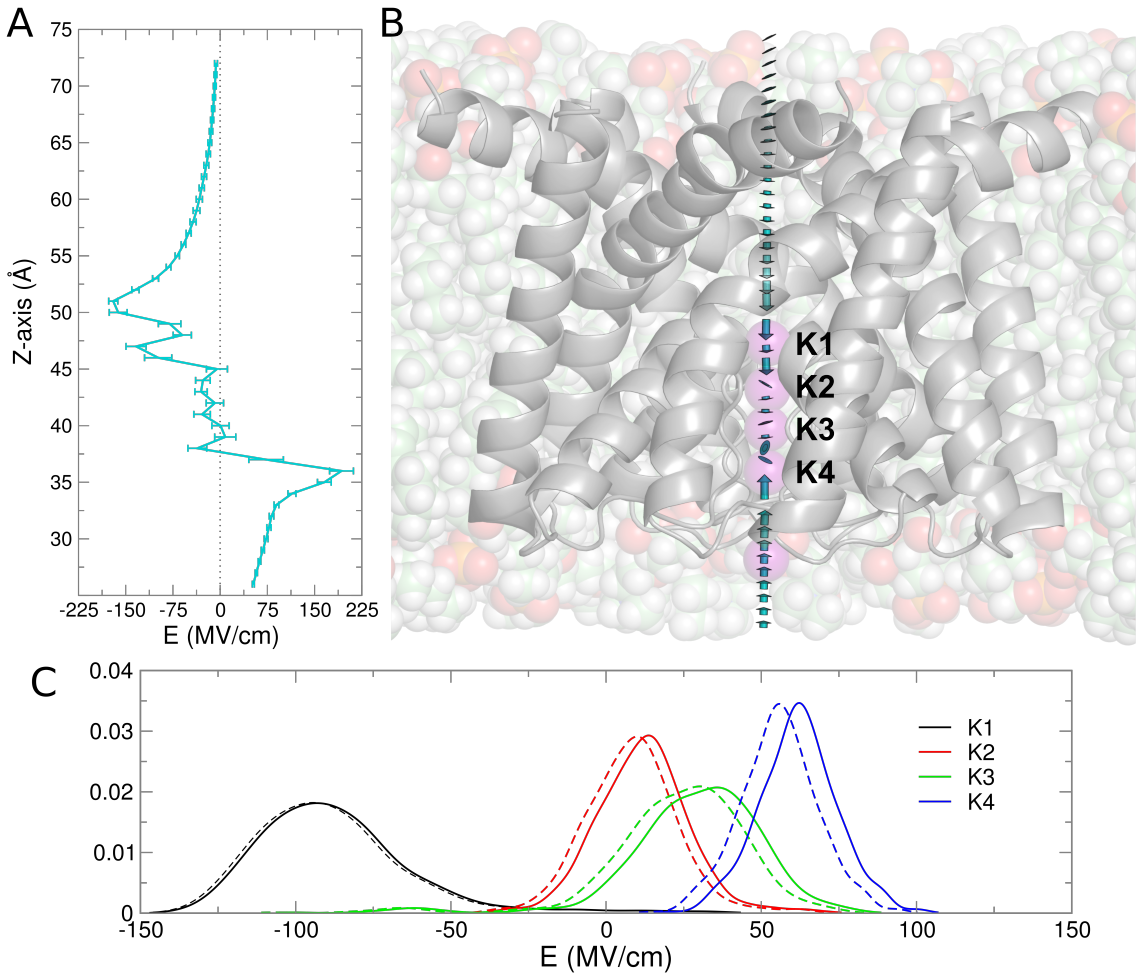


Figure 5: Electric field properties of NaK2K ion channel. A) $\vec{E}_{PMz}$ throughout the ion channel as depicted in panel B. Negative values represent a vector pointing downward, while positive values represent a vector pointing upward. B) $\vec{E}_{PMz}$ orientation throughout the ion channel core. Coordinated K^+ ions K1 to K4 are shown as translucent magenta spheres. C) $\vec{E}_{PMKz}$ distribution exerted on coordinated K^+. Full lines represent $\vec{E}_{PMKz}$ when accounting for the contribution of protein, membrane, and other coordinated ions. Dashed lines represent $\vec{E}_{PKz}$, the electric field without the membrane contribution.

the electric force sensed by K1-K4 ions during dynamics also considers contributions of the other coordinated K^+ ions. Hence, we used the ATOM MODE to define each of the K^+ as probes while making sure to include the other ions in the environment set when calculating the electric field ($\vec{E}_{PMKz}$). Configuration files for that calculation can also be found in the Supplementary Material. As shown in Figure 5C, $\vec{E}_{PMKz}$ exerted on K1 was consistently negative and it achieved larger magnitudes than those of the other coordinates ions. $\vec{E}_{PMKz}$ exerted on K2 and K3 had smaller magnitudes and adopted both positive and negative values, with K3 being more consistently positive than K2. Lastly, $\vec{E}_{PMKz}$ exerted on K4 was consistently positive with a somewhat large magnitude.

Moreover, we also evaluated the membrane contribution to $\vec{E}_{PMKz}$. For this, we calculated the electric field exerted by the membrane atoms ($\vec{E}_{Mz}$) on the coordinated ions and obtained the contribution of NaK2K and coordinates ions by evaluating $\vec{E}_{PKz} = \vec{E}_{PMKz} - \vec{E}_{Mz}$. As seen in Figure 5C, the membrane contribution was larger at K4 than the rest of the coordinated atoms, but $\vec{E}_{PKz}$ was still the major contribution to the total field sensed by those ions throughout the simulation. Even though our MD protocol restrained the conformational dynamics of the ion channel, our results highlight how understanding the directionality of such fields and forces can contribute to the study of ion channels and the relationship between their conformational dynamics and ion transport.

Long-range contributions to electric field

In TUPÃ, $\vec{E}$ is calculated assuming a real-space electrostatic contribution of all atoms in the system. While an earlier approach[7] used a similar assumption, it did not account for contributions arising from particles in images around the central

periodic simulation box. TUPÃ, however, can handle such contributions by applying MDAnalysis selections, which automatically evaluate for particles in periodic images within the defined radial cutoff.

Coulombic forces are typically calculated in MD simulations using particle-mesh Ewald (PME)[42,43] or reaction-field (RF)[44] methods for particles beyond a certain cutoff (usually between 10-12 Å) instead of real-space calculations such as in TUPÃ. In fact, Fried et al.[5] noted this difference of approach and proposed to derive $\vec{E}$ by using the actual electric force on each atom as calculated and output by the MD engine. In doing so, $\vec{E}$ is calculated on each atom i by $\vec{E} = \vec{F}_{electric}/Q_i$. Such an approach limits the calculation of $\vec{E}$ to atoms and their geometric centers, resembling the ATOM and BOND calculation modes in TUPÃ. However, due to the r^{-2} dependency of $\vec{E}$, the contribution of particles beyond the commonly used electrostatic cutoff should be minor when compared to the expected fluctuation of $\vec{E}$ due to the conformational dynamics of nearby particles.

To determine the effect of long-range contributions to the total electric field, we analyzed the T4 lysozyme system and selected all atoms within increasing radii of the benzene molecule to compose the environment set. The T4 lysozyme is an interesting system for this analysis since the benzene ligand is accessible to the solvent throughout the entire simulation. As shown in Figure 6A, $|\vec{E}|$ is mostly converged beyond 15 Å, demonstrating that the contribution of distant particles is negligible regardless of the environment set composition change. In fact, that is precisely the rationale for applying PME or RF methods in MD simulations. In contrast, the composition of the environment set in close proximity to the probe (less than 10 Å) and particles dynamics have a greater impact in $|\vec{E}|$ (Figure 6B). Overall, care must be taken when defining the environment set to guarantee an accurate reproduction of $\vec{E}$.

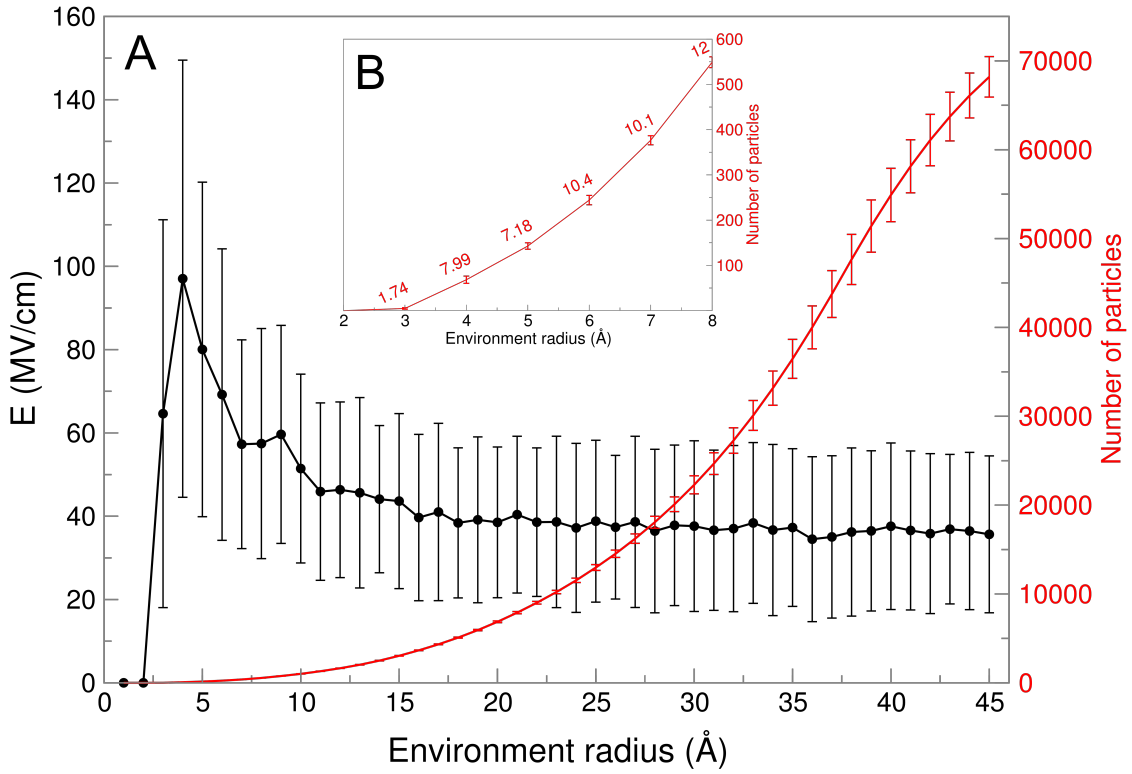


Figure 6: Electric fields of the T4 lysozyme system. A) Electric field magnitudes (in black) and the respective environment size (in red) as a function of environment radial composition. Standard deviations shown as bars. B) Variation of the environment set with the standard deviation explicitly shown.

Generally, we do not expect that our approach would significantly differ from the actual computed electric force and field in MD trajectories. By taking advantage of the real-space electrostatic contribution, TUPÃ can evaluate how $\vec{E}$ is oriented at any given point within the simulation box using the COORDINATE calculation mode.

CONCLUSIONS

Due to the increasing interest in studying electric field effects in biological systems, the features implemented in TUPÃ were designed to provide compatibility with a wide range of trajectory and topology formats, while also granting analytical flexibility to analyze multiple types of biological systems via the three implemented calculation modes. By taking advantage of the all-atom MM convention, TUPÃ is compatible with both additive and polarizable force fields that rely on the same convention, such as the well-known AMBER, CHARMM, OPLS-AA, and Drude force fields. Here, we have demonstrated the robustness and utility of this software across multiple systems of biological interest with both additive and polarizable force fields. We expect TUPÃ to assist researchers worldwide when analyzing electrostatic-driven events in molecular simulations.

ACKNOWLEDGMENTS

The authors thank Haley M. Michel for assistance in preparing TUPÃ's logo. This work was supported by the National Institutes of Health (grant R35GM133754) and USDA-NIFA (project VA-160092).

References

1. L. Delemotte, M. L. Klein, and M. Tarek, Frontiers in Pharmacology p. 97 (2012).

2. W. Kopec, B. S. Rothberg, and B. L. de Groot, Nature Communications 2019 10:1 **10**, 1 (2019).

3. F. Che, J. T. Gray, S. Ha, N. Kruse, S. L. Scott, and J. S. McEwen, ACS Catalysis **8**, 5153 (2018).

4. N. G. Léonard, R. Dhaoui, T. Chantarojsiri, and J. Y. Yang, ACS Catalysis **11**, 10923 (2021).

5. S. D. Fried, S. Bagchi, and S. G. Boxer, Journal of the American Chemical Society **135**, 11181 (2013).

6. D. L. Ensign and L. J. Webb, Proteins: Structure, Function and Bioinformatics **79**, 3511 (2011).

7. A. T. Fafarman and S. G. Boxer, Journal of Physical Chemistry B **114**, 13536 (2010).

8. V. Vaissier, S. C. Sharma, K. Schaettle, T. Zhang, and T. Head-Gordon, ACS Catalysis **8**, 219 (2018).

9. S. D. Fried and S. G. Boxer, Accounts of Chemical Research **48**, 998 (2015).

10. V. V. Welborn, L. Ruiz Pestana, and T. Head-Gordon, **1**, 649 (2018).

11. R. T. Bradshaw, J. Dziedzic, C. K. Skylaris, and J. W. Essex, Journal of Chemical Information and Modeling **60**, 3131 (2020).

12. S. D. Fried, S. Bagchi, and S. G. Boxer, Science **346**, 1510 (2014).

13. Y. Wu and S. G. Boxer, Journal of the American Chemical Society **138**, 11890 (2016).

14. V. V. Welborn and T. Head-Gordon, Journal of the American Chemical Society **141**, 12487 (2019).

15. Y. Wu, S. D. Fried, and S. G. Boxer, Journal of the American Chemical Society **142**, 9993 (2020).

16. J. Nash, T. Barnes, and V. Welborn, Journal of Open Source Software **5**, 2576 (2020).

17. J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J. P. Piquemal, P. Ren, and J. W. Ponder, Journal of Chemical Theory and Computation **14**, 5273 (2018).

18. J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. Distasio, et al., Journal of Physical Chemistry B **114**, 2549 (2010).

19. Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, Journal of Chemical Theory and Computation **9**, 4046 (2013).

20. C. Zhang, C. Lu, Z. Jing, C. Wu, J. P. Piquemal, J. W. Ponder, and P. Ren, Journal of Chemical Theory and Computation **14**, 2084 (2018).

21. G. Van Rossum and F. L. Drake, *Python 3 Reference Manual* (CreateSpace, Scotts Valley, CA, 2009), ISBN 1441412697.

22. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, Journal of Computational Chemistry **32**, 2319 (2011).

23. R. J. Gowers, M. Linke, J. Barnoud, T. J. E. Reddy, M. N. Melo, S. L. Seyler, J. Domański, D. L. Dotson, S. Buchoux, I. M. Kenney, et al., Proceedings of the 15th Python in Science Conference pp. 98–105 (2016).

24. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, et al., Nature **585**, 357 (2020), URL `https://doi.org/10.1038/s41586-020-2649-2`.

25. E. Baldwin, W. A. Baase, X. J. Zhang, V. Feher, and B. W. Matthews, Journal of Molecular Biology **277**, 467 (1998).

26. F. Y. Lin, J. Huang, P. Pandey, C. Rupakheti, J. Li, B. Roux, and A. D. Mackerell, Journal of Chemical Theory and Computation **16**, 3221 (2020).

27. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller, and A. D. MacKerell, Nature Methods **14**, 71 (2016).

28. M. G. Derebe, D. B. Sauer, W. Zeng, A. Alam, N. Shia, and Y. Jiang, Proceedings of the National Academy of Sciences of the United States of America **108**, 598 (2011).

29. B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, et al., Journal of Computational Chemistry **30**, 1545 (2009).

30. D. P. Sun, D. I. Liao, and S. J. Remington, Proceedings of the National Academy of Sciences of the United States of America **86**, 5361 (1989).

31. G. Lamoureux, E. Harder, I. V. Vorobyov, B. Roux, and A. D. MacKerell, Chemical Physics Letters **418**, 245 (2006).

32. A. Warshel, Journal of Biological Chemistry **273**, 27035 (1998).

33. A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. Olsson, Chemical Reviews **106**, 3210 (2006).

34. S. D. Fried and S. G. Boxer, Annual Review of Biochemistry **86**, 387 (2017).

35. L. Wang, S. D. Fried, and T. E. Markland, Journal of Physical Chemistry B **121**, 9807 (2017).

36. R. N. Roy, K. Hendriks, W. Kopec, S. Abdolvand, K. L. Weiss, B. L. De Groot, A. Lange, H. Sun, and L. Coates, IUCrJ **8**, 421 (2021).

37. A. Mironenko, U. Zachariae, B. L. de Groot, and W. Kopec, **433**, 167002 (2021).

38. C. Öster, K. Hendriks, W. Kopec, V. Chevelkov, C. Shi, D. Michl, S. Lange, H. Sun, B. L. de Groot, and A. Lange, Science Advances **5** (2019).

39. V. Carnevale, L. Delemotte, and R. J. Howard, **46**, 621 (2021).

40. S. Pérez-Conesa, E. G. Keeler, D. Zhang, L. Delemotte, and A. E. McDermott, Journal of Chemical Physics **154**, 165102 (2021).

41. A. Sridhar, S. C. Lummis, D. Pasini, A. Mehregan, M. Brams, K. Kambara, D. Bertrand, E. Lindahl, R. J. Howard, and C. Ulens, Journal of Biological Chemistry **297**, 100899 (2021).

42. T. Darden, D. York, and L. Pedersen, The Journal of Chemical Physics **98**, 10089 (1993).

43. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, Journal of Chemical Physics **103**, 8577 (1995).

44. I. G. Tironi, R. Sperb, P. E. Smith, and W. F. van Gunsteren, Journal of Chemical Physics **102**, 5451 (1995).