Commun. Comput. Phys.
doi: 10.4208/cicp.170711.111111s

Vol. **13**, No. 1, pp. 256-268
January 2013

# DelEnsembleElec: Computing Ensemble-Averaged Electrostatics Using DelPhi

Lane W. Votapka[1], Luke Czapla[1], Maxim Zhenirovskyy[2] and Rommie E. Amaro[1,*]

[1] *Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, USA.*
[2] *Computational Biophysics and Bioinformatics, Department of Physics, Clemson University, Clemson, SC 29634, USA.*

Received 17 July 2011; Accepted (in revised version) 11 November 2011

Available online 12 June 2012

**Abstract.** A new VMD plugin that interfaces with DelPhi to provide ensemble-averaged electrostatic calculations using the Poisson-Boltzmann equation is presented. The general theory and context of this approach are discussed, and examples of the plugin interface and calculations are presented. This new tool is applied to systems of current biological interest, obtaining the ensemble-averaged electrostatic properties of the two major influenza virus glycoproteins, hemagglutinin and neuraminidase, from explicitly solvated all-atom molecular dynamics trajectories. The differences between the ensemble-averaged electrostatics and those obtained from a single structure are examined in detail for these examples, revealing how the plugin can be a powerful tool in facilitating the modeling of electrostatic interactions in biological systems.

**PACS**: 02.70.Ns

**Key words**: DelEnsembleElec, ensemble-averaged electrostatics, Poisson Boltzmann equation, DelPhi, molecular dynamics, influenza glycoproteins.

## 1 Introduction

Electrostatic interactions play an essential role in the dynamics of biological systems. These forces are the predominant long-range interactions influencing the dynamics within and between biomolecules, thus the accurate treatment of electrostatics is necessary for detailed models of these systems. The electrostatic interactions associated with ionic and polar chemical groups found in proteins, nucleic acids, lipids, and other biomolecular systems are essential to both their structure and function [1].

*Corresponding author. *Email address:* ramaro@ucsd.edu (R. E. Amaro)

 

A variety of methods exist for computing the electrostatic interactions within the context of classical Molecular Mechanics (MM) models. The electrostatic energy and potential may be obtained by computing the pairwise Coulomb interaction between all atoms, and many approximate methods exist to efficiently estimate these interactions. While the computation of pairwise interactions is computationally expensive, scaling as the square of the number of atoms in a system, methods such as the Particle-Mesh Ewald (PME) [2] algorithm for periodic systems and the Fast Multipole Method (FMM) [3] can significantly reduce the complexity of computation by introducing well-controlled approximations.

The Poisson-Boltzmann (PB) equation is an attractive method for computing the mean-field potential of biomolecules [4]. By modeling the solvent environment as a continuum dielectric and salt ion distribution, the electrostatic potential and electrostatic free energy of a biomolecule may be estimated under the given conditions. In most implementations, the solvent provides dielectric screening of the biomolecule potential, represented as a uniform dielectric constant, while the salt is modeled based on the potential values within the continuum solvent by a nonlinear term in the PB equation. A linearized PB treatment of salt effects is accurate for proteins with modest net charges in monovalent salt solutions, while the more sophisticated treatment of salt with the nonlinear form of the PB equation is often used in treating multivalent salt solutions and in modeling highly charged polyelectrolytes such as nucleic acids [5]. Many methods of solving the Poisson-Boltzmann equation such as presented here in DelPhi utilize the finite difference method to efficiently solve the electrostatic potential on a discretized grid [6].

Realistic treatment of biomolecular dynamics requires ensemble sampling to understand the conformational flexibility of these molecules within their chemical environment. All-atom explicit solvent Molecular Dynamics (MD) is a widely accepted method for generating a canonical distribution of states for biomolecules modeled using Molecular Mechanics force fields such as AMBER [7] and CHARMM [8]. By averaging over the trajectory of states produced in these simulations, ensemble properties such as the electrostatic potential may be obtained, giving insight into thermodynamic properties of these systems. This approach has been useful in previous studies of biomolecular systems, such as in understanding the electrostatic environment of protein-bound drugs [9] and in understanding the transport of ions and biomolecules through membrane pore proteins [10]. Notably, the use of ensemble information in particular, as opposed to a single static structure, when computing the electrostatic properties of dynamic biomolecules has been shown to increase agreement of the computed values with experiment [9].

To facilitate the calculation of ensemble-averaged electrostatic properties with the Poisson-Boltzmann equation, we have developed a plugin interface for the visualization software package VMD [11], which interfaces with the DelPhi numerical PB equation-solving software package [6, 12]. This plugin computes the ensemble electrostatic potential and free energy of biomolecules from their trajectories in VMD-compatible formats such as those used in the MD packages AMBER and CHARMM. A graphical user frontend provides a simplified interface for specifying all the individual options supported

by DelPhi, and the plugin creates the proper biomolecule coordinate inputs for DelPhi from the system partial charges and radii obtained by VMD from the input structure file, as well as automating the process of ensemble averaging. The plugin executes DelPhi and then visualizes the electrostatic potential or ensemble-averaged potential from the DelPhi grid outputs in a variety of VMD formats, such as color-coded surface, 2D plane projection, and isopotential surface plots. The plugin supports both the linear and non-linear forms of the PB equation provided in DelPhi.

The application of these Poisson-Boltzmann calculations with the DelPhi Electrostatics plugin to biomolecular systems of current research interest such as the influenza virus major glycoproteins, hemagglutinin [13, 14] and neuraminidase [15–17], are presented in this work as examples of the versatility of calculations that can be performed with this convenient new tool. The ensemble-averaged mean-field potential of these proteins are obtained in identical salt and solvent conditions as the corresponding MD simulation trajectories. The electrostatic potential of any AMBER or CHARMM system may be computed from its structure file and coordinates with this plugin, importing the parameters into the DelPhi calculation in Protein Databank (PDB) format. As the software plugin and source code is freely available under the GNU Public License (GPL), it is anticipated that this tool will be of general interest to the physics-based modeling research community.

## 2 Methods

### Molecular dynamics simulations

Trajectories for two biomolecules were generated using explicitly solvated, all-atom molecular dynamics (MD) simulations. We chose two proteins important in influenza pathogenesis, neuraminidase and hemagglutinin, as subject molecules on which to perform the electrostatic calculations. Three different commonly used MD-related software packages were used for the simulation of each molecule, as described below.

### 2.1 Influenza hemagglutinin

Atomic coordinates were taken from accession number 1HGF in the Protein Data Bank (PDB). The protonation states for histidine and other titratable residues were determined at pH 7.0 using the PDB2PQR [18] server using PROPKA [19]. Crystallographic water molecules were retained. The AMBER Tools11 [20] program sLeap was used to connect disulfide bridges and also to add a TIP3P water box with a 10 Å box spacing along each edge beyond the dimensions of the protein and 20mM NaCl to act as an explicit solvent along with K+ counterions to neutralize the system. The composite system contained 198,533 atoms. The hemagglutinin trimer was then minimized for 45,000 steps and equilibrated using four stages of harmonic constraints at 250,000 steps each, starting at 4 kcal $mol^{-1}$ $Å^{-2}$ reducing by 1 kcal $mol^{-1}$ $Å^{-2}$ each time. The systems were then simulated using NAMD2.8 [21] with the AMBER FF99SB [22] force field under periodic boundary

conditions with the isothermal-isobaric (NPT) ensemble at a temperature of 310 K. Pressure was maintained at 1 atm using a Nosé-Hoover Langevin Piston [23] and Particle-Mesh Ewald [24] was used to treat long-range electrostatic interactions. Bonds involving hydrogen positions were constrained using the RATTLE algorithm [25]. Ranger, a massively parallel Teragrid computing platform, was utilized to perform the calculations (benchmark: 3.9 ns/day using 512 cores). The final trajectory contained 100 nanoseconds of simulation, which was reduced to 500 frames with a stride of 200 picosecond between each frame for analysis of ensemble electrostatics.

### 2.2 Influenza neuraminidase

Details for the preparation of this system has been described previously [26]. Atomic coordinates were taken from PDB accession code 3NSS. The protonation states for histidines and other titratable groups at pH 6.5 were determined using the PDB2PQR [18] server. The system was prepared using the Ambertools11 program xLeap with a padding of 10-12 Å of water molecules prepared in a similar fashion to the hemagglutinin system mentioned above. The neuraminidase tetramer was then energy minimized using the PMEMD module from AMBER11 with 2000 steps of steepest descent, followed by 5000 steps of conjugate gradient minimization with 5.0 kcal $\text{mol}^{-1}$ $\text{Å}^{-2}$ harmonic restraints. Then 25,000 more conjugate gradient steps were performed without restraints. The neuraminidase system was then also equilibrated by gradual heating to 310 K in the isothermal/constant volume (NVT) ensemble using a Langevin thermostat with a collision frequency of 5.0 $\text{ps}^{-1}$. Three subsequent 250 ps runs were performed at 310 K in the isothermal/isobaric (NPT) ensemble, with 4 kcal $\text{mol}^{-1}$ $\text{Å}^{-2}$ restraints being reduced by 1 kcal $\text{mol}^{-1}$ $\text{Å}^{-2}$ in each consecutive run. A Berendsen barostat [27] with a coupling constant of 1 ps and a target pressure of 1 atm was used to maintain pressure, followed by a final 250 ps segment of NPT dynamics without restraints. Production dynamics was then performed for 100 ns with conditions similar to that of the hemagglutinin system above. Minimization, equilibration, and production were all performed on the NCIS Cray XT4 and SDSC Trestles high performance compute platforms (benchmark: 10.1 ns/day using 256 cores [1 core per node] on NICS Athena). We pruned the final trajectory to 500 frames with a stride of 200 ps between each frame for analysis of ensemble electrostatics.

### 2.3 Poisson-Boltzmann ensemble-averaged electrostatics

The PB equation (2.1) uses an implicit and continuum-based model of the solvent and counterion environment surrounding a biomolecule to give a detailed description of its electrostatics [28]. Though many variations exist, it assumes a spatially varying dielectric constant and takes into account the shape and irregular charge distribution of the biomolecule [29]. A general form is given below [12]:

$$\nabla\cdot[\epsilon(x)\nabla\phi(x)]+\frac{e}{\epsilon_0}\sum_j c_j z_j \exp\left(-\frac{ez_j\phi(x)}{kT}\right)=-\frac{e}{\epsilon_0}\sum_i q_i\delta(x-x_i). \tag{2.1}$$

Here, $\phi(x)$ represents the electrostatic potential at position $x$, $\epsilon(x)$ is the spatially-varying value of the dielectric constant (in units relative to $\epsilon_0$), $q_i$ are the individual $N_i$ partial charges associated with the atoms of the biomolecule (with positions specified by $x_i$), and $e$ is the elementary electronic charge. The value $kT$ is the Boltzmann factor (Boltzmann constant times temperature) and $\delta(x)$ is the Dirac delta function. The second term in (2.1) is associated with the $N_j$-component ion density distribution with components having concentration $c_j$ at a distance infinity away from the biomolecule and having valence $z_j$, at the specified salt conditions. Introducing the approximation

$$\frac{e}{\epsilon_0}\sum_j c_j z_j \exp\left(-\frac{ez_j\phi(x)}{kT}\right) = -\kappa^2\phi(x) \tag{2.2}$$

(where $\kappa^2 = \frac{e^2}{kT\epsilon_0}\sum_j c_j z_j^2$) into Eq. (2.1) yields the linear PB equation, which is accurate for conditions such as modestly charged molecules (having relatively low electrostatic potentials in the surrounding medium) in a 1:1 monovalent salt environment.

$$\nabla\cdot[\epsilon(x)\nabla\phi(x)] - \kappa^2\phi(x) = -\frac{e}{\epsilon_0}\sum_i q_i\delta(x-x_i). \tag{2.3}$$

Typically a computational method, such as implemented in the program DelPhi, is used to solve the equation numerically in a system of biological scale, as its canonical form is a nonlinear partial differential equation. In systems without regions of large potential values, the equation can be simplified to a linear form (2.3). The variety of approaches to approximating the PB equation allow differing degrees of accuracy and efficiency [28], usually one at the expense of the other.

## 3 User interface

DelEnsembleElec provides a graphical interface that can be accessed within the "Extensions" menu of VMD, simplifying the process of specifying DelPhi options and then computing single-point and ensemble electrostatic calculations, using the many VMD display drawing methods to visualize the resulting potential grids. The code is written in the Tcl script language and is compatible with Microsoft Windows and UNIX-based platforms (Mac, Linux). When DelEnsembleElec is started, the main window appears (Fig. 1), which contains the most basic customizable options. Once one or more trajectories or structures are loaded into VMD, or if they have been loaded already, a menu button allows the user to select the molecule on which to perform the run. Entry fields allow the user to specify the atom selection of the molecule, as well as an option to write a Gaussian cube file upon completion of the run. Two additional checkboxes allow the user to specify whether and how the completed map data will be loaded into VMD for visualization.

In addition, two dropdown menu options provide additional functionality. Under the menu Edit > Settings, a window appears where one can specify the working directory

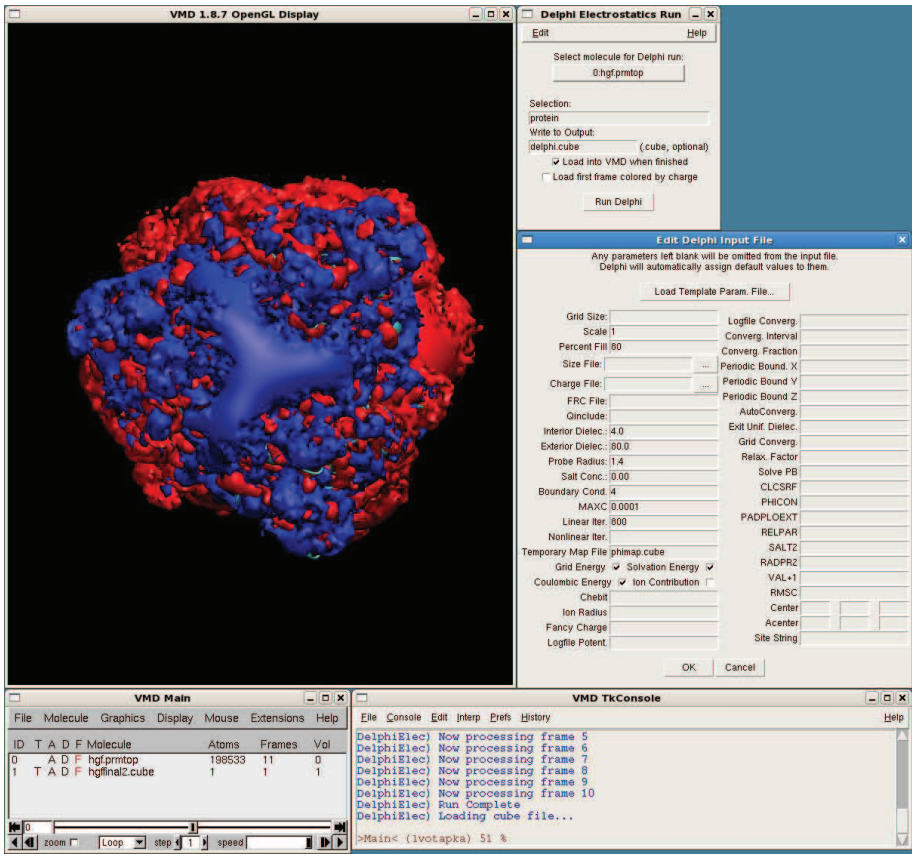


Figure 1: DelEnsembleElec plugin interface. A screenshot of the DelEnsembleElec plugin is shown; the top right window is the main DelEnsembleElec window. The window immediately below it is for setting the parameters in the DelPhi input file. The remaining windows are VMD windows; main display window at the top left, molecule loader and animation control at the bottom left, and console at the bottom right.

of the Delphi runs(s) (set automatically to the operating system's temporary directory) as well as the location of the Delphi program. The user may optionally create the setup file only, but not actually initiate a Delphi run. Under the menu Edit > Input File, a window enables the user to customize most of the Delphi input parameters, either by entering the fields manually or by loading a template Delphi parameter file.

Once satisfied with the run specifications, the user may click the "Run Delphi" button at the bottom of the Main Window to execute the calculation. The time taken for execution depends on several factors, including the number of atoms in the system, the number of frames in the trajectory, as well as the amount of processor speed dedicated to the calculation. A benchmark on a single Intel Xeon 2.67 GHz processor takes approximately 15 minutes per frame to solve the linear PB equation for a system containing 20,000 atoms on a $169\times169\times169$ static grid. Upon completion of the run, depending on the options selected by the user, the map data may be saved as a Gaussian cube file or loaded as a new

molecule into the VMD main window, where the VMD representation window allows one to manipulate the graphical representation in a variety of formats. The plugin also has the capability to compute and display "difference maps" between the ensemble and single frame electrostatic grids (ensemble-averaged grid potential minus a single-frame grid potential).

When a single structure or a trajectory is loaded that does not contain atomic charge or radius information in a way that VMD can recognize (such as a pdb trajectory), this information must be provided to Delphi as force field parameter files. The paths to these files may be specified in the Edit > Input File window under the "Size File" and "Charge File" fields. If the user neglects to specify a path to a size file, these parameters are assigned automatically using typical CHARMM radii (e.g. carbon: 2 Å, hydrogen: 0.7 Å, oxygen: 1.7 Å, nitrogen: 1.8 Å, sulfur: 2.0 Å, phosphorous: 2.15 Å). If the user desires different atomic radii, a radii size file must be specified in the input parameters.

# 4 Results

## 4.1 Influenza hemagglutinin electrostatics

Explicitly solvated all-atom MD simulations were performed on the hemagglutinin trimer protein for 100 ns using NAMD2.8. Afterwards, the protein snapshots were extracted at 200 ps intervals and aligned. Using DelEnsembleElec, a MD trajectory containing 500 frames of approximately 20,000 atoms required nearly 50 hours to complete using a single Intel Xeon: 2.67 GHz processor, as Delphi must calculate the PB equation for each individual frame of the trajectory. Once complete the Gaussian cube is loaded into VMD, it is represented by the VMD graphical representations isosurface and solvent-accessible surface area colored by loaded data (Figs. 2-4). The grid scale DelPhi parameter was set to 1.0 Å. The exterior dielectric was set to 80.0, with the interior set to 2.0. The salt concentration in the continuum solvent was set to 20 mM; consistent with the concentration in the MD simulations. The probe radius defining the dielectric boundary was 1.4 Å. Since the solvent was a 1:1 monovalent salt distribution, the PB equation was solved linearly by reaching a convergence of less than a 0.0001 kT/e change of potential. For comparison, DelEnsembleElec was also used to perform the same calculation on only a single snapshot (frame 0, representing the equilibrated hemagglutinin).

The hemagglutinin glycoprotein contains a long "stalk" region, which anchors the protein to the membrane at one end. The receptor-binding domain, which is the major antigenic area, resides at the opposite end and controls the entry of viral particles through productive binding events between sialic acid receptors on the host cell. In both the single frame and ensemble-averaged electrostatic maps, it is apparent that hemagglutinin exhibits a dipole, with the receptor-binding domain having a more positive potential and the bottom of the stalk having a more negative potential (Fig. 2). The ensemble-averaged potentials projected onto the surface of the protein indicate an increased polarity, as indicated by the increased regions of negative (red) charge colored on the surface.

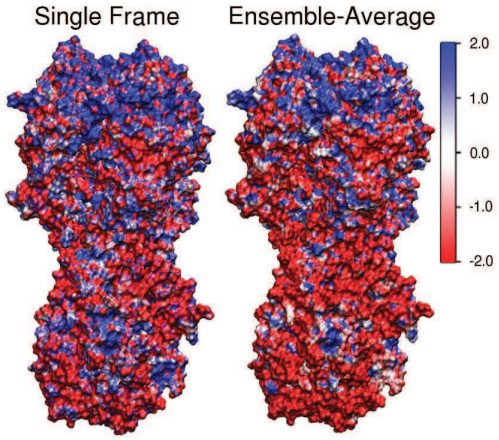


Figure 2: Hemagglutinin stalk electrostatics. Side view of the hemagglutinin trimer depicted by the solvent-accessible surface area colored by electrostatic potential calculated on a single frame (left panel) and on the ensemble-averaged 500-frame trajectory (right panel). Units are in $kT/e$ at $T=300$K; these units are used in all of the presented results.

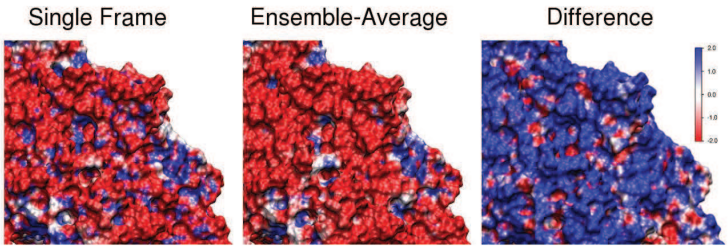


Figure 3: Close-up of hemagglutinin stalk electrostatics. Close-up of the hemagglutinin stalk solvent-accessible surface area colored by electrostatic potential. The left panel depicts electrostatics calculated on a single frame. The center panel depicts the same surface for the ensemble-averaged 500-frame trajectory. The right panel depicts the difference in potential as the single-frame potential subtracted from the ensemble-averaged potential. The smoother distribution of surface potential values from the ensemble-averaged electrostatics more accurately represents the effective potential encountered over the trajectory of sampled protein fluctuations. Units are in $kT/e$ at $T=300$K.

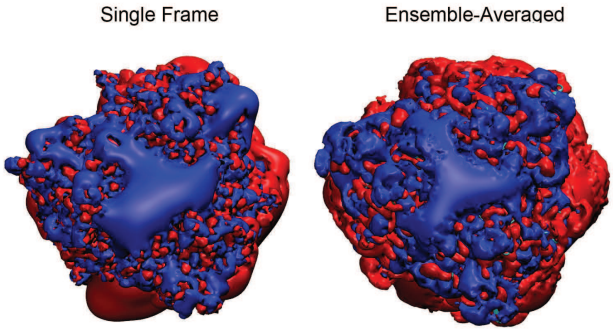


Figure 4: Hemagglutinin receptor binding site electrostatics. Top view of hemagglutinin with potential fields at positive (blue) and negative (red) $2kT/e$. The left panel shows the result when the calculation is performed with only a single frame. The right panel shows the result for the ensemble-averaged 500-frame trajectory.

In general, the ensemble averaging tends to "blur out" small pockets of variable charges that are present due to specific point charges on the protein surface (Fig. 3). Such effects are indicated by larger patches of unobstructed electrical charge on the surface of the protein. When the values present in the single-frame grid are subtracted from the ensemble-averaged grid, the resulting "difference map" highlights the dramatic differences between the two (Fig. 3 right panel). Grid points located near atomic centers have very extreme values in the single frame, and these points are greatly dampened in the ensemble-average. This accounts for the large difference between equivalent grid points in the difference map.

The ensemble-averaged electrostatic potentials also indicate a more symmetric electric field, as shown by isopotential values, at the receptor binding domain end of hemagglutinin (Fig. 4). This dynamic electrostatic information may be of critical value in the rational design of improved vaccines or better understanding the interactions of hemagglutinin with glycan receptors on the surface of the host cell [14, 20].

### 4.2 Influenza neuraminidase electrostatics

One hundred nanoseconds of explicitly solvated, all-atom MD simulations of the tetrameric neuraminidase protein were performed using AMBER11. Protein snapshots were again extracted at 200 ps intervals and aligned based on alpha carbons to remove rotational and translational motion. DelEnsembleElec was used to compute the electrostatics for a single frame (frame 0, representing the equilibrated protein) and the MD trajectory (500 frames). The DelPhi parameters used for neuraminidase were the same as those used for the hemagglutinin system.

The influenza neuraminidase protein controls viral particle exit from the host cell by cleaving the terminal sialic acid linkage on the host cell glycan receptors, and as such, is currently the major target for small molecule antiviral compounds [30]. Although most small molecule drug discovery efforts have focused exclusively on optimizing ligand-protein interactions within the sialic acid binding site, a secondary sialic acid binding site, whose exact function is yet unclear, exists on the periphery of the neuraminidase active site. It was recently shown through Brownian dynamics (BD) simulations that this secondary sialic acid site may affect the association kinetics (rate) of both sialic acid and the current clinically-used drug, oseltamivir (Tamiflu, Roche) [17]. The BD calculations utilized a single crystallographic snapshot. Using DelEnsembleElec, it is clear that the mean field electrostatic potential exhibited through the ensemble-based approach substantially affects the surface potential at this secondary site (Fig. 5). Although it remains to be seen what effects would result by repeating the BD calculations using the ensemble averaged electrostatic potential values as opposed to the single static structure, we hypothesize that the ensemble-based environment would more closely align with the actual electrostatic conditions *in vivo*. Such claims have already been substantiated for other systems through work shown in [9]. Notably, the general trend of the charge fields in the region of the secondary sialic binding site of neuraminidase can be more clearly

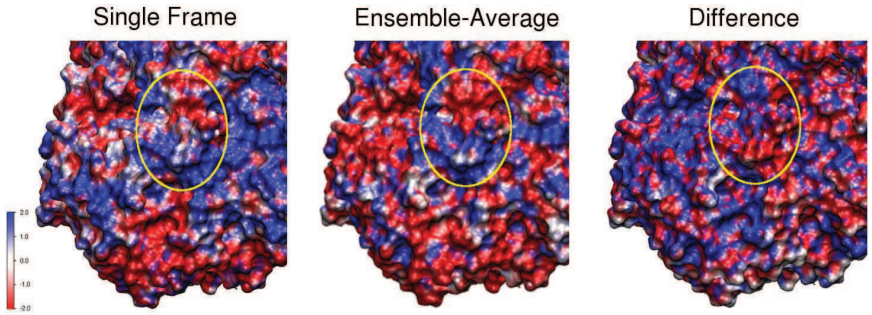


Figure 5: Neuraminidase secondary sialic acid binding site electrostatics. Close-up of the secondary sialic acid binding site of neuraminidase [17] represented as the solvent-accessible surface area and colored by electrostatic potential. The left panel depicts the surface area using potentials calculated using a single frame. The center panel depicts the result for the ensemble-averaged 500-frame trajectory. The right panel depicts the difference in potential as the single-frame potential subtracted from the ensemble-averaged potential. Units are in $kT/e$ at $T=300$K.

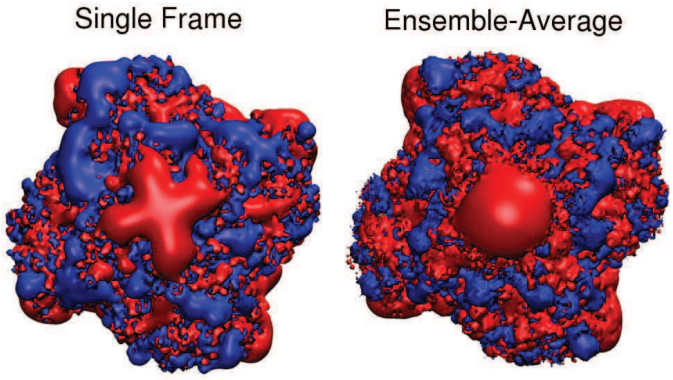


Figure 6: Overall electrostatics of neuraminidase. Top view of the neuraminidase tetramer with potential fields at positive (blue) and negative (red) $2kT/e$. The left panel depicts the electrostatics of neuraminidase when calculated using a single frame. The right panel depicts the electrostatics of the ensemble-averaged 500-frame trajectory.

identified in the ensemble-averaged view than the single frame view, where the influence of individual atoms obscure the electrical characteristic of the areas of the active site (Fig. 5). Looking more globally at the neuraminidase electrostatics, the ensemble-average electrostatic field again becomes more symmetric with the dynamic structural information (Fig. 6). Interestingly, it also appears to become slightly dampened, as compared to the single frame electrostatics calculation.

## 5 Discussion

The electrostatics of the two major influenza glycoproteins, hemagglutinin and neuraminidase, are of general interest to public health as they are both vaccine and small

molecule drug targets, respectively. In this work, we present ensemble-averaged electrostatics calculations for both proteins using the newly developed DelEnsembleElec plugin for VMD. These ensemble-averaged calculations showcase the utility of utilizing ensemble-based structural information in adding insight to outstanding biological questions.

Performing ensemble-averaged electrostatics on the hemagglutinin and neuraminidase trajectory with DelEnsembleElec and comparing the results with a single-frame DelPhi run shows that the ensemble-averaged potential surfaces are generally more symmetrical, with areas of consistent charge bias more easily visible. Additionally, the instantaneous locations of the residues in a single frame appear to cause odd shapes in the potential surface (Figs. 4 and 6, left panels). This effect is alleviated in the ensemble-averaged electrostatic calculations, since random outlier charge values often exist within individual frames of a trajectory. Though some "noise" still remains from these outlier values, including more protein frames in the averaging can potentially decrease such effects. By loading the data into a surface representation of the molecule, it becomes readily apparent that ensemble averaging filters out the influence of individual point charges. Such results provide a clearer indication of the general charge of a specific area on the molecule.

The DelEnsembleElec plugin presented in this work allows interactive and customizable preparation for running ensemble-averaged PB electrostatics using the program DelPhi. It eases the preparation of trajectories and parameters with a graphical user interface, while maintaining much of the functionality of a command-line invocation of DelPhi. The increased accuracy of the PB equation combined with the benefit of ensemble averaging allows great precision in the prediction of a biomolecule's overall electrostatics. Though we only present protein examples here, we have verified that DelEnsembleElec can handle nucleic acid- and lipid-containing systems as well, as long as the charge and radius parameters are available within the loaded structure or additional parameter files are specified.

DelEnsembleElec is freely available under the Gnu Public License. The language, applications, and libraries on which it depends are also freely available. Download instructions and a tutorial can be found at `http://amarolab.ucsd.edu/delensembleelec.html`. A link to DelEnsembleElec is provided at the Delphi webpage as well, `http://www.ces.clemson.edu/compbio/tools.html`.

## Acknowledgments

This work was funded in part by the National Institutes of Health through the NIH Director's New Innovator Award Program 1-DP2-OD007237 and through the NSF TeraGrid Supercomputer resources grant LRAC CHE060073N to R.E.A. M.Z. was supported by a grant from the Institute of General Medical Sciences, National Institutes of Health, award number 1R01GM093937-01. We thank Prof. Emil Alexov for scientific discussions.

**References**


[1] Honig, B. & Nicholls, A. Classical electrostatics in biology and chemistry. Science 268, 1144-1149 (1995).

[2] York, D. M., Darden, T. A. & Pedersen, L. G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals - a comparison of the Ewald and truncated list methods. J Chem Phys 99, 8345-8348 (1993).

[3] Greengard, L. & Rokhlin, V. A fast algorithm for particle simulations. J Comput Phys 73, 325-348 (1987).

[4] Fogolari, F., Brigo, A. & Molinari, H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. J Mol Recognit 15, 377-392, doi:10.1002/jmr.577 (2002).

[5] Harris, R. C., Bredenberg, J. H., Silalahi, A. R. J., Boschitsch, A. H. & Fenley, M. O. Understanding the physical basis of the salt dependence of the electrostatic binding free energy of mutated charged ligand-nucleic acid complexes. Biophys Chem 156, 79-87, doi:10.1016/j.bpc.2011.02.010 (2011).

[6] Rocchia, W. et al. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem 23, 128-137, doi:10.1002/jcc.1161 (2002).

[7] Case, D. A. et al. The Amber biomolecular simulation programs. J Comput Chem 26, 1668-1688, doi:10.1002/jcc.20290 (2005).

[8] Brooks, B. R. et al. Charmm - a Program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4, 187-217 (1983).

[9] Suydam, I. T. S., C. D.; Pande, V. S.; Boxer, S. G. Electric fields at the active site of an enzyme: direct comparison of experiment with theory. Science 313, 200-204 (2006).

[10] Aksimentiev, A. & Schulten, K. Imaging alpha-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map. Biophys J 88, 3745-3761, doi:10.1529/biophysj.104.058727 (2005).

[11] Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. J Mol Graph 14, 33-38, 27-38 (1996).

[12] Rocchia, W., Alexov, E. & Honig, B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. J Phys Chem B 105, 6507-6514 (2001).

[13] Newhouse, E. I. Mechanism of glycan receptor recognition and specificity switch for avian, swine, and human adapted influenza virus hemagglutinins: a molecular dynamics perspective. J Am Chem Soc 131, 17430-17442 (2009).

[14] Xu, D. et al. Distinct glycan topology for avian and human sialopentasaccharide receptor analogues upon binding different hemagglutinins: a molecular dynamics perspective. J Mol Biol 387, 465-491, doi:10.1016/j.jmb.2009.01.040 (2009).

[15] Amaro, R. E., Cheng, X., Ivanov, I., Xu, D. & McCammon, J. A. Characterizing loop dynamics and ligand recognition in human- and avian-type influenza neuraminidases via generalized born molecular dynamics and end-point free energy calculations. J Am Chem Soc 131, 4702-4709, doi:10.1021/ja8085643 (2009).

[16] Lawrenz, M. et al. Impact of calcium on N1 influenza neuraminidase dynamics and binding free energy. Proteins 78, 2523-2532, doi:10.1002/prot.22761 (2010).

[17] Sung, J. C., Van Wynsberghe, A. W., Amaro, R. E., Li, W. W. & McCammon, J. A. Role of secondary sialic acid binding sites in influenza N1 neuraminidase. J Am Chem Soc 132,

2883-2885, doi:10.1021/ja9073672 (2010).
[18] Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 32, W665-667, doi:10.1093/nar/gkh381 (2004).
[19] Li, H., Robertson, A. D. & Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. Proteins 61, 704-721, doi:10.1002/prot.20660 (2005).
[20] Newhouse, E. I. et al. Mechanism of glycan receptor recognition and specificity switch for avian, swine, and human adapted influenza virus hemagglutinins: a molecular dynamics perspective. J Am Chem Soc 131, 17430-17442, doi:10.1021/ja904052q (2009).
[21] Phillips, J. C. et al. Scalable molecular dynamics with NAMD. J Comput Chem 26, 1781-1802, doi:Doi 10.1002/Jcc.20289 (2005).
[22] Hornak, V. et al. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins-Structure Function and Bioinformatics 65, 712-725, doi:Doi 10.1002/Prot.21123 (2006).
[23] Scott, E. F., Yuhong, Z., Richard, W. P., Bernard, R. B. Constant pressure molecular dynamics simulation: the Langevin piston method. J Chem Phys 103, 4613-4621 (1995).
[24] Darden, T. Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. J Chem Phys 98, 10089-10092 (1993).
[25] Andersen, H. C. Rattle: A "velocity" version of the Shake algorithm for molecular dynamics calculations. J Comput Phys 52, 24-34 (1983).
[26] Amaro, R. E. S., R. V.; Votapka, L. V.; Li, W. W.; Walker, R. C.; Bush, R. Mechanism of 150-cavity formation in influenza neuraminidase. Nature Commun 2, 388, doi:10.1038/ncomms1390 (2011).
[27] Berendsen, H. J. C. Molecular Dynamics with coupling to an external bath. J Chem Phys 81, 3684 (1984).
[28] Baker, N. A. Poisson-Boltzmann methods for biomolecular electrostatics. Methods Enzymol 383, 94-118, doi:10.1016/S0076-6879(04)83005-2 (2004).
[29] Baker, N. A. & McCammon, J. A. Structural Bioinformatics. 427-440 (John Wiley & Sons, 2003).
[30] von Itzstein, M. The war against influenza: discovery and development of sialidase inhibitors. Nat Rev Drug Discov 6, 967-974, doi:10.1038/nrd2400 (2007).