

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

PB-AM: An open-source, fully analytical linear poisson-boltzmann solver.

### Permalink

<https://escholarship.org/uc/item/7qn8w312>

### Journal

Journal of computational chemistry, 38(15)

### ISSN

0192-8651

### Authors

Felberg, Lisa E  
Brookes, David H  
Yap, Eng-Hui  
[et al.](#)

### Publication Date

2017-06-01

### DOI

10.1002/jcc.24528

Peer reviewed

# PB-AM: An Open-Source, Fully Analytical Linear Poisson-Boltzmann Solver

Lisa E. Felberg,<sup>[a]</sup> David H. Brookes,<sup>[b]</sup> Eng-Hui Yap,<sup>[c]</sup> Elizabeth Jurrus,<sup>[d,e]</sup>  
Nathan A. Baker,<sup>[f,g]</sup> and Teresa Head-Gordon<sup>\*[a,b,h,i]</sup>

We present the open source distributed software package Poisson-Boltzmann Analytical Method (PB-AM), a fully analytical solution to the linearized PB equation, for molecules represented as non-overlapping spherical cavities. The PB-AM software package includes the generation of outputs files appropriate for visualization using visual molecular dynamics, a Brownian dynamics scheme that uses periodic boundary conditions to simulate dynamics, the ability to specify docking criteria, and offers two different kinetics schemes to evaluate biomolecular association rate constants. Given that PB-

AM defines mutual polarization completely and accurately, it can be refactored as a many-body expansion to explore 2- and 3-body polarization. Additionally, the software has been integrated into the Adaptive Poisson-Boltzmann Solver (APBS) software package to make it more accessible to a larger group of scientists, educators, and students that are more familiar with the APBS framework.  
© 2016 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24528

## Introduction

The preferential association of biomolecules in the cell, ranging from proteins to nucleic acids to small metabolites, is crucial for cellular function. The first events in the molecular recognition process among multiple molecular constituents involve long-range diffusional association over large spatial scales. However, these dynamical encounters have proved difficult to model at an atomistic level as molecular dynamics is too computationally intensive to capture the large spatial and temporal scales over which many macromolecule associations occur. Fortunately coarse graining of the participating biomolecules and their environment, combined with simulations using stochastic dynamics, can be just as insightful as the all-atom deterministic dynamics. At these scales electrostatic interactions dominate, such that the coarse-graining strategy can rely on the popular Poisson-Boltzmann (PB) continuum mean field treatment that forms the basis of Gouy–Chapman theory<sup>[1,2]</sup> in electrochemistry, Derjaguin–Landau–Verwey–Overbeek theory in colloid chemistry,<sup>[3,4]</sup> and under the low field linearized Poisson-Boltzmann approximation, the Debye–Hückel theory in solution chemistry.<sup>[5]</sup>

There is a range of software packages currently available for solving the linearized Poisson-Boltzmann equation (LPBE). Numerical methods, such as finite-difference (FD) and finite-element (FE) packages including Adaptive Poisson-Boltzmann Solver (APBS),<sup>[6]</sup> DelPhi,<sup>[7]</sup> MEAD,<sup>[8]</sup> UHBD,<sup>[9]</sup> ZAP,<sup>[10]</sup> and modules in Amber<sup>[11,12]</sup> and CHARMM,<sup>[13,14]</sup> handle arbitrary dielectric boundaries by solving the PBE on a grid or mesh. Due to their numerical challenges, many essential features of the PBE were not treated robustly in these software packages; for example, the point charge singularity and the enforcement of electrostatic continuity across the dielectric boundary. Another limiting factor of these methods is the tradeoff between grid refinement and accuracy given the  $O(m^3)$  scaling for 3D electrostatics, where  $m$  is the number of grid points in one dimension. Some of these

[a] L. E. Felberg, T. Head-Gordon

Department of Chemical and Biomolecular Engineering, University of California Berkeley, Berkeley, California 94720  
E-mail: thg@berkeley.edu

[b] D. H. Brookes, T. Head-Gordon

Department of Chemistry, University of California Berkeley, Berkeley, California 94720

[c] E.-H. Yap

Department of Systems and Computational Biology, Albert Einstein College of Medicine, Bronx, New York 10461

[d] E. Jurrus

Division of Computational and Statistical Analytics, Pacific Northwest National Laboratory, Richland, Washington 99352

[e] E. Jurrus

Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah 84112

[f] N. A. Baker

Advanced Computing, Mathematics, and Data Division, Pacific Northwest National Laboratory, Richland, Washington 99352

[g] N. A. Baker

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

[h] T. Head-Gordon

Department of Bioengineering, University of California Berkeley, Berkeley, California 94720

[i] T. Head-Gordon

Chemical Sciences Division, Lawrence Berkeley National Labs, Berkeley, California 94720

Contract grant sponsor: Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy (to T.H.G. and D.H.B.); Contract grant number: DE-AC02-05CH11231; Contract grant sponsor: National Science Foundation Graduate Research Fellowship (to L.F.); Contract grant number: DGE 110640; Contract grant sponsor: National Biomedical Computation Resource (APBS); Contract grant sponsor: National Institutes of Health; Contract grant number: GM069702; Contract grant sponsor: National Institutes of Health (VMD software); Contract grant numbers: 9P41GM104601 and 5R01GM098243-02; Contract grant sponsor: National Energy Research Scientific Computing Center, a DOE Office of Science User Facility Office of Science of the U.S. Department of Energy; Contract grant number: DE-AC02-05CH11231

© 2016 Wiley Periodicals, Inc.

issues have been mitigated for the energy: charge singularities have been treated with regularization schemes,<sup>[15]</sup> induced surface charges<sup>[16]</sup> and other decomposition methods,<sup>[17]</sup> molecular surface definitions and discontinuities across that boundary have been addressed with the matched interface and boundary method<sup>[18]</sup> and advanced grid methods are also implemented.<sup>[12,19]</sup> However, unlike the energetic terms, force calculations still pose many numerical issues for FE and FD methods. The most definitive method, known as the “virtual work” approach<sup>[20]</sup> has been implemented to account for full mutual polarization in FD and FE methods, but is limited by computational efficiency as gradients are calculated numerically, requiring a recalculation for the electrostatic energy after small displacements in each Cartesian direction.

Another approach is the boundary element method (BEM) that is the main feature of software packages such as AFMPB,<sup>[21]</sup> PyGBe,<sup>[22,23]</sup> and TABI.<sup>[24]</sup> BEM approaches offer some advantages over FE and FD methods as there is an analytical treatment of singular charges, and explicit treatment of the dielectric discontinuity and boundary conditions. However, BEMs also have some limitations, including an increasingly large dense matrix that scales with system size and has severe memory requirements, singular surface integrals, and issues with mesh generation. In contrast to FD and FE methods, BEMs can efficiently compute forces, but they only include self-polarization as opposed to full mutual polarization, and have yet to be demonstrated as efficient enough for dynamical simulations.

An alternative to the numerical FE, FD, and BEM methods is to consider analytical solutions instead. As far as we are aware, the first analytical LPBE solution was solved by Kirkwood for one spherical macromolecule.<sup>[25]</sup> However, the treatment to describe mutual polarization requires at least two distinct spheres, and many different partial and approximate solutions to the mutual polarization of two or more macromolecules using the LPBE have been proposed.<sup>[26–30]</sup> In 2006, Lotan and Head-Gordon derived the first completely general analytical solution to the LPBE,<sup>[31]</sup> including the forces and torques due to this potential, that permits analytical calculations to arbitrary large assemblies of interacting molecules. While the primary drawback of the analytical solutions are that they are restricted to idealized geometries such as spheres, they do have the benefit that the boundary conditions are solved completely, and for the Poisson-Boltzmann analytical model (PB-AM)<sup>[31]</sup> the mutual polarization is accounted for accurately.

While the PB-AM model<sup>[31]</sup> is computationally efficient, we believe that the model’s theoretical formalism has made it inaccessible to many and inhibited its potential use for application on a large scale. Here we present an expanded, open-source software implementation of PB-AM, as well as a number of python-based utilities for creating files for multi-dimensional visualization for the electrostatic potential (ESP) and forces and torques, a Brownian Dynamics simulation module that models the stochastic dynamics of multiple macromolecule interactions,<sup>[32–34]</sup> the ability for users to define docking criteria and to simulate biomolecular association rate constants, as well as exploration of the many-body expansion (MBE) for the mutual polarization.<sup>[35]</sup> The PB-AM open source software is easy to

download and install and has a simple input file structure and modular design to encourage both ease of use and extensibility to other software packages. Additionally, the PB-AM model has been integrated into the APBS (<http://www.poissonboltzmann.org/>) to allow accessibility of the PB-AM capabilities to a larger group of scientists, educators and students.

The article is organized as follows. After a brief review of the PB-AM theory developed by Lotan and Head-Gordon in PB-AM Theory section, we describe the software framework of the PB-AM stand-alone code and its integration into the APBS software package in Software Architecture section. In Electrostatic Potential Visualization section, we demonstrate the PB-AM software’s visualization outputs for the ESP using 2D cross-sections as well as 3D isosurfaces that feed into the visual molecular dynamics<sup>[36]</sup> (VMD) program. In Dynamical Trajectories and Analysis Tools section, we introduce the dynamics capabilities using PB-AM that utilizes a standard Brownian dynamics scheme, permits users to define docking criteria, and offers two distinct ways to calculate bimolecular rate constants, one of which is especially suited to simulations involving large number of molecules. In Many-Body Expansion of the Electrostatic Energy section, we introduce the MBE that allows the mutual polarization to be deconstructed into a direct polarization model, as well as approximate mutual polarization models. Finally, we provide a brief summary in Conclusion section and plans for future work.

## PB-AM Theory

The derivation details of PB-AM have been reported previously,<sup>[31]</sup> but the main points can be summarized as follows. The ESP of the system at any point  $\mathbf{r}$  is governed by the linearized form of the PB equation

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})] - \epsilon(\mathbf{r}) \kappa^2 \Phi(\mathbf{r}) = 4\pi \rho(\mathbf{r}) \quad (1)$$

For the case of spherical cavities, we can solve eq. (1) at a point  $t = [r, \theta, \phi]$ , by dividing the system into inner sphere and outer sphere regions, and enforcing a set of boundary conditions that stipulate the continuity of the ESP and the electrostatic field at the surface of each sphere. The ESP inside molecule  $i$  is described by:

$$\Phi_m^{(i)}(t) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \left( \frac{E_{n,m}^{(i)}}{\epsilon_p r^{n+1}} + B_{n,m}^{(i)} r^n \right) Y_{n,m}(\theta, \phi) \quad (2)$$

where  $\epsilon_p$  is the interior dielectric,  $B^{(i)}$  is a vector of unknowns that will be determined through the application of the boundary conditions,  $E^{(i)}$  is the multipole expansion of the charges inside molecule  $i$ .

$$E_{n,m}^{(i)} = \sum_{j=1}^{M_i} q_j^{(i)} \left( \rho_j^{(i)} \right)^n Y_{n,m}(\vartheta_j^{(i)}, \varphi_j^{(i)}) \quad (3)$$

where  $M_i$  is the number of charges in molecule  $i$ ,  $q_j^{(i)}$  is the magnitude of the  $j$ -th charge, and  $\rho_j^{(i)} = [\vartheta_j^{(i)}, \varphi_j^{(i)}]$  is its position in spherical coordinates and  $Y_{n,m}$  are the spherical harmonics. The general form of the potential outside all

molecules (in a coordinate frame whose origin is the center of molecule  $i$ ) is:

$$\Phi_{\text{out}}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^n \left( \frac{A_{n,m}^{(i)}}{r^{n+1}} e^{-\kappa r} \hat{k}_n(\kappa r) + L_{n,m}^{(i)} r^n \hat{i}_i(\kappa r) \right) Y_{n,m}(\theta, \phi) \quad (4)$$

where the coefficients  $L^{(i)}$  in eq. (4) are a re-expansion of the external potential coefficients  $A^{(j)}$ ,  $j \neq i$  of all other molecules in the system. It is defined as:

$$L^{(i)} = \sum_{j=1, j \neq i}^N T^{(ij)} \cdot A^{(j)} \quad (5)$$

where  $T^{(ij)}$  is the linear re-expansion operator that transforms a multipole expansion at  $\mathbf{c}^{(j)}$  to a local (Taylor) expansion at  $\mathbf{c}^{(i)}$ . This operator is described in detail in our previous work.<sup>[31]</sup> The use of the  $T^{(ij)}$  operators allows us to represent the potentials due to all molecules in the coordinate frame of a single molecule, a mathematical feature that is central to obtaining an analytical solution. As the  $B^{(i)}$  depend on  $A^{(i)}$ , the application of the boundary conditions leads to the following compact solution

$$\mathbf{A} = \Gamma \cdot (\Delta \cdot \mathbf{T} \cdot \mathbf{A} + \mathbf{E}) \quad (6)$$

where  $\mathbf{A}$  is a matrix of vectors, one for each molecule in the system, representing the effective multipole expansion of the charge distributions of each molecule, the  $\Gamma$  matrix is a dielectric boundary-crossing operator, and the  $\Delta$  matrix is a cavity polarization operator.

Using this formalism, physical properties of the system, such as interaction energy, forces and torques may be computed. The interaction energy of each molecule,  $\Omega^{(i)}$ , is the product of the molecule's total charge distribution (from fixed and polarization charges) with the potential due to external sources. This is computed as the inner product between the molecule's multipole expansion, and the multipole expansions of the other molecules in the system,  $L^{(i)}$  as follows:  $\mathbf{A}^{(i)}$

$$\Omega^{(i)} = \frac{1}{\epsilon_s} \langle L^{(i)}, \mathbf{A}^{(i)} \rangle \quad (7)$$

which allows us to define the force, which is computed as the gradient of the interaction energy with respect to the position of the center of molecule  $i$ :

$$\mathbf{F}^{(i)} = -\nabla_i \Omega^{(i)} = \frac{-1}{\epsilon_s} \left( \langle \nabla_i L^{(i)}, \mathbf{A}^{(i)} \rangle + \langle L^{(i)}, \nabla_i \mathbf{A}^{(i)} \rangle \right) \quad (8)$$

By definition, the torque on a charge in the molecule is the cross product of its position relative to the center of mass of the molecule with the force it experiences. The total torque on the molecule is a linear combination of the torque on all charges of the molecule.

$$\boldsymbol{\tau}^{(i)} = \frac{1}{\epsilon_s} \left[ \mathbf{H}^{(i),x}, \mathbf{H}^{(i),y}, \mathbf{H}^{(i),z} \right] \times \left[ \nabla_i L^{(i)} \right] \quad (9)$$

Please see previous publication<sup>[31]</sup> for details on the PB-AM solver.

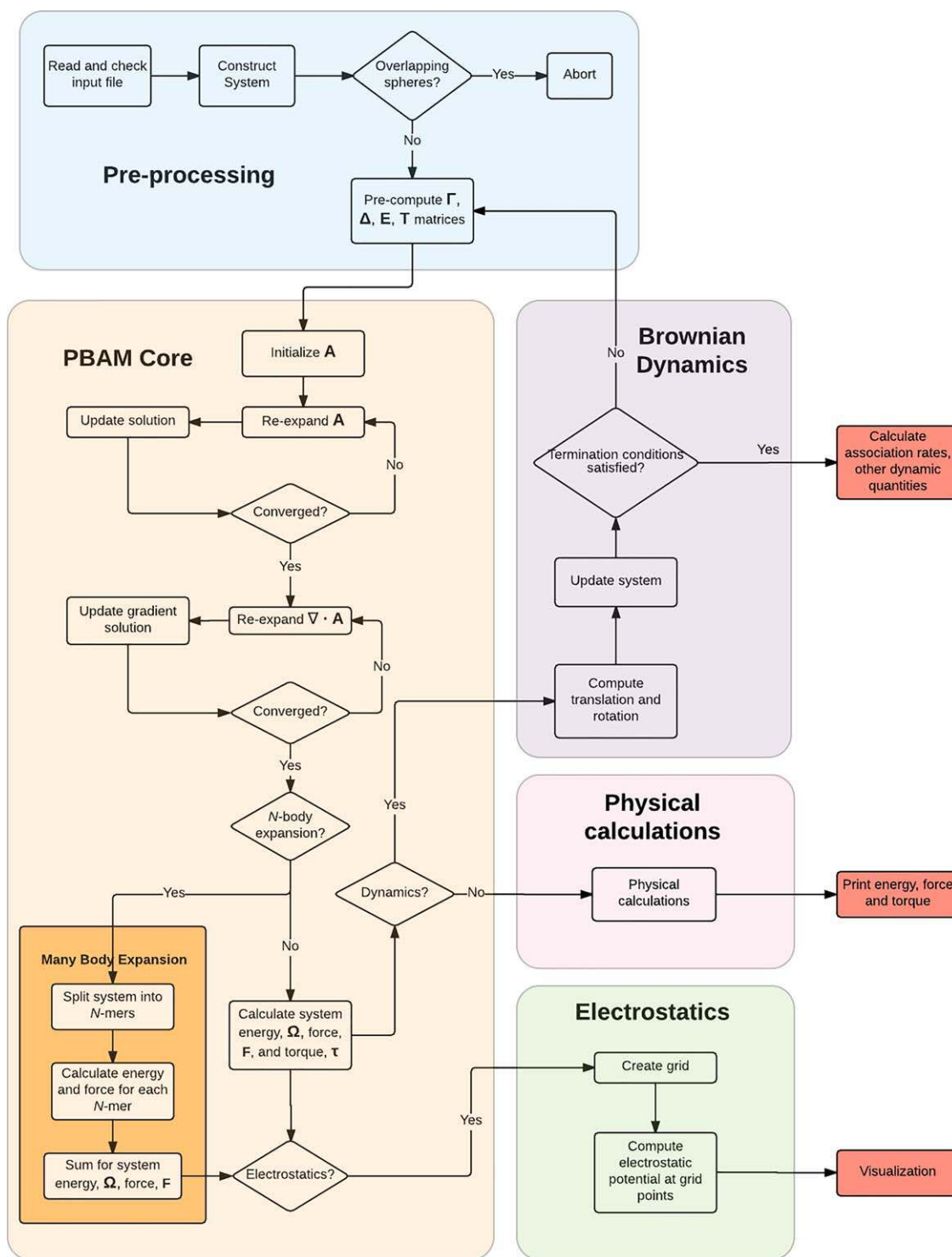
## Software Architecture

The stand-alone PB-AM code has been completely re-written from the original code base development<sup>[31]</sup> in C++11 to include a modular interface, simplified input files, expanded examples, unit testing, and an automated build system (CMake). It also includes a set of utilities, written in Python, to aid in visualization tasks and perform additional simulations. The new Application Programming Interface has been designed to focus on five important utilities of the program: (1) ESP visualization, (2) energy, force, and torque calculation for molecule-molecule interactions, (3) use of the MBE to formulate direct and mutual polarization models, (4) dynamical simulations using Brownian dynamics, and (5) different ways to calculate rate constants for biomolecule association. The input file structure is simple; each line containing a *keyword value* pair, in most cases, and for most basic usage, the program can be run with an input file of three lines. Additionally, the outputs have been expanded to formats readable in VMD,<sup>[36]</sup> such as *dx* and *xyz* files. Finally, an extensive user manual and website ([https://davas301.github.io/pb\\_solvers/](https://davas301.github.io/pb_solvers/)) have been developed with examples and post-processing examples. PB-AM is available for download through the GitHub site: [https://github.com/davas301/pb\\_solvers](https://github.com/davas301/pb_solvers). Figure 1 shows the workflow of the software and its components.

In addition to the stand-alone code, the PB-AM code has been integrated into the APBS software package that solves the equations of continuum electrostatics for large biomolecular assemblages. This software is the central computational element for computational research involving the PB equation and is, a unique software package that solves the equations of continuum electrostatics for large biomolecular assemblages. APBS plays an important role in the structural and computational biology research community and the proposed research ensures its continued availability and support. Specifically, APBS addresses three key technology challenges for understanding solvation and electrostatics in biomedical applications: (1) accurate and efficient theories and models for biomolecular solvation and electrostatics, (2) robust and scalable software for applying those theories to biomolecular systems, and (3) mechanisms for sharing and analyzing biomolecular electrostatics data in the scientific community. PB-AM has been fully integrated into the APBS software. The user invokes the code with a keyword in the ELEC section of an APBS input file. Some additional keywords are required and are documented on the APBS website. The incorporation of PB-AM into APBS will allow for PB-AM to be available to an audience of thousands of users, and is made available through the APBS site: <http://www.poissonboltzmann.org/>.

## ESP Visualization

One of the most popular uses of PBE solvers is to provide rapid analysis of the ESPs around macromolecules and their assemblies. Therefore, like many other PBE solvers, our PB-AM approach has the capability of producing output files for visualizing the ESP of a system of arbitrary complexity. To this end, we have provided several visualization options that a user may



**Figure 1.** Software workflow for the PB-AM model and its utilities. The stand alone PB-AM code has been completely re-written from the original development<sup>[31]</sup> in C++ 11, and has four important utilities: electrostatic potential visualization, energy, force, and torque calculations for molecule-molecule interactions, use of the many-body expansion to formulate approximate direct and mutual polarization models, dynamical simulations using Brownian dynamics, which can generate many interesting outputs, for example, different ways to calculate rate constants for biomolecule association. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

select, including two-dimensional ESP cross-sections and isosurface potentials (Fig. 2), which may be loaded and viewed in the program VMD.<sup>[36]</sup> In addition to the ESPs, it is also quite straightforward to visualize of forces and torques on and around the center-of-mass.

## Dynamical Trajectories and Analysis Tools

In addition to electrostatics calculations, the PB-AM software includes dynamics capabilities. The computed expressions for force and torque have been incorporated into a Brownian dynamics simulation protocol adapted from Ermak and McCammon.<sup>[32]</sup>

Each molecule in the system is treated as an independent rigid particle, and molecule-molecule overlap is not allowed. Assuming no hydrodynamic effects, the translational and rotational displacement,  $\Delta \mathbf{r}_i$  and  $\Delta \theta_i$ , respectively, are computed as:

$$\Delta \mathbf{r}_i = \frac{D_{i,\text{trans}}}{k_B T} \Delta t \mathbf{f}_i + \mathbf{S}_i(\Delta t) \quad (10a)$$

$$\Delta \theta_i = \frac{D_{i,\text{rot}}}{k_B T} \Delta t \boldsymbol{\tau}_i + \boldsymbol{\Theta}_i(\Delta t) \quad (10b)$$

where  $\mathbf{S}_i$  is the stochastic displacement, and  $\boldsymbol{\Theta}_i$  is the stochastic rotation, have the following properties in dimension  $\alpha = x, y, z$ :

$$\langle S_\alpha \rangle = 0, \quad \langle S_\alpha^2 \rangle = 2 D_{i,\text{trans}} \Delta t \quad (11a)$$

$$\langle \Theta_\alpha \rangle = 0, \quad \langle \Theta_\alpha^2 \rangle = 2 D_{i,\text{rot}} \Delta t \quad (11b)$$

The PB-AM software allows a generalized simulation procedure for a variety of terminations conditions, including (a) time: the simulation will terminate when it has run for  $t$  picoseconds, (b) coordinates: the simulation will terminate when a specified particle has diffused beyond a certain point in space specified by the user, (c) docking: the simulation will terminate when a pair of specified atoms in the simulated molecules are within a specified cutoff; this docking criteria is specified by the user as a contact list at the start of the simulation. For details on BD simulation in PB-AM, please see our user manual.

One application of dynamics of particular interest is the calculation of bimolecular association rate constants and the ability to specify the docking criteria when association is complete. For the case of PB-AM, which represents each molecule as one hard spherical object, the boundary of the sphere prevents the physical proximity of contact pairs to be defined based on an atomistic geometry. Instead, the locations of the atomistic contact points, from the contact list, are projected onto the surface of the sphere to create a reactive surface patch, and the BD simulation terminates at the docking criteria when these surface patches are within 0.1 Å.

For calculations of the rate constants, we have incorporated into the PB-AM software package two independent approaches. The first is the Northrup–Allison–McCammon (NAM) method.<sup>[39,40]</sup> The NAM method is based on the Smoluchowski equation, and its underlying assumption is that the mobile molecule  $B$  only experiences forces from stationary molecule  $A$ ; that is, concentration effects that modulate intra-species interactions  $A:A$  and  $B:B$  are not taken into account. The NAM method models the rate calculation as an analysis across a series of flux surfaces: an inner region, bounded by a radial distance  $b$ , is defined where inter-molecular forces are anisotropic and must be evaluated explicitly with BD. In this region, we can numerically evaluate the docking frequency,  $\delta$ , that is, the ratio of successfully docked trajectories to total trajectories simulated. For simulated distances outside  $b$ , any inter-molecular forces are approximately centrosymmetric and the diffusion rate constant  $k_D(b)$  can be evaluated analytically using the Smoluchowski equation.<sup>[39]</sup>

$$k(R) = \left[ \frac{1}{4\pi D_{\text{trans}}} \int_R^\infty \frac{\exp\left(\frac{U(r)}{k_B T}\right)}{r^2} dr \right]^{-1} \quad (12)$$

Beyond the surface at  $R = b$ , a second surface at  $R = q$  is chosen such that  $q \gg b$ , and the simulation can be truncated, and the diffusional rate beyond  $q$  can also be evaluated from the Smoluchowski equation,  $k_D(q)$ . By accounting for re-crossings across, the two flux surfaces at  $b$  and  $q$ , we can evaluate intrinsic bimolecular collision rate,  $k$ , is as follows:

$$k = \frac{k_D(b)\delta}{1 - (1-\delta)k_D(b)/k_D(q)} \quad (13)$$

A more accurate rate constant calculation method must also simulate a system with multiple copies of  $A$  and  $B$  that are allowed to interact with each other. Given that such interactions are dominated by electrostatics at large distances, this demands an efficient algorithm capable of computing the electrostatic forces and torques for multiple molecules on the fly as in PB-AM. In this second approach, it is possible to evaluate the mean first passage time (MFPT) under periodic boundary conditions, by considering the second-order rate equation of the association of molecules  $A$  and  $B$ :

$$\frac{d[AB]}{dt} = k[A][B] \quad (14)$$

In this case, the user can specify the concentrations  $[A]$  and  $[B]$  and the fraction of trajectories docked at time  $t$  is given by

$$P_{\text{docked}}(t) = \frac{\text{No. of systems with } t_i \leq t}{N} \quad (15)$$

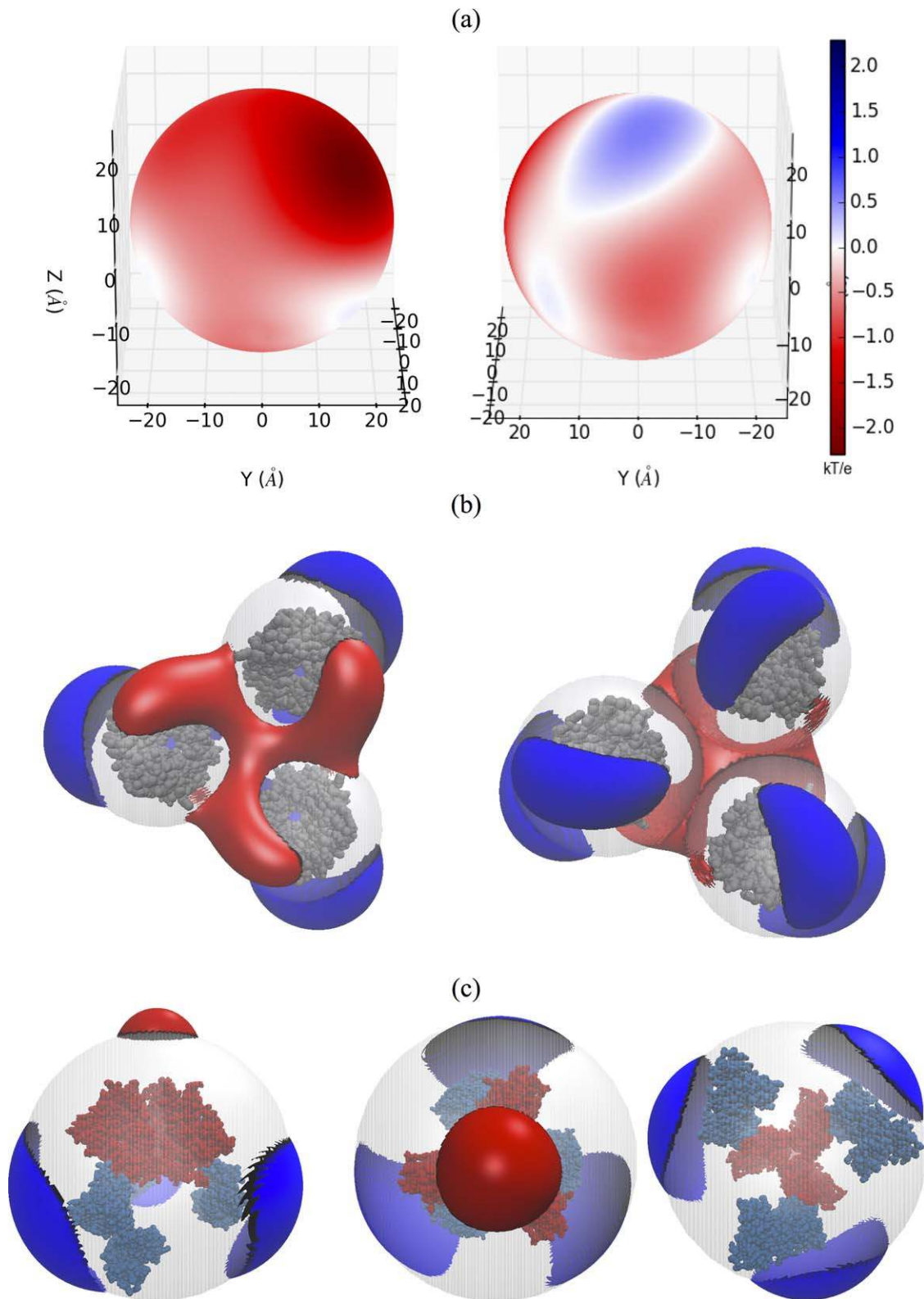
At present one can use PB-AM to evaluate the pseudo-first order rate constant  $k'$  by fitting the plotted data of  $P_{\text{docked}}(t)$  against  $t$ , to the functional form

$$P_{\text{docked}}(t) = 1 - \exp(-k'(t - \tau_d)) \quad (16)$$

where  $\tau_d$  is the dead time required for the system to equilibrate, and the bimolecular rate constant is computed from

$$k = \frac{k'}{[A]_{t=0}} \quad (17)$$

The barnase-barstar association kinetics has been extensively characterized, both experimentally<sup>[41]</sup> and computationally using the NAM method.<sup>[42]</sup> We compare the NAM method versus a multi-molecular simulation involving multiple barnase and barstar proteins ( $\sim 125$  proteins total). Both systems were performed under the following conditions:  $T = 298.15$  K, dielectric constants of 78 (protein) and 4 (solvent), salt concentration of 0.05 M. A variable time step with a minimum of 2 ps was used. At each time step, the system is solved with a polarization cutoff of 10 Å and a force cutoff of 100 Å. The multi-molecular simulation was run with periodic boundary conditions with a box length of 320 Å to ensure that the minimum image conventions were obeyed. In Table 1, we compare these results against PB-AM using both NAM



**Figure 2.** Different visualizations of the electrostatic potentials from the PB-AM model. All systems are at 0.0M salt concentration, 7 pH, with protonation states calculated using PROPKA,<sup>[37]</sup> interior dielectric of 2, and solution dielectric of 78. a) Three-dimensional surface potential from two different vantage points of the coarse-grained Barstar protein. The molecule has been coarse-grained into a single sphere of radius 21.8 Å that encompasses all molecule atoms, and the potential is depicted at the surface of the coarse-grain sphere. The color scale for each image pair is given by color bar on the right, in units of  $k_B T/e$ . (b) Three-dimensional isosurfaces for Omp32 Porin trimer from the view of (c) the exoplasmic face (channel entrance). The negative isosurface (red) forms a funnel that can direct anions from the environment toward the channel. The periplasmic face (channel exit). The positive surface (blue) at the channel exit may further enhance anion transport through the channel. The blue isosurface is drawn at  $1.0 k_B T/e$  and the red at  $-1.0 k_B T/e$ . The channel proteins are represented by a gray solvent-excluded surface. (d) The HIV glycoprotein binds to CD4 proteins (PDB structure 4NCO<sup>[38]</sup>) from three vantage points. The upper hemisphere is a virus membrane-spanning portion of the glycoprotein, while the lower hemisphere is connected to 3 CD4 proteins that extend from the surface of the cell wall. These two distinct binding regions of the protein are distinguished here by the location of the electrostatic potential isosurfaces, with the CD4 creating a positive potential, and the portion attached to virus membrane creating a negative potential. The blue isosurface is drawn at  $1.0 k_B T/e$  and the red at  $-1.0 k_B T/e$ . [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**Table 1.** The barnase-barstar rate of association using different PBE models and rate constant protocols NAM and MFPT against experiment and using an atomistic docking criteria from Gabdouliline and Wade

Model and rate method	[Barnase] (concentration [M]): [Barstar] (concentration [M])	Rate constant values, $k$ [ $M^{-1}s^{-1}$ ]
Experiment <sup>[39]</sup>	N/A	$2.86 \pm 0.28 \times 10^8$
Gabdouliline and Wade <sup>[40]</sup>	1 ( $5.068 \times 10^{-5}$ ): 1 ( $5.068 \times 10^{-5}$ )	$3.88 \pm 0. \times 10^8$
PB-AM using NAM	1 ( $5.068 \times 10^{-5}$ ): 1 ( $5.068 \times 10^{-5}$ )	$7.53 \pm 0.25 \times 10^7$
PB-AM using MFPT	124 ( $6.284 \times 10^{-3}$ ): 1 ( $5.068 \times 10^{-5}$ )	$9.58 \pm 0.47 \times 10^7$

and the multi-protein MFPT simulations, showing that the simplified dielectric boundary is in very good qualitative agreement with past efforts, and thus, a useful way to rapidly evaluate kinetics for biomolecule association. In comparison with the results of Gabdouliline and Wade, we believe that differences in computed rate constants are likely due to differences in molecular representations. All other system conditions held constant (protein model, force field, docking definition), the crowding conditions increase the docking rate by approximately ~20%.

## MBE of the Electrostatic Energy

The formalism of the PB-AM model allows mutual polarization of all molecules in the system to be treated analytically, but as the number of molecules in the system increases, this N-body problem can become increasingly time intensive. The MBE allows us to expand the energy, and forces and torques as well, can be calculated in terms of simpler and independent 1-body, 2-body, 3-body, and so forth components.

$$U_N = U_1 + U_2 + U_3 + \dots \quad (18)$$

where

$$U_1 = \sum_{i=1}^N U(i); \quad \Delta U_2 = \sum_{i=1}^{N-1} \sum_{j=i+1}^N U(i,j) - U(i) - U(j) \quad (19)$$

$$\Delta U_3 = \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^N U(i,j,k) - U(i,j) - U(i,k) - U(j,k) + U(i) + U(j) + U(k) \quad (20)$$

We have shown that the MBE converges quickly at the 2-body level with small Debye lengths, whereas the 3-body truncation of eq. (12) is excellent for larger Debye lengths.<sup>[35]</sup> This allows for a different type of refactoring of the calculations while still leading to a highly accurate approximation to the PB-AM model. The computational benefits to the approach are evident from Table 2, using a cubic grid of barnase and barstar molecules, in similar configurations to the MFPT initialization. We can achieve a 2–3X speed-up over the full-calculation using only four cores, and if we take advantage of the independent nature of the dimer and trimer calculations and increase the core count to 64, we can increase the speed-up to 8X.

**Table 2.** Comparison of timings for full mutual polarization and the MBA

[Barnase]: [Barstar]	Timings (s)		
	Full Mutual (4 cores)	3-body MBE (4 cores)	3-body MBE (64 cores)
7:1	1.07	3.19	0.47
63:1	376.22	185.11	50.81
124:1	2307.71	1612.34	297.01

The systems are comprised of a cubic grid of barnase and barstar molecules, with the given ratio of Barnase to Barstar molecules.

## Conclusion

We have described the release of the PBE solver, PB-AM, a complete analytical solution for both self- and full-mutual polarization, which allows users to analyze electrostatics and stochastic dynamics of complex biomolecular systems, under the assumption of idealized spherical geometries. PB-AM can be downloaded stand-alone or as part of the distributed APBS software package.

Obviously, the simple geometric shape of the low dielectric spherical cavity containing the complex charge distribution can result in differences with that of a more detailed molecular shape handled by other LPBE formulations. We have overcome this limitation using our semi-analytic solution to the PBE (PB-SAM)<sup>[43,44]</sup> that expands the analytic formalism to describe molecular boundaries, and it will be made available in a future release of the PB-AM software package. Even so, we have demonstrated that meaningful rate constants can be evaluated under the idealized geometries inherent to the PB-AM model, as they were shown to be in good agreement with experiment and the PB-SAM solver that uses realistic geometric boundaries.

As a freely available, open-source package, we hope that other users will find value in expanding the capabilities of the PB-AM code. In addition to the future release of PB-SAM, which describes more realistic molecular boundaries, there are other important additions that should be included such as hydrodynamic interactions and other kinetic schemes.

## Acknowledgments

THG wishes Charles Brooks III the best on the occasion of his 60th birthday. Since I worked on the Poisson Boltzmann problem as one of the first projects as a graduate student in his group I think it fitting that the current status on PBE solvers be included in his *Festschrift!*

**Keywords:** electrostatics · Brownian dynamics · linearized Poisson-Boltzmann equation

How to cite this article: L. E. Felberg, D. H. Brookes, E.-H. Yap, E. Jurrus, N. A. Baker, T. Head-Gordon. *J. Comput. Chem.* **2017**, *38*, 1275–1282. DOI: 10.1002/jcc.24528

[1] D. L. Chapman, *Philos. Mag. Ser. 6* **1913**, *25*, 475.



- [2] G. Gouy, *C. R. Hebd. Seances Acad. Sci.* **1910**, 150, 1652.
- [3] B. Derjaguin, L. Landau, *Prog. Surf. Sci.* **1993**, 43, 30.
- [4] E. J. W. Verwey, *Phillips Res. Rep.* **1945**, 1, 33.
- [5] P. Debye, E. Hueckel, *Phys. Z.* **1923**, 24, 185.
- [6] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 10037.
- [7] L. Li, C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, L. Wang, N. Smith, M. Petukh, E. Alexov, *BMC Biophys.* **2012**, 5, 9.
- [8] D. Bashford, In Scientific Computing in Object-Oriented Parallel Environments: First International Conference, ISCOPE 97 Marina del Rey, California, USA December 8–11, 1997 Proceedings; Y. Ishikawa, R. R. Oldehoeft, J. V. W. Reyniers, M. Tholburn, Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, **1997**; pp. 233–240.
- [9] M. E. Davis, J. D. Madura, B. A. Luty, J. A. Mccammon, *Comput. Phys. Commun.* **1991**, 62, 187.
- [10] J. A. Grant, B. T. Pickup, A. Nicholls, *J. Comput. Chem.* **2001**, 22, 608.
- [11] D. A. Case, R. M. Betz, W. Botello-Smith, D. S. Cerutti, T. E. Cheatham, III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, AMBER 2016, University of California, San Francisco, **2016**.
- [12] R. Luo, L. David, M. K. Gilson, *J. Comput. Chem.* **2002**, 23, 1244.
- [13] W. Im, D. Beglov, B. Roux, *Comput. Phys. Commun.* **1998**, 111, 59.
- [14] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, 4, 187.
- [15] W. Geng, S. Yu, G. Wei, *J. Chem. Phys.* **2007**, 127, 114106.
- [16] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, B. Honig, *J. Comput. Chem.* **2002**, 23, 128.
- [17] Q. Cai, J. Wang, H. K. Zhao, R. Luo, *J. Chem. Phys.* **2009**, 130, 145101.
- [18] Y. C. Zhou, S. Zhao, M. Feig, G. W. J. Wei, *Comput. Phys.* **2006**, 213, 1.
- [19] M. J. Holst, In Numerical Computing Group; University of Illinois at Urbana-Champaign: Urbana-Champaign, IL, **1993**.
- [20] M. E. Davis, J. A. Mccammon, *Chem. Rev.* **1990**, 90, 509.
- [21] B. Lu, X. Cheng, J. Huang, J. A. McCammon, *Comput. Phys. Commun.* **2010**, 181, 1150.
- [22] C. D. Cooper, L. A. Barba, *Comput. Phys. Commun.* **2016**, 202, 23.
- [23] C. D. Cooper, J. P. Bardhan, L. A. Barba, *Comput. Phys. Commun.* **2014**, 185, 720.
- [24] W. H. Geng, R. Krasny, *J. Comput. Phys.* **2013**, 247, 62.
- [25] J. G. Kirkwood, *J. Chem. Phys.* **1934**, 2, 351.
- [26] A. T. Fenley, J. C. Gordon, A. Onufriev, *J. Chem. Phys.* **2008**, 129, 075101.
- [27] R. B. McClurg, C. F. Zukoski, *J. Colloid Interface Sci.* **1998**, 208, 529.
- [28] G. D. Phillies, *J. Chem. Phys.* **1974**, 60, 976.
- [29] J. E. Sader, A. M. Lenhoff, *J. Colloid Interface Sci.* **1998**, 201, 233.
- [30] R. R. Gabdouliline, R. C. Wade, *J. Phys. Chem.-US* **1996**, 100, 3868.
- [31] I. Lotan, T. Head-Gordon, *J. Chem. Theory Comput.* **2006**, 2, 541.
- [32] D. L. Ermak, J. A. Mccammon, *J. Chem. Phys.* **1978**, 69, 1352.
- [33] T. Head-Gordon, C. L. Brooks, *J. Phys. Chem.* **1987**, 91, 3342.
- [34] T. Head-Gordon, C. L. Brooks, *J. Phys. Chem.* **1989**, 93, 490.
- [35] O. Demerdash, E. H. Yap, T. Head-Gordon, *Annu. Rev. Phys. Chem.* **2014**, 65, 149.
- [36] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, 14, 33, 27.
- [37] H. Li, A. D. Robertson, J. H. Jensen, *Proteins Struct. Funct. Bioinf.* **2005**, 61, 704.
- [38] J. P. Julien, A. Cupo, D. Sok, R. L. Stanfield, D. Lyumkis, M. C. Deller, P. J. Klasse, D. R. Burton, R. W. Sanders, J. P. Moore, A. B. Ward, I. A. Wilson, *Science* **2013**, 342, 1477.
- [39] S. H. Northrup, S. A. Allison, J. A. Mccammon, *J. Chem. Phys.* **1984**, 80, 1517.
- [40] J. D. Madura, J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, J. A. Mccammon, *Comput. Phys. Commun.* **1995**, 91, 57.
- [41] G. Schreiber, A. R. Fersht, *Nat. Struct. Biol.* **1996**, 3, 427.
- [42] R. R. Gabdouliline, R. C. Wade, *J. Mol. Biol.* **2001**, 306, 1139.
- [43] E. H. Yap, T. Head-Gordon, *J. Chem. Theory Comput.* **2010**, 6, 2214.
- [44] E. H. Yap, T. L. Head-Gordon, *J. Chem. Theory Comput.* **2013**, 9, 2481.

Received: 4 August 2016

Revised: 18 September 2016

Accepted: 1 October 2016

Published online on 2 November 2016