Adaptive Goal-Oriented Solver for the Linearized Poisson-Boltzmann Equation

Svetoslav Nakov,^{*,†} Ekaterina Sobakinskaya,^{*,†} Thomas Renger,[†] and Johannes Kraus[‡]

[†]*Institute for Theoretical Physics, Johannes Kepler University, Altenbergerstraße 69, 4040 Linz, Austria*

[‡]*Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany*

E-mail: svetoslav.nakov@jku.at; ekaterina.sobakinskaya@jku.at

Abstract

An adaptive finite element solver for the numerical calculation of the electrostatic coupling between molecules in a solvent environment is developed and tested. The new solver is based on a derivation of a new goal-oriented a posteriori error estimate for the electrostatic coupling. This estimate involves the consideration of the primal and adjoint problems for the electrostatic potential of the system, where the goal functional requires pointwise evaluations of the potential. A common practice to treat such functionals is to regularize them, for example, by averaging over balls or by mollification, and to use weak formulations in standard Sobolev spaces. However, this procedure changes the goal functional and may require the numerical evaluation of integrals of discontinuous functions. We overcome the conceptual shortcomings of this approach by using weak formulations involving nonstandard Sobolev spaces and deriving a representation of the error in the goal quantity which does not require averaging and directly exploits the original goal functional. The accuracy of this solver is evaluated by numerical experiments on a series of problems with analytically known solutions. In addition, the solver is used to calculate electrostatic couplings between two chromophores, linked to polyproline helices of different lengths. All the numerical experiments are repeated by using the well-known finite difference solvers MEAD and APBS, revealing the advantages of the present finite element solver.

1 INTRODUCTION

Almost all chemical processes in living organisms take place in aqueous solution. The solvation properties of the aqueous environment significantly contribute to the reaction mechanisms and kinetics, and influence the function of biomolecules in general. Dehydrated enzymes are reactivated when they are covered by at least one layer of water molecules and “dynamical” transitions of proteins are triggered by their strong coupling with the hydration water.^{1,2} The presence of water contributes also to protein folding by stabilizing secondary structure elements and their association.³ Mobility and rotational degrees of freedom of the water molecules are responsible for a strong shielding of electrostatic interactions between biomolecules that can influence

protein-ligand binding, DNA-protein recognition, etc.

Contemporary models of solvation account for the aqueous environment either explicitly, by inclusion of water molecules in atomic detail, or implicitly, by approximating the mean force exerted by the solvent on a biomolecule. The latter continuum solvation models, represent a solvated molecule at an atomic level of detail inside a molecule-shaped cavity. The electrostatic properties of the molecule are modelled by partial charges placed at the atom positions, embedded in a low dielectric region. The degrees of freedom of water molecules in the surrounding solvent are approximated by that of an isotropic dielectric continuum with high dielectric permittivity.

The electrostatic properties of such a model are described by the Poisson-Boltzmann equation (PBE), which also accounts for the presence of mobile ions in the solvent. The latter are introduced as a continuous Boltzmann-distributed charge density. Apart from the approximations mentioned above, the PBE assumes a linear response of the dielectric medium and neglects correlations between ions. Despite its relative simplicity, the PBE provides an effective way to study electrostatic effects in biomolecular systems and is widely used in different areas of research:^{4,5} biomolecular recognition, kinetics of molecular association, enzyme catalysis, etc. Analytical solutions of the PBE are available only for specific simple systems. For complex biomolecular surfaces and arbitrary distribution of charges, numerical methods have to be employed to solve the PBE. From a mathematical point of view the PBE represents a second-order semilinear elliptic partial differential equation with a measure-valued right hand side. The PBE imposes several challenges impacting the accuracy of numerical algorithms. These challenges include discontinuous dielectric coefficients at interfaces (for example, water-biomolecule), singularities as a result of the representation of the partial charges by delta functions, and the complicated geometry of molecular domains in three spatial dimensions. Overcoming each of these difficulties is the key for an accurate numerical solution.

Most of the common approaches for the numerical solution of the PBE fall in one of the following three categories: finite difference (FD), boundary element (BE) and finite element (FE) methods. The majority of the PBE solvers like DelPhi,⁶ GRASP,⁷ MEAD,⁸ UHBD,⁹ the PBEQ module

of CHARMM,¹⁰ and the PMG solver in the APBS software suite¹¹ are based on FD methods, utilizing regular 3D lattices, due to the ease of numerical implementation.

Finite difference schemes rely on solving the PBE in a lattice representation. The charges and dielectric permittivity are mapped onto a rectangular grid and a discrete approximation employing difference quotients is applied to the partial differential equation. As a result, a system of linear/nonlinear algebraic equations is build up, which is often solved by iterative methods, e.g., the Conjugate Gradient method.

To speed up the convergence and reduce the memory requirements, a focusing technique was developed in Refs. 6 and 12. This procedure involves solving the equation on a coarse grid, covering a large region, followed by a solution on a finer grid on a smaller region, with a boundary condition, interpolated from the values of the coarse grid solution. This approach was further improved in Ref. 13 by implementing it in parallel and then was applied to the electrostatics of microtubules and ribosomes. Due to the highly improved performance, focusing schemes are utilized in most of the FD-based solvers for the PBE in which regular cubic lattices are used.

As the underlying grid in regular 3D lattices is not adapted to the molecular surface, the accuracy of the solution is decreased. The correct assignment of the dielectric coefficient in points close to the solute-solvent interface is crucial for an optimal convergence of FD methods. One way to achieve it is the use of certain averaging procedures. More precisely, for each integer or half-integer grid point the values of the dielectric coefficient ϵ , mapped on the grid, are defined by averaging its actual values at the neighboring points, which depend on the modelling assumptions concerning the definition of the solute-solvent interface and the dielectric permittivity in the molecular and solvent regions.^{12,14} The inverse Debye length $\bar{k}$ is defined in a similar fashion. Due to the averaging, there is a region around the true interface in which these model parameters vary between their respective values on both sides of the interface. The thickness of this layer depends on the grid spacing h and this is where the discretization error of the FD method is typically much larger than that produced by the FE method, discussed in the end of this section.

Finite difference methods also exhibit difficulties caused by nonsmooth source terms, such as the delta distributions introduced to model point charges. A redistribution of the charges around their closest grid points is necessary in order to minimize the errors occurring in the electrostatic potential (ESP) at small distances from the charges (see Refs. 12,14 for different approaches to such a redistribution). However, redistribution techniques in general lead to so-called grid artifacts arising from the self interaction of the charges.^{15,16} The latter can be avoided by using Galerkin type FE approximations and combining them with proper splitting techniques for the full potential. By applying such decompositions of the full potential, the delta distributions describing the discrete charge density due to the point charges in the solute are transformed to a more regular distribution, which is interpreted as a certain charge density concentrated on the molecular surface. We note that these splitting techniques can also be used in conjunction with FD methods (see, e.g. Refs. 17,18). However, the surface charge density on the interface between the molecular region and the solution region remains hard to handle with FD methods.

Another disadvantage of FD methods is that the errors, caused by the specific way of handling interfaces and the partial charge distribution in the solute, strongly depend on the particular position and orientation of the molecular system relative to the FD grid. One approach to cope with this problem is to use finer grids, which naturally leads to a

larger number of grid points. Another approach is to use a rotational averaging scheme. The idea is to make several calculations with slightly different mutual orientation between the molecular system and the FD grid, and finally take an average (see Refs. 12,18).

A third approach, which could cope better with curved interfaces and regions where the solution is sharply varying (mainly around the partial charges and across interfaces), is to involve some kind of adaptivity in the FD method in conjunction with splitting techniques for the full potential. Finite difference methods based on adaptive Cartesian grids (ACG) have been developed for solving nonlinear equations in domains with curved boundaries.¹⁹⁻²¹ Here, the grid representation is based on quadrees in two dimensions and octrees in three dimensions. The ACG is recursively adapted by identifying mesh cells, which intersect the molecular surface, and test whether they satisfy one of the following criteria: finest grid-spacing is reached or intersected cell lies more than a prescribed distance from the nearest atomic charge site. The cell, which does not meet either of these criteria is uniformly subdivided into four (in 2D) or eight (in 3D) smaller cells. Using a coarser mesh away from the surface reduces the total number of grid points, which for ACG often is several orders of magnitudes less than that required in a conventional lattice grid. However, the underlying mesh still does not conform to the solute surface and thus an accurate evaluation of the ESP near the interface remains problematic.

This issue is even more pronounced in the presence of non-smooth interfaces which are quite common, e.g., when employing van der Waals' surfaces. In order to avoid this shortcoming, the technique proposed in Ref. 22 introduces additional mesh points at the Cartesian grid/surface intersection and uses a reconstruction of the ESP from the solution obtained in the exterior and interior domains. Note that, in this approach the continuity of the normal component of the displacement field is not enforced in a strong (pointwise) sense, which is a potential source of error. Moreover, the accuracy typically deteriorates when a charge approaches the surface and the mesh does not adequately resolve the rapid variations of the potential induced by such a charge. Recently, an FD method based on ACG was implemented in the finite difference solver CPB, particularly aiming at the approximation of the PBE for complex biomolecular structures.²³ This solver also utilizes a splitting technique for the full potential to eliminate the singular behavior at the charge sites.

The boundary element method (BEM) is another alternative discretization technique for partial differential equations. It was first adapted to the computation of the ESP of macromolecules in Ref. 24. Here, the PBE is reformulated into boundary integral equations for the boundary of the solute domain by employing Green's theorem. Therefore, only the two-dimensional molecular surface needs to be discretized. The resulting boundary integral equations are solved by using collocation, Galerkin, or least squares methods. The first of these approaches is the most commonly used by the PBE community since it is the simplest to apply in practice. Further developments of the BEM with the same focus of application can be found in Refs. 25 and 26 and deal mainly with algorithmic improvements. The BE has some advantages over other numerical methods like the FE and FD methods, which involve volume-domain discretization. Since only the molecular surface is discretized, the number of unknowns is greatly reduced. Moreover, boundary conditions at infinity are treated exactly as opposed to the FE and FD methods. In contrast to the FDM, the jump in the normal component of the electrostatic field across the molecular surface (interface) is explicitly accounted for, which results in a more accurate solution

near the solute surface.

The BEM also possesses certain disadvantages, mainly related to the requirement of an accurate surface mesh, the efficient evaluation of singular and hypersingular integrals and numerous other integral operations.²⁷ The conventional BEM requires the evaluation of mutual interactions between all elements. For a problem with N surface elements, the calculation of the potential in all elements then involves $O(N^2)$ interactions to be stored and $O(N^3)$ operations to be performed. Furthermore, the evaluation of each single integral, representing an interaction, is expensive. Thus, significant efforts have been made in developing fast algorithms for performing this task. In the presence of an ionic charge density of the solvent, BE methods usually neglect the ion exclusion (Stern) layer and thus only take into account a single surface separating the solute and the solvent.

One of the promising strategies to improve the efficiency of the BEM is to combine it with a fast multipole method (FMM).^{28,29} Further improvements can be achieved by using an adaptive FMM for accelerating the convolution type matrix-vector multiplications.³⁰ Solvers based on these concepts benefit from a more efficient use of the memory and improved load-balance between the local- and far-field calculations, resulting in a significant speed up of the computations.

The third approach to the numerical solution of the PBE is the finite-element (FE) method, which approximates the potential by a superposition of a set of basis functions. The FE method, which is employed in the present work, is the most successful for elliptic PDEs, since it combines geometrical flexibility and satisfactory convergence analysis with the ability to handle nonlinear problems involving interface jump conditions and nonsmooth source terms. Moreover, it enjoys a wide range of efficient iterative solvers for the resulting sparse linear systems, where the PBE is typically solved on unstructured finite element meshes. The latter in general provide higher geometric flexibility and variable mesh spacing that enable the algorithm to adapt the mesh better to the shape of the molecule so that no edges or elements intersect its surface. This adaptation allows for inherently more accurate estimates of quantities of interest on the surface, particularly of the electrostatic field. The finite element discretization results in a linear or nonlinear algebraic system of equations for the expansion coefficients in the linear combination of basis functions that represents the approximate potential. This system is then solved usually by means of iterative methods such as the Conjugate Gradient or the Newton-Raphson method.

One of the first applications of the FE method to continuum electrostatics of solubilized biomolecules appeared in Ref. 31. Since then, the popularity of this approach has increased synchronically with the increase of computational power and advancement of FE mesh generation techniques, see, e.g. Refs. 11,32–35. As representatives of the solvers, implementing the FE method for the numerical approximation of the PBE and the linearized PBE (LPBE), we mention APBS¹¹ (utilizing adaptive mesh refinement based on a residual type error indicator) and mFES,¹⁵ respectively.

However, application of FEM solvers for the PBE and the LPBE to reliable calculation of the electrostatic interaction between molecules becomes extremely difficult in the absence of appropriate error estimators and adaptive mesh refinement techniques. The two FEM solvers mentioned above face the same issues. More precisely, mFES does not implement error estimation and adaptive mesh refinement techniques, besides the fact that the initial mesh is somehow adapted to the shape and position of the solute. More precisely, the tetrahedral volume mesh inside the solute has approximately constant edge size related to the mesh size on the solute-solvent interface, whereas the mesh size out-

side the solute is gradually increasing up to the boundary of the computational domain. On the other hand, APBS utilizes adaptive mesh refinement based on a residual type error indicator. However, in APBS, the error is controlled in the so called energy norm, which is a global (integral) quantity, whereas for the evaluation of the electrostatic coupling between molecules the ESP needs to be evaluated only at particular points of interest. Therefore, adaptive mesh refinement based on energy norm error indicators attempts to enhance the global quality of the solution. This approach typically yields optimal rates of convergence with respect to the global energy norm, while the improvement in the potential at the points of interest happens at a much slower rate. Thus, an adaptive FE solver which is based on energy norm error indicators is inherently unreliable and ineffective for the computation of the electrostatic interaction between molecules.

In the present paper, we report an Adaptive Refinement Goal-Oriented Solver³⁶ (ARGOS) for the calculation of the electrostatic interaction between molecules. In order to have a consistent description of the electrostatic interaction, which is symmetric in nature, one needs to consider the linearized PBE with a homogeneous Dirichlet boundary condition as the primal problem governing the ESP. Besides the smallness of the ESP, this is why ARGOS, in its present implementation, is based on the linearized PBE. The most notable feature of this solver is the goal-oriented a posteriori error control, which allows adaptive refinement of the mesh specifically aiming at the reduction of the error in the electrostatic coupling, herewith significantly improving the accuracy of the numerical solution and the computational efficiency. We compare the accuracy of computed interaction energies obtained via this new approach with those from FD solvers (MEAD and APBS) both for practically relevant problems and for model problems with known analytical solutions.

The structure of the paper is as follows. In METHODS, the theory of a goal-oriented error estimator for the electrostatic coupling between molecules and its algorithmic realization are discussed. In RESULTS, we start with an application of our new ARGOS to three test cases with analytically known solutions and evaluate the accuracy of ARGOS in comparison to the FD solvers MEAD and APBS. We proceed with the application of the different solvers to the computation of the interaction energy between two chromophores, that are attached to a polypeptide. Besides reevaluating the coupling between the transition densities of the chromophores,³⁷ we study the interaction between their ground state charge densities. The latter holds the key for an implicit description of the aqueous solvent that has been described explicitly in our earlier molecular dynamics (MD) simulations.³⁷ The superior performance of ARGOS is used to evaluate the accuracy and numerical efficiency of the FD solvers MEAD and APBS.

2 METHODS

2.1 Theoretical background and implementation of the solver

2.1.1 The linearized Poisson-Boltzmann equation

The solution of the Poisson-Boltzmann equation (PBE) gives the ESP φ of a discrete charge distribution that is placed in a dielectric continuum containing a continuous ionic charge distribution. In the special case of only two univalent ion species in the solvent with the same concen-

tration but opposite charges, the PBE is given by

$$-\nabla \cdot (\epsilon \nabla \varphi) + \frac{k_B T}{e_0} \bar{k}^2 \sinh\left(\frac{e_0 \varphi}{k_B T}\right) = 4\pi e_0 \sum_{i=1}^N z_i \delta(x - x_i) \quad (1)$$

where k_B is the Boltzmann constant, T the temperature, e_0 the elementary charge, N the number of charges in the molecule, z_i the valency of the i -th partial charge, and x_i its coordinate in the molecule. The dielectric coefficient ϵ is piecewise constant and defined by

$$\epsilon(x) = \begin{cases} \epsilon_m, & x \in \Omega_m, \\ \epsilon_s, & x \in \Omega_s, \end{cases} \quad (2)$$

where Ω_m is the molecular domain and Ω_s is the solvent domain composed of Ω_{ions} and the ionic exclusion layer (IEL) Ω_{IEL} . The latter is a layer around the molecules which no ions can penetrate. It is defined as the difference between the union of the van der Waals spheres of the atoms, where the atomic radii have been increased by a counterion radius R_{ion} , and the molecular region defined by the Connolly surface formed by the contact points of the van der Waals surface and a (solvent) probe sphere, rolling over it (see Refs. 54–58). The probe radius is set to 1 Å in all calculations. It is known that the electrostatic interaction energy depends on the definition of the solute surface (see, for example, Ref. 53), but the study of this problem is beyond the scope of this paper. For a system of N_{mol} molecules, Ω_m consists of the domains $\Omega_{m_1}, \dots, \Omega_{m_{N_{mol}}}$, where Ω_{m_i} corresponds to the domain of the i -th molecule (see Figure 1). The coefficient $\bar{k}^2$ is also piecewise constant and given by

$$\bar{k}^2(x) = \begin{cases} 0, & x \in \Omega_m \cup \Omega_{IEL}, \\ \frac{8\pi N_A e_0^2 I_s}{1000 k_B T}, & x \in \Omega_{ions}, \end{cases} \quad (3)$$

where the ionic strength I_s is defined as

$$I_s = \frac{1}{2} \sum_{i=1}^2 c_i (-1)^2$$

with the molar concentration of each ion type $c_1 = c_2$ and N_A denoting Avogadro's number. Assuming the potential φ to be small ($\varphi \ll \frac{k_B T}{e_0}$), the PBE (1) can be linearized. The resulting equation

$$-\nabla \cdot (\epsilon \nabla \varphi) + \bar{k}^2 \varphi = 4\pi e_0 \sum_{i=1}^N z_i \delta(x - x_i) \quad (4)$$

is referred to as the linearized PBE (LPBE) and will be used in this work. For the special case $I_s = 0$, equation (4) becomes the Poisson equation:

$$-\nabla \cdot (\epsilon \nabla \varphi) = 4\pi e_0 \sum_{i=1}^N z_i \delta(x - x_i). \quad (5)$$

Equations (1), (4), and (5) include a natural interface Γ between the molecular and the solvent domains due to different dielectric coefficients $\epsilon = \epsilon(x)$ (see (2)). The potential φ and the normal component of the displacement field $\epsilon \nabla \varphi$ must

be continuous across the interface, i.e.,

$$[\varphi]_\Gamma = 0, \quad (6)$$

$$[\epsilon \nabla \varphi \cdot n]_\Gamma = 0, \quad (7)$$

where $[\cdot]_\Gamma$ denotes the jump of the enclosed quantity across Γ . In order to find a unique solution of (1) and (4), a Dirichlet boundary condition for φ must be imposed. The usual approach in physics prescribes a vanishing potential at infinite distance from the boundary of Ω_m . In practical applications, one works in a bounded computational domain Ω and imposes the boundary condition

$$\varphi = g \quad \text{on } \partial\Omega, \quad (8)$$

where g is usually close to zero and calculated accurately enough by solving a simpler problem, possibly with a known analytical solution.

Figure 1: Schematic representation of the different regions in the presence of ionic charge density in the solvent.

2.1.2 Integral formulation of the problem

The right-hand side of (4) is a functional, consisting of a linear combination of delta functions, while the left hand side defines a proper function everywhere except at the interface Γ , where ϵ is discontinuous and $\nabla \cdot (\epsilon \nabla \varphi)$ is not well defined. A natural way to interpret (4) is to rewrite it in its integral (weak) formulation, where the left-hand side and the right-hand side are comparable objects with well-defined meaning. To obtain the weak formulation, we multiply both sides of (4) by a function v equal to zero on $\partial\Omega$, also called a test function, and then integrate the equation over the domain Ω :

$$\int_{\Omega} \nabla \cdot (\epsilon \nabla \varphi) v dx + \int_{\Omega} \bar{k}^2 \varphi v dx = 4\pi e_0 \sum_{i=1}^N z_i v(x_i). \quad (9)$$

On the left-hand side we apply the divergence theorem for integration by parts and obtain (for details see Ref. 38)

$$\int_{\Omega} \epsilon \nabla \varphi \cdot \nabla v dx + \int_{\Omega} \bar{k}^2 \varphi v dx = 4\pi e_0 \sum_{i=1}^N z_i v(x_i) \quad (10)$$

for all test functions v .

It can be rigorously shown that under certain conditions on the regularity of the interface Γ there exists a unique potential φ that satisfies the integral formulation (10) for any

test function v in a certain function space.^{39,40} By directly discretizing (10), one obtains an approximation φ_h of the exact potential φ .

Another widespread approach, from which both the numerical treatment and analysis of the LPBE (also the PBE) benefits, involves splitting the potential φ into two parts G and u , i.e., $\varphi = G + u$ (see, e.g. Refs. 14,18,41,42), where

$$G = \sum_{i=1}^N \frac{z_i e_0}{\epsilon_m |x - x_i|} \quad (11)$$

is the singular Coulomb potential of the point charges in a uniform dielectric medium with $\epsilon = \epsilon_m$ without any molecular cavities and u is the reaction field potential that takes into account the polarization of the solvent. Thus, one has to find only the reaction field potential u_h instead of the potential φ that has singularities at the positions of the charges. This approach is particularly beneficial when one needs the reaction field potential only or when calculating the solvation energy, that is, the free energy difference between the solute in vacuum and in a solvent environment

$$\begin{aligned} \Delta G^{\text{vac} \rightarrow \text{solv}} &= G^{\text{vac}}(\epsilon_{\text{vac}}, \epsilon_m) - G^{\text{solv}}(\epsilon_s, \epsilon_m) \\ &= \frac{1}{2} \sum_{i=1}^N q_i (u_{\text{vac}}(\vec{r}_i) - u_{\text{solv}}(\vec{r}_i)), \end{aligned} \quad (12)$$

where $q_i = z_i e_0$, u_{vac} and u_{solv} are the reaction field potentials for the molecule in vacuum and the solvent, respectively. It should be noted that by separating the singular potential G from the full potential φ one avoids the grid artifacts arising from the unphysical self interaction of charges (see Refs 15 and 16).

2.1.3 Goal-oriented a posteriori error estimates

The adaptive finite element method with a posteriori error estimation is designed for achieving high accuracy of a solution at minimum computational costs. Finding the mesh, providing the desired properties, is a nontrivial task, since the solution is unknown. If one is interested in the global quality of the solution over the whole computational domain, the desired mesh is obtained under the condition of error equidistribution, that is, the discretization errors should be approximately equal on each computational element. The errors are estimated by error indicators that are derived based on a particular a posteriori error estimation technique.

In many applications one is interested not in the solution itself, but rather in a specific quantity that depends linearly on the solution. Thus, the target quantity is expressed in terms of a linear functional, called a goal functional. In this case, the error estimations are provided only for the goal quantity. The resulting a posteriori estimates are used in the feedback process for mesh refinement in the computational domains, where the error is high, thus, improving the accuracy of the target quantity.

To outline the method, let us consider a general case of a linear differential equation with a solution w (see, e.g. Ref. 43):

$$Aw = f, \quad (13)$$

where A is a linear differential operator and f is a generalized force, exerted by the environment. Problem (13) is called primal. Let $\mathcal{L}(w)$ be a quantity of interest obtained from the solution w by applying a goal functional $\mathcal{L}$ and w_h an approximation of w that depends on a discretization

parameter h . For the purpose of this introduction we will assume that all involved functions are smooth. The strategy involves controlling the error

$$\mathcal{L}(e) = \mathcal{L}(w) - \mathcal{L}(w_h) \text{ with } e := w - w_h \quad (14)$$

in terms of the local residuals $\rho_K(w_h)$ computable on each mesh cell K :

$$\rho_K(w_h) = (f - Aw_h)_K. \quad (15)$$

The error equation

$$A(w - w_h) = \rho(w_h) \quad (16)$$

demonstrates that the cell residue $\rho_K(w_h)$ influences the error $e_{K'} = (w - w_h)_{K'}$ in another cell K' through the Green's function of the continuous problem (13). In practical applications, usually it is impossible to reveal the contribution of the local residuals ρ_K to the error $\mathcal{L}(e)$ in the goal quantity, and thus it has to be determined by computation. The idea is to consider the adjoint (or dual) problem

$$A^* z = j, \quad (17)$$

where A^* is the adjoint operator defined via $\langle Au, v \rangle = \langle A^* v, u \rangle$ for all functions u and v , where $\langle \cdot, \cdot \rangle$ is the standard integral inner product, and j is the density function associated with $\mathcal{L}$. For example, if one is interested in the average value of a solution w , the goal functional acting on w can be represented by

$$\mathcal{L}(w) = \int_{\Omega} j w dx \quad (18)$$

with $j(x) = 1/|\Omega|$, where $|\Omega|$ denotes the volume of Ω . Taking into account the definition of the adjoint operator, (17), and (14), the error in the goal quantity is obtained as

$$\begin{aligned} \mathcal{L}(e) &= \langle A^* z, w \rangle - \langle A^* z, w_h \rangle \\ &= \langle Aw, z \rangle - \langle Aw_h, z \rangle = \langle f - Aw_h, z \rangle \\ &= \langle \rho(w_h), z \rangle, \end{aligned} \quad (19)$$

where z is unknown and has to be substituted with an approximate solution. By analogy with equation (16) for the error $w - w_h$, we see that the error $\mathcal{L}(e)$ of the goal quantity is determined by the local cell values of the residual $\rho_K(w_h)$ through the solution z of the adjoint problem (17). In fact, this kind of goal error representation can be obtained with much less regularity, where f in (13) is exchanged with $\mathcal{F}$ in a space of distributions. Likewise, the approximation w_h and also $\mathcal{L}$ can be less regular. In this case, one obtains

$$\mathcal{L}(e) \approx \langle \rho(w_h), s_h(z) \rangle, \quad (20)$$

where $s_h(z)$ is a sensitivity factor obtained from the approximate solution of the adjoint problem (17). Thus, the adjoint solution demonstrates how the information from the residual $\rho(w_h)$ in the primal problem is propagated to the error $\mathcal{L}(e)$ of the quantity of interest. This approach to a posteriori error estimation with respect to a quantity of interest is called the dual weighted residual (DWR) method since the residual $\rho(w_h)$ is weighted with the quantity $s_h(z)$ obtained by a solution of the adjoint problem (17).

The target quantity in our case is the electrostatic interaction energy E between two molecules, M_1 and M_2 . It can

be expressed through the potential φ_1 , produced by M_1 at the positions $x_{i,2}$ of the charges $q_{i,2}$ of molecule M_2 :

$$E_{1-2} = \sum_{i=1}^{N_2} \varphi_1(x_{i,2}) q_{i,2}, \quad (21)$$

where $q_{i,2} = z_{i,2}e_0$ and N_2 is the number of charges in M_2 . The energy can also be expressed by the potential φ_2 of M_2 at the position of the charges $q_{i,1}$ of M_1 , i.e.,

$$E_{2-1} = \sum_{i=1}^{N_1} \varphi_2(x_{i,1}) q_{i,1}, \quad (22)$$

where $q_{i,1} = z_{i,1}e_0$ and N_1 is the number of charges in M_1 . Here, we assume φ_1 and φ_2 to be the solutions of the LPBE (4) together with the interface conditions (6) and (7), and a homogeneous Dirichlet boundary condition $g = 0$ in (8). Note that both, the linearity of the LPBE and the homogeneous Dirichlet boundary condition on $\partial\Omega$, are necessary in order to reflect the symmetry of the interaction, i.e., $E_{2-1} = E_{1-2}$ ($= E$) (see Proposition 5.12 in Ref. 38 for a proof of this fact).

In what follows, without loss of generality, we define the primal problem to be

$$-\nabla \cdot (\epsilon \nabla \varphi_2) + \bar{k}^2 \varphi_2 = 4\pi \sum_{i=1}^{N_2} q_{i,2} \delta(x - x_{i,2}). \quad (23)$$

It describes the potential created by the charges of molecule M_2 . For convenience, we denote the right-hand side of (23) as the functional

$$\mathcal{F} = 4\pi \sum_{i=1}^{N_2} q_{i,2} \delta(x - x_{i,2}). \quad (24)$$

Applying (10), we obtain the integral formulation of the primal problem (23), i.e.,

$$\int_{\Omega} \epsilon \nabla \varphi_2 \cdot \nabla v dx + \int_{\Omega} \bar{k}^2 \varphi_2 v dx = \mathcal{F}(v) \quad (25)$$

for all test functions v .

It is easy to see that the goal functional related to our target quantity (22) is given by

$$\mathcal{L} = \sum_{i=1}^{N_1} q_{i,1} \delta(x - x_{i,1}). \quad (26)$$

Therefore, the interaction energy (22) can be written as

$$E_{2-1} = \mathcal{L}(\varphi_2). \quad (27)$$

For further analysis, it is necessary to express the adjoint problem in an integral formulation similar to (25). Recalling the definition (17) and taking into account that in our case $A^* = A$ we obtain

$$\int_{\Omega} \epsilon \nabla v \cdot \nabla z dx + \int_{\Omega} \bar{k}^2 v z dx = \mathcal{L}(v) \quad (28)$$

for all test functions v ,

where z is the solution of (28) and is defined to be equal to zero on the boundary $\partial\Omega$. Here, we would like to emphasize the difference between the primal and adjoint problems: the solution φ_2 of the primal problem is generated by the charges of the molecule M_2 , whereas the solution z of the adjoint problem is determined by the charges of M_1 (see (24) and (26)). Now, by taking into account the definitions (24) and (26) of $\mathcal{F}$ and $\mathcal{L}$, respectively, and of the potential z , it is clear that the interaction energy E_{1-2} ($= E_{2-1}$) can also be calculated as

$$E_{1-2} = \mathcal{F}(z). \quad (29)$$

The optimal adaptive algorithm balances the error distribution of the calculated potential over the domain Ω in such a way that for a given number of degrees of freedom (DOFs) the error in the goal quantity

$$\mathcal{L}(e) = \mathcal{L}(\varphi_2) - \mathcal{L}(\varphi_{2,h}) = \sum_{i=1}^{N_1} q_{i,2} e_i, \quad (30)$$

with $e := \varphi_2 - \varphi_{2,h}$ and $e_i := \varphi_2(x_{i,1}) - \varphi_{2,h}(x_{i,1})$ is minimal. In general, this minimization problem cannot be solved exactly, but adaptive algorithms aim at approximating its solution well. Hereinafter, the index h will indicate that the physical quantity to which it refers to is computed approximately by a numerical algorithm on a mesh with mesh parameter h .

A difficulty in the problem under consideration is that the goal functional defining the quantity of interest, i.e., the electrostatic interaction, requires point evaluations. Hence, it is not representable in the form (18) through a density function j , but defined in terms of the delta functions $\delta(x - x_{i,1})$, cf. (26). A standard approach in such a situation employs regularization of the goal functional by replacing the point evaluations with averages over small balls $B(x_{i,1}, \rho)$, where ρ is the radius around the points of interest $x_{i,1}$. The application of this approach, however, changes the goal functional when the averaged functions are not harmonic and requires the evaluation of integrals of discontinuous functions over balls. For achieving high enough accuracy, the mesh has to be a priori adapted around the points of interest and special integration rules have to be used. In order to avoid these conceptual shortcomings, here, we derive nontrivial representations of the error in the goal quantity which do not require averaging and directly exploit the original goal functional.

Error estimator

The following approach does not involve any splitting of the full potential and is based on a direct discretization of the problem defining φ_2 . Hence, the primal and adjoint problems are given by the expressions (25) and (28), respectively. The goal quantity is defined in (27), and the approximate interaction energy is simply

$$E_{2-1}^{Prim} = \mathcal{L}(\varphi_{2,h}), \quad (31)$$

where $\varphi_{2,h}$ is the solution of the discrete primal problem

$$\int_{\Omega} \epsilon \nabla \varphi_{2,h} \cdot \nabla v dx + \int_{\Omega} \bar{k}^2 \varphi_{2,h} v dx = \mathcal{F}(v) \quad (32)$$

for all test functions $v \in V_h^1$.

Here, $\varphi_{2,h}$ is a linear combination of continuous piecewise linear functions from the finite element space V_h^1 defined on the mesh $\mathcal{T}_h$. By taking into account (27), (31), and the

fact that the goal functional $\mathcal{L}$ is linear, the error in the interaction energy is obtained as

$$E_{2-1} - E_{2-1}^{Prim} = \mathcal{L}(\varphi_2 - \varphi_{2,h}). \quad (33)$$

Further, to obtain an error estimator in terms of the solution to the adjoint problem, we will use the following equalities (for a proof see Proposition 5.12 in Ref. 38):

$$\begin{aligned} \int_{\Omega} \epsilon \nabla(\varphi_2 - \varphi_{2,h}) \cdot \nabla z dx + \int_{\Omega} \bar{k}^2 (\varphi_2 - \varphi_{2,h}) z dx \\ = \mathcal{L}(\varphi_2 - \varphi_{2,h}), \end{aligned} \quad (34)$$

$$\int_{\Omega} \epsilon \nabla \varphi_2 \cdot \nabla z dx + \int_{\Omega} \bar{k}^2 \varphi_2 z dx = \mathcal{F}(z). \quad (35)$$

Equality (34) essentially means that one can test (28) with the function $\varphi_2 - \varphi_{2,h}$ and equality (35) shows that (25) can be tested with the adjoint solution z . Successive application of (34) and (35), similarly to the general approach presented by (19), yields

$$\begin{aligned} \mathcal{L}(\varphi_2 - \varphi_{2,h}) &= \mathcal{F}(z) - \int_{\Omega} \epsilon \nabla \varphi_{2,h} \cdot \nabla z dx \\ &- \int_{\Omega} \bar{k}^2 \varphi_{2,h} z dx =: \mathcal{B}(\varphi_{2,h}, z). \end{aligned} \quad (36)$$

Evaluation of the error in (36) requires the solution of the adjoint problem. In practice, one finds a good approximation of z in a richer space than the one which $\varphi_{2,h}$ belongs to. In particular, we use an approximation $z_h^{(2)}$ from the space V_h^2 consisting of continuous piecewise quadratic functions defined on the same mesh $\mathcal{T}_h$. Thus, the solution $z_h^{(2)}$, which is equal to zero on $\partial\Omega$, is obtained from

$$\begin{aligned} \int_{\Omega} \epsilon \nabla v \cdot \nabla z_h^{(2)} dx + \int_{\Omega} \bar{k}^2 v z_h^{(2)} dx &= \mathcal{L}(v) \\ \text{for all test functions } v \in V_h^2. \end{aligned} \quad (37)$$

In this case, by using (32) (for similar techniques of error estimator localization see, e.g. Refs. 44 and 43) we obtain

$$\begin{aligned} \mathcal{B}(\varphi_{2,h}, z_h^{(2)}) &= \mathcal{F}(z_h^{(2)} - I_h z_h^{(2)}) - \int_{\Omega} \epsilon \nabla \varphi_{2,h} \cdot \nabla (z_h^{(2)} - I_h z_h^{(2)}) dx \\ &- \int_{\Omega} \bar{k}^2 \varphi_{2,h} (z_h^{(2)} - I_h z_h^{(2)}) dx \\ &= \sum_{i=1}^{N_V} \eta_i^{PU}(\varphi_{2,h}, z_h^{(2)}), \end{aligned} \quad (38)$$

where I_h is the nodal interpolation operator in V_h^1 (and thus $I_h z_h^{(2)} \in V_h^1$ can be taken as a test function in (32)), N_V is the number of nodes (vertices) in the mesh $\mathcal{T}_h$, and

$$\begin{aligned} \eta_i^{PU}(\varphi_{2,h}, z_h^{(2)}) &= \langle \mathcal{F}, (z_h^{(2)} - I_h z_h^{(2)}) \psi_i \rangle \\ &- \left(\epsilon \nabla \varphi_{2,h} \cdot \nabla ((z_h^{(2)} - I_h z_h^{(2)}) \psi_i) \right) \\ &- \left(\bar{k}^2 \varphi_{2,h} (z_h^{(2)} - I_h z_h^{(2)}) \psi_i \right). \end{aligned} \quad (39)$$

In (39), $\{\psi_i\}_{i=1}^{N_V}$ are the nodal P_1 basis functions equal to 1 at the i -th node and zero elsewhere (for details, see Section 2.1.6). The functions $\{\psi_i\}_{i=1}^{N_V}$ form a partition of unity, i.e., their sum $\sum_{i=1}^{N_V} \psi_i(x) = 1$ for each $x \in \Omega$, which justifies the second equality in (38). The nodal interpolant $I_h f$ of a continuous function f is a finite element function in the space V_h^1 , equal to f at all nodes V_i of the mesh $\mathcal{T}_h$.

The quantities $\eta_i^{PU}(\varphi_{2,h}, z_h^{(2)})$, $i = 1, \dots, N_V$ are nodewise error indicators that are used to drive the adaptive mesh refinement. Note, that the term $\langle \mathcal{F}, (z_h^{(2)} - I_h z_h^{(2)}) \psi_i \rangle$ in the definition of η_i^{PU} can be nonzero only if the support of ψ_i contains a partial charge from molecule M_2 (recall that the charges of M_2 define the primal problem). Finally, for the error estimate we have

$$\mathcal{L}(\varphi_2 - \varphi_{2,h}) = \mathcal{B}(\varphi_{2,h}, z_h^{(2)}) + \mathcal{E}(\varphi_2, \varphi_{2,h}, z, z_h^{(2)}), \quad (40)$$

where $\mathcal{E}(\varphi_2, \varphi_{2,h}, z, z_h^{(2)})$ is a higher-order term that can be neglected. The exact expression for $\mathcal{E}(\varphi_2, \varphi_{2,h}, z, z_h^{(2)})$ can be deduced from (40) by using the expressions (36) and (38) for $\mathcal{L}(\varphi_2 - \varphi_{2,h})$ and $\mathcal{B}(\varphi_{2,h}, z_h^{(2)})$, respectively. To see why $\mathcal{E}(\varphi_2, \varphi_{2,h}, z, z_h^{(2)})$ is a higher-order term with respect to the mesh parameter h we refer to Ref. 38. For this reason, in practical applications we will employ the approximate equality

$$E_{2-1} - E_{2-1}^{Prim} \approx \mathcal{B}(\varphi_{2,h}, z_h^{(2)}) \quad (41)$$

and use the elementwise localization (38) of $\mathcal{B}(\varphi_{2,h}, z_h^{(2)})$ as an error indicator for our adaptive algorithm.

In our calculation of interaction energies, we will use the approximate solution $z_h^{(2)}$ of the adjoint problem since it is more accurate than $\varphi_{2,h}$. By first taking into account (32) together with the fact that $I_h z_h^{(2)} \in V_h^1$, and then noting that $\varphi_{2,h} \in V_h^1 \subset V_h^2$, from (38), we find that

$$\mathcal{B}(\varphi_{2,h}, z_h^{(2)}) = \mathcal{F}(z_h^{(2)}) - \mathcal{L}(\varphi_{2,h}). \quad (42)$$

Therefore, from (41), the interaction energy can be expressed as

$$E_{2-1} \approx E_{2-1}^{Prim} + \left(\mathcal{F}(z_h^{(2)}) - \mathcal{L}(\varphi_{2,h}) \right) = \mathcal{F}(z_h^{(2)}). \quad (43)$$

Now, from (42) and (43), we obtain

$$E_{1-2}^{Adj} := \mathcal{F}(z_h^{(2)}) = E_{2-1}^{Prim} + \mathcal{B}(\varphi_{2,h}, z_h^{(2)}). \quad (44)$$

This expression shows that the interaction E_{1-2}^{Adj} computed by means of the more accurate $z_h^{(2)}$ is equal to the interaction E_{2-1}^{Prim} corrected with the approximately evaluated error estimator $\mathcal{B}(\varphi_{2,h}, z_h^{(2)})$. The quantity E_{1-2}^{Adj} is a more accurate approximation of the electrostatic interaction E_{1-2} (see (29)), which will be used in our applications.

We will refer to the corresponding solver as ARGOS and the solution obtained by ARGOS will be denoted as $E^{ARGOS} := E_{1-2}^{Adj}$. If, for the definition of the primal prob-

lem (23), we use the charges of molecule M_1 instead, then the definition of the adjoint problem (28) will involve the charges of molecule M_2 . In this case, the numerical approximations of the interaction energies E_{1-2} and E_{2-1} obtained by ARGOS are denoted as E_{1-2}^{Prim} and E_{2-1}^{Adj} , respectively. The reasoning here is that this time the energy E_{1-2} is computed by using the solution of the primal problem, whereas E_{2-1} is obtained from the more accurate solution of the adjoint problem.

We note that, a similar approach to the estimation of the error in the goal quantity E can be executed in conjunction with the splitting of the full potential $\varphi_2 = G_2 + u_2$, where G_2 is the Coulomb potential of the charges of M_2 in a uniform dielectric medium with $\epsilon = \epsilon_m$ and u_2 is the corresponding reaction field potential. An adaptive algorithm based on goal oriented error estimation together with a splitting of the potential can be very reliable for the calculation of the solvation energy difference (12) and will be included in future releases of ARGOS.

2.1.4 Convergence order

It can be shown that, provided φ_2 and z are regular solutions, and under uniform mesh refinement in 3D, the error estimates satisfy

$$\left| E_{2-1} - E_{2-1}^{Prim} \right| = |\mathcal{B}(\varphi_{2,h}, z)| \leq C_1 h^2 \approx \frac{C_1}{(\#DOFs)^{2/3}}, \quad (45)$$

$$\left| E_{2-1} - E_{1-2}^{Adj} \right| = |\mathcal{E}(\varphi_2, \varphi_{2,h}, z, z_h^{(2)})| \leq C_2 h^3 \approx \frac{C_2}{\#DOFs}, \quad (46)$$

where $C_1, C_2 > 0$ are constants independent of the maximum mesh size h .³⁸ Here, we introduced $\#DOFs$ as the number of degrees of freedom since it is more appropriate to measure the approximation error in terms of $\#DOFs$ than in h under adaptive mesh refinement. From this expression, it is clear that the interaction energy obtained with the adjoint solution z converges faster to the exact interaction energy than the one with the primal solution. We will demonstrate these convergence rates in the RESULTS section.

2.1.5 Adaptive algorithm

In general, adaptive algorithms can be represented by the following flow chart (Figure 2):

Figure 2: Schematic representation of adaptive refinement as explained in the text.

Instead of the mesh parameter h , we will use a subindex $\ell = 1, 2, \dots$ to denote the refinement level and we recall that $\sum_{i=1}^{N_V} \eta_i^{PU} = \mathcal{B}(\varphi_{2,h}, z_\ell^{(2)})$. The adaptive finite element solver based on the error estimator described above can be summarized as follows.

Solve

The solution $\varphi_{2,\ell}$ from the space V_ℓ^1 , defined on the current mesh $\mathcal{T}_\ell$, is found by solving (32).

Estimate

Next, we find the solution $z_\ell^{(2)}$ of the adjoint problem by solving the finite dimensional problem (37). From $z_\ell^{(2)}$, we compute the nodewise error indicators $\eta_i^{PU}(\varphi_{2,\ell}, z_\ell^{(2)})$.

Mark

Based on the values of η_i^{PU} , a set of node patches $\mathcal{N}_\ell$ is marked for refinement. This set can be chosen through different marking strategies (see the SI).

Refine

Having performed a refinement based on the marked patches $\mathcal{N}_\ell$, a new mesh $\mathcal{T}_{\ell+1}$ is obtained.

The detailed algorithm of ARGOS reads:

ARGOS

- 1: $\ell \leftarrow 0$, initialize $\mathcal{T}_0, V_0^1, V_0^2$
 - 2: **repeat**
 - 3: $\varphi_{2,\ell} \leftarrow \text{Solve}^{Prim}(\mathcal{T}_\ell, V_\ell^1)$
 - 4: $z_\ell^{(2)} \leftarrow \text{Solve}^{adj}(\mathcal{T}_\ell, V_\ell^2)$
 - 5: $\{\eta_i\}_{i=1}^{PU} \leftarrow \text{Estimate}(\varphi_{2,\ell}, z_\ell^{(2)}, \mathcal{T}_\ell)$
 - 6: $\mathcal{N}_\ell \leftarrow \text{Mark}(\{\eta_i\}_{i=1}^{N_V})$
 - 7: $\mathcal{T}_{\ell+1} \leftarrow \text{Refine}(\mathcal{T}_\ell, \mathcal{N}_\ell)$
 - 8: $\ell \leftarrow \ell + 1$
 - 9: **until** convergence or maximum number of refinement levels reached
-

2.1.6 Finite element method

The finite element (FE) method aims at approximating the exact solution of a differential equation by a piecewise polynomial function. For this purpose, the computational domain is split into individual small (finite) elements, such as triangles in 2D or tetrahedrons in 3D. The construction is such that the FE space used for the approximation is finite dimensional with a basis consisting of locally (on patches of elements) supported piecewise polynomials. The latter have to satisfy certain interpolation conditions, e.g. continuity across element interfaces, as used in the adaptive algorithms presented in this paper.

In case of the lowest approximation order, i.e., piecewise linear polynomials, each basis function is associated with a node in the mesh, that is, a vertex of an element. The approximate solution is represented in the chosen basis with unknown expansion coefficients, e.g. the nodal values of the approximate solution. Based on a variational formulation of the differential equation, taking into account boundary conditions, a system of algebraic equations is derived for these expansion coefficients.

In our case, the FE space V_h^1 for the primal problem is based on the Lagrange P_1 element and consists of continuous piecewise linear functions. The FE space V_h^2 for the adjoint problem is based on the Lagrange P_2 element and consists of continuous piecewise quadratic functions. In the case of V_h^1 , there is one basis function ψ_i for each vertex (node) V_i of the mesh $\mathcal{T}_h$, which is equal to 1 at V_i and 0 at all other vertices, see Figure 3. Here, we will outline the construction of the system of linear algebraic equations arising from (32). The system for the adjoint problem can be constructed in a similar way. Recall, that equation (32) has to be satisfied for all test functions $v \in V_h^1$. Equivalently, this means that the equation must hold for all basis functions ψ_i , $i = 1, \dots, N_\Omega$ that are zero on the boundary $\partial\Omega$. In this notation, the number of nodes N_V in the mesh $\mathcal{T}_h$ equals the number of degrees of freedom $N_{DOFs}^{Prim} (= N_\Omega)$ plus the number of nodes $N_{\partial\Omega}$ on the boundary $\partial\Omega$. Hence, in the primal problem, we have $N_{DOFs}^{Prim} = N_\Omega$ and we obtain N_{DOFs}^{Prim} equations for

$\varphi_{2,h}$:

$$\int_{\Omega} \epsilon \nabla \varphi_{2,h} \cdot \nabla \psi_i dx + \int_{\Omega} \bar{k}^2 \varphi_{2,h} \psi_i dx = \mathcal{F}(\psi_i), \quad (47)$$

for $i = 1, \dots, N_{DOFs}^{Prim}$.

Moreover, $\varphi_{2,h}$ belongs to V_h^1 and $\varphi_{2,h} = 0$ on $\partial\Omega$. Therefore, $\varphi_{2,h}$ can be represented as a linear combination of the basis functions ψ_j which are equal to zero on the boundary $\partial\Omega$ with unknown expansion coefficients φ_2^j , i.e.,

$$\varphi_{2,h}(x) = \sum_{j=1}^{N_{DOFs}^{Prim}} \varphi_2^j \psi_j(x). \quad (48)$$

The goal is to find the vector of unknowns $\Phi_2 = (\varphi_2^1, \dots, \varphi_2^{N_{DOFs}^{Prim}})^T$ corresponding to the interior vertices.

For this purpose, we substitute expression (48) for $\varphi_{2,h}$ in (47) and use the linearity of the integral, gradient operator, and right-hand side to obtain the following system of linear equations

$$\sum_{j=1}^{N_{DOFs}^{Prim}} \left(\int_{\Omega} \epsilon \nabla \psi_j \cdot \nabla \psi_i dx \right) \varphi_2^j + \sum_{j=1}^{N_{DOFs}^{Prim}} \left(\int_{\Omega} \bar{k}^2 \psi_j \psi_i dx \right) \varphi_2^j = \mathcal{F}(\psi_i), \text{ for } i = 1, \dots, N_{DOFs}^{Prim}.$$

In matrix notation, the above equations read

$$\mathbf{A} \Phi_2 = \mathbf{F}_2, \quad (49)$$

where $\mathbf{A}$ is the symmetric positive definite matrix with entries

$$a_{i,j} = \int_{\Omega} \epsilon \nabla \psi_j \cdot \nabla \psi_i dx + \int_{\Omega} \bar{k}^2 \psi_j \psi_i dx, \quad i, j = 1, \dots, N_{DOFs}^{Prim}$$

and

$$\mathbf{F}_2 = \left(\mathcal{F}(\psi_1), \dots, \mathcal{F}(\psi_{N_{DOFs}^{Prim}}) \right)^T.$$

The system (49) is then solved either by a direct or by an iterative method. In our numerical tests we have used the conjugate gradient method with a sufficiently high accuracy in order to make the error due to the iterative solution of the linear system subordinate to the overall error. In this way, it is possible to study the effect of the adaptive mesh refinement in a reliable way.

Figure 3: Interior P_1 basis function in 2D.

Our computations are performed with two levels of geometric approximation of the solvent excluded surface

(SES). The triangulated surface meshes are generated with NanoShaper,⁴⁵ where the quality of the approximation is controlled with the parameter *GridScale*. A *GridScale* parameter equal to 2 refers to the cubic grid with spacing of $h = 0.5$ Å used in the construction of the surface mesh. *GridScale* = 4 corresponds to a spacing $h = 0.25$ Å, see Figure 4. A grid scale of 2 produces an average edge length of the triangulated surface mesh of around 0.45 Å, whereas *GridScale* = 4 results in an average edge length of around 0.225 Å.

Once the surface meshes are generated, we use TetGen⁴⁶ to create the initial tetrahedral meshes. For the adaptive mesh refinement (AMR) we use mng3d.⁴⁷ The implementation of the prototype version of the adaptive solver is realized in FreeFem++.⁴⁸

(a) *GridScale* = 2. (b) *GridScale* = 4.

Figure 4: Triangulated surfaces of a sphere with radius $R = 1$ Å, using NanoShaper.

2.2 MEAD solver

Here, we will give a brief overview on the Macroscopic Electrostatics with Atomic Detail (MEAD) software solving the linearized PBE (for more details, see Ref. 16). Taking into account the finite size of the grid, it prescribes a Dirichlet boundary condition for the ESP given by the analytical expression obtained for molecular point charges in a homogeneous dielectric environment with the dielectric coefficient of the solvent. This condition requires the grid to span a large area, increasing the cost of calculations drastically. To avoid this problem, MEAD adopts a multigrid focusing method. Here, the initial calculations are made with a coarse-grained grid covering a larger domain, followed by calculations on finer grids covering smaller subdomains, centered on the region of interest.

In our calculations, we used MEAD version 2.2.8a and its programs *potential* and *multiflex*. The former is applied to calculate the ESP at the coordinates specified by the user. We used *multiflex* to calculate the interaction energy. An important issue in the MEAD procedure involves calculating both energies E_{1-2} and E_{2-1} , expressed through potentials φ_1 (see (21)) and φ_2 (see (22)). The final result is obtained as the average of the two, i.e.,

$$\bar{E}^{MEAD} = \frac{E_{1-2}^{MEAD} + E_{2-1}^{MEAD}}{2}. \quad (50)$$

The implementation of MEAD allows the user to set model parameters including the dielectric coefficients inside and outside the molecular cavities, the ionic strength of the solvent, the grids dimensions and spacings, as well as to revert to the old way of testing for convergence of the iterative solver Successive Over Relaxation (SOR). The latter is a variant of the Gauss-Seidel method for solving a linear system of equations. The algorithm is controlled by the following parameters, hidden from the user: **minits**, **maxrmsdiff** and **maxiter**. The parameter **minits** specifies the minimum number of iterations and is defined as $30(N_d - 1)/2$, where

N_d is the number of grid points in each coordinate direction for both the coarse-grained and the fine mesh. The parameter `maxrmsdiff` is the tolerance for the maximum root mean square of the difference (`rmsdiff`) between two consecutive iterates φ_k and φ_{k+1} generated by the SOR method. By default, `maxrmsdiff` is set to $2 \cdot 10^{-5}/(N_d - 1)$. The parameter `maxiter` denotes the maximum number of iterations and is set to `10 · minits`. The SOR iterative method is terminated when either the maximum number of iterations is reached, or the current iteration number is greater than or equal to `minits`, while `rmsdiff` is less than or equal to `maxrmsdiff`. It turns out that the default values of these parameters are inappropriate in the majority of our tests. Therefore, in this paper, we also present results of MEAD for a set of manually improved values of SOR parameters, which requires a recompilation of the source code. The selected values of these and other parameters will be given for every problem separately. It is important to note that MEAD performs the calculations in single-precision arithmetics.

2.3 APBS solver

Similar to MEAD, the Adaptive Poisson-Boltzmann Solver (APBS) solves the linearized PBE¹, applying an FD method.¹¹ However, unlike MEAD, it employs double precision arithmetics in combination with a multigrid iterative solver and a better stopping criterion.

In analogy with the MEAD calculations, for each frame we perform two computations with APBS: one to find the potential φ_1^{APBS} generated by the charges of molecule M_1 at the atom positions $x_{i,2}$ of M_2 and another to find the potential φ_2^{APBS} by the charges of molecule M_2 at the atom positions $x_{i,1}$ of M_1 (subroutine *multival*). The corresponding electrostatic interaction energies E_{1-2}^{APBS} and E_{2-1}^{APBS} are obtained using (21) and (22), respectively. The final result is obtained by the average of the two, in analogy with (50):

$$\overline{E}^{APBS} = \frac{E_{1-2}^{APBS} + E_{2-1}^{APBS}}{2}. \quad (51)$$

To present the best case scenario for APBS, we use the finite difference solver `mg-auto` with parameters which provide a relatively good computational efficiency and a very high accuracy.

Below, we list the keywords, most relevant to our computations, from the input ".in" file of APBS, where "#" denotes the beginning of a comment in the input file. The full list of keywords together with particular values of the parameters they define is given in the SI.

`mg-auto` # specifies the type of calculations that the finite difference multigrid PMG code performs.

`dime nx ny nz` # are the number of grid points in the x-, y-, and z-direction, respectively, for all focused FD grids.

`cgcent mol 1` # indicates that the coarse grid is centred at the centre of the minimal box, with sides parallel to the coordinate axes, such that it contains all the atoms' centers of the system under study (mol 1 is its internal number).

`cglen xlen ylen zlen` # defines the computational domain which is a rectangular parallelepiped with side lengths *xlen*, *ylen*, and *zlen* in Å, in the x-, y-, and z-directions, respectively.

`fgcent mol 1` # indicates that all focused grids, including the finest one, are centred at the centre of the minimal box, with sides parallel to the coordinate axes, such that it con-

tains all the atoms' centres of the system under study (mol 1 is its internal number).

`fglen xlen ylen zlen` # specifies the fine mesh domain, a rectangular parallelepiped with side lengths *xlen*, *ylen*, and *zlen* in Å, in the x-, y-, and z-directions, respectively.

`etol tol` # specifies the tolerance for iterations of the partial differential equation solvers. This keyword is optional for the finite difference solvers and in particular for `mg-auto`. Its default value is *tol* = 1e-06 according to the APBS manual.

Note that, the values of the arguments *nx*, *ny*, *nz* of the keyword `dime` are dependent on the depth ℓ of the multilevel hierarchy used in the multigrid solver `mg-auto` by the formula $n_i = c2^{\ell+1} + 1$, where *i* is *x*, *y* or *z* and *c* is a nonzero positive integer. The choice $n = 321$ is compatible with $\ell = 5$, whereas $n = 385$ is in line with $\ell = 4$.

2.4 System under study

In this work we consider electrostatic interactions between the chromophores (dyes) Alexa 488 and Alexa 594, attached to all-trans polyproline helices with 6 and 20 residues (POLY6 and POLY20), placed in aqueous solution with 300 mM NaCl at temperature $T = 300$ K (see Figure 5), which was investigated earlier with MD simulations and electrostatic calculations, using MEAD.³⁷ While in our earlier work only the electrostatic interaction between transition charges of the dyes was investigated, here we include also the interaction between ground-state charge densities in order to compare the performance of different solvers for a larger variety of conditions. In addition, this interaction is relevant for an implicit description of the solvent. The coupling be-

Figure 5: Polyproline helix with 6 amino acid residues (POLY6), labeled with the two chromophores Alexa 594 and Alexa 488 at the two ends of the helix.

tween the dyes was computed for certain time frames taken from all-atom molecular dynamics (MD) simulations explicitly including the water molecules and ions of the solvent, performed with the NAMD software package.⁴⁹ For both polyproline lengths (6 and 20 residues), we generated 10 trajectories, each with a time window of 200 ns. The standard CHARMM force field (*version v35b3*)⁵⁰ was used for polyproline. The force field parameters of the Alexa chromophores were created by an analogy approach from that of similar chemical groups in the CHARMM force field (for details on the MD simulations including force field parameters see Ref. 37). The water molecules in the MD simulations were described, using the TIP3P model.⁵¹

The calculation of the interaction between Alexa chromophores by MEAD, APBS and ARGOS follows the same procedure, starting with computing the ESP from the Poisson or LPBE (see (5) and (4)). This potential is then used to calculate the coupling (see (21) and (22)).

¹APBS can solve both the PBE and the linearized PBE with the FD method. The application of the finite element method, however, is still at a very experimental stage and in our calculations we could not use it.

3 RESULTS

3.1 Evaluation of the accuracy of the numerical methods

We start with three models, for which analytical solutions for the ESP exist. By comparing the numerical with the analytical solutions, the sources of numerical errors and their magnitude will be determined. The numerical accuracy of the solvers (MEAD, APBS and ARGOS) is evaluated by the relative errors (in %) defined as

$$\begin{aligned} Err^{Adj} = Err^{ARGOS} &= \frac{|E - E_{1-2}^{Adj}|}{|E|} \cdot 100, \\ Err^{Prim} &= \frac{|E - E_{2-1}^{Prim}|}{|E|} \cdot 100, \\ Err^{MEAD} &= \frac{|E - \bar{E}^{MEAD}|}{|E|} \cdot 100, \\ Err^{APBS} &= \frac{|E - \bar{E}^{APBS}|}{|E|} \cdot 100, \end{aligned} \quad (52)$$

where E is the exact (analytical) solution, $Err^{Adj} = Err^{ARGOS}$ and Err^{Prim} correspond to the relative errors of the numerical solutions $E_{1-2}^{Adj} (= E^{ARGOS})$ and E_{2-1}^{Prim} of the adjoint and the primal problem, respectively, obtained with ARGOS. $\bar{E}^{MEAD}$ and $\bar{E}^{APBS}$ are the numerical solutions, obtained with MEAD and APBS, respectively, and Err^{MEAD} and Err^{APBS} are the corresponding relative errors. The different models, studied below, were constructed to focus on particular aspects that limit the accuracy of the numerical solution. The first model includes chromophores without interfaces and examines the effect of discretization of the charges and the imposed boundary conditions on the ESP φ . Models 2 and 3 employ a Born ion model, represented by a spherical cavity with a central charge, and are designed to probe the role of the interface and of the ionic charge density of the solvent.

3.1.1 Model-1: Boundary conditions and discretization of charges

The setup consists of a polyproline helix (POLY6 or POLY20), labelled with charge-neutral Alexa dyes in the electronic ground state. The system is placed in vacuum ($\epsilon_s = 1$). The dielectric coefficient ϵ_m of the chromophores and polyproline are also set to 1. Therefore, the model does not include any interfaces and the analytical solution of the electrostatic coupling between the dyes is just a sum over pairwise Coulomb interactions between point charges. The couplings were calculated for 19 snapshots (frames), randomly selected from the MD simulations: frames 1 to 10 and 11 - 19 contain POLY6 and POLY20, respectively. The sizes of the coarse-grained and fine (focused) meshes in the MEAD and APBS calculations are selected taking into account the dimensions of the system. The coarse-grained, the intermediate-focused, and the fine-focused meshes for MEAD are centered at the geometric center of the system, whereas the meshes in the APBS calculations are centered at the center of the minimal box (with sides parallel to the coordinate planes) that contains all the atoms of the molecule. Since the default stopping criterion for the iterative solver SOR in MEAD is not adequate in most of our tests, we used 2 different sets of SOR parameters (see Table 1 and METHODS section for more details). The complete specification of configurations M1-1, M1-2 for MEAD and A1 for APBS,

used in the computations of Model-1 are given in Tables 1 and 2, respectively.

Table 1: Computational configurations of MEAD for Model-1

Configuration	Coarse-grained mesh number of nodes/spacing (Å)	Intermediate mesh number of nodes/spacing (Å)	Fine mesh number of nodes/spacing (Å)	SOR maxiter	SOR maxrmdiff
M1-1	101 / 3	201 / 1	POLY6: 321 / 0.25 POLY20: 401 / 0.25	default (POLY6: 1500 and 4800 POLY20: 1500 and 6000)	default $2 \cdot 10^{-5}/(N_d - 1)$
M1-2	101 / 3	201 / 1	POLY6: 321 / 0.25 POLY20: 401 / 0.25	for all cases 12000	$2 \cdot 10^{-5}/(N_d - 1)$

Table 2: Configuration A1 of APBS for Model-1.

	dime	cglen grid spacing (Å)	fglen grid spacing (Å)	Automatically placed intermediate focused mesh: size / grid spacing (Å)
POLY6	321 321 321	900 900 900 / 2.813	80 80 80 / 0.25	268.33 268.33 268.33 / 0.839
POLY20	385 385 385	900 900 900 / 2.344	100 100 100 / 0.26	300 300 300 / 0.781

The fine mesh domain for both MEAD and APBS is a box with a side length of 80 Å for POLY6 and 100 Å for POLY20, and encompasses the polyproline and the chromophores with an average margin (over all frames) of at least 10 Å from all sides. We have chosen the values of the parameters of **dime** and of **fglen** in APBS so that they are as close as possible to the parameters in the grid configurations M1-1 and M1-2 for MEAD. The values of **cglen**, however, are chosen much larger, i.e., 900 Å, which can improve the accuracy of the APBS computations compared to MEAD. We also note that the **mg-auto** solver automatically chooses how many intermediate focusing grids are used between the coarse-grained and the fine mesh. In the case of configuration A1, with **cglen** 900 900 900 and **fglen** 80 80 80 or **fglen** 100 100 100, **mg-auto** places one intermediate focusing grid with dimensions given in Table 2.

The computational domain for ARGOS is a cube with a side length of around 160 000 Å for all frames with POLY6 and around 400 000 Å for all frames with POLY20. To demonstrate the convergence of the ARGOS results with respect to the number of degrees of freedom (DOFs), we compute the relative errors for the interaction energies obtained from the solutions of the adjoint and primal problems, Err^{ARGOS} and Err^{Prim} defined in (52), and average them over all frames (Figure 6). Both Err^{Adj} and Err^{Prim} exhibit the order of convergence, predicted by (45) and (46). The solution of the adjoint problem exhibits a higher convergence order $O(\#DOFs^{-1})$ (see (46)), and thus, much smaller relative errors, ranging from 2×10^{-1} % down to 2×10^{-4} % for the largest number of DOFs (10^7). For a nonhomogeneous Dirichlet boundary condition, the relative errors Err^{ARGOS} , Err^{Prim} and the respective convergence orders for charged and neutral chromophores are given in Figures S5, S6, and S7. This boundary condition is given by the Coulomb potential (see (11)) for the point charges defining the respective problem (primal or adjoint) with a uniform dielectric coefficient $\epsilon_m = \epsilon_s = 1$. In this case, it is given by the exact potential on the boundary of the computational domain Ω and coincides with the boundary condition used by MEAD and APBS.

The relative errors of MEAD, APBS, and ARGOS are compared in Figure 7. The FE-solver ARGOS is orders of magnitude more accurate than the FD-solvers MEAD and APBS. In the case of MEAD, adjusting the SOR parameters in M1-2 helps to decrease the relative errors for many frames, except frame 12. The maximum error for configuration M1-2 is limited to 2%. We conclude that, if the number of iterations in SOR is too small, the iterative error dominates

Figure 6: Model-1. Average (over 19 frames) relative errors in the interaction energies E_{1-2}^{Adj} and E_{2-1}^{Prim} (Eq. (52)) computed by ARGOS as a function of the number of DOFs (averaged over the same 19 frames). The dashed lines represent the optimal convergence rates for the finite element spaces used, i.e P_1 for the primal and P_2 for the adjoint problem (see (45) and (46)).

the discretization error. A significant increase of the number of SOR iterations helps to improve the accuracy up to a certain point. Further increase results in an accumulation of round off errors, due to the single precision arithmetics used and no further decrease of the overall error is possible.

APBS performs much better than MEAD for almost all frames. The error of the former does not exceed 0.4% and can be lower than the one of ARGOS for MRL = 0. However, further refinements gradually decrease the relative errors of ARGOS, reaching about 5×10^{-3} % at most for MRL = 4 for ARGOS. If one continues with the mesh refinement procedure in ARGOS, the relative errors drop even further, see Figure S4.

Figure 7: Model-1. Relative errors of MEAD (red and green symbols), APBS (blue symbols), and ARGOS (grey symbols) (Eq. (52)) for the coupling between chromophores in charge-neutral ground state in a uniform dielectric for different configurations of MEAD and APBS, and different MRLs of ARGOS. The different frames correspond to different molecular geometries taken from MD simulations.

3.1.2 Model-2: Interface description

As discussed in the introduction, the approximation of the geometric features of the interfaces between different dielectric media, as well as the approximation of the continuity conditions, (6) and (7), for the potential φ and for the normal component of the displacement field $\epsilon \nabla \varphi$, are critical for the accuracy, particularly of the FD solvers. The description of sharp interfaces, where the dielectric coefficient

undergoes a considerable jump, is especially important for the computation of the electrostatic interaction. Furthermore, the presence of charges close to the interface can also have a strong impact on the accuracy.

In order to have an analytical solution to compare with, we investigate the interaction between a charge $q_1 = 1e_0$ in the center of a spherical cavity and a probe charge $q_2 = -1e_0$ outside the cavity. Different dielectric coefficients $\epsilon_m = 4$ inside and $\epsilon_s = 80$ outside the cavity are assumed. The spherical cavity with the central charge in a homogeneous dielectric is known from the Born model of ion solvation.

In a first model, a cavity with radius $R = 2 \text{ \AA}$ is assumed (Figure 8). The coordinates of the ion centre are $(0.05, -0.03, -0.25)$. The electrostatic interaction energy was calculated for different distances r of the probe charge from the centre of the ion for $1 \text{ \AA} < r < 35 \text{ \AA}$. To track the dependence of the MEAD and APBS results on the relative orientation of the FD-solver's grid and the probe charge position, we computed the electrostatic interaction for 10 different directions $\vec{d}_i$, $i = 1, 2, \dots, 10$ from the ion's centre, along which the probe charge was placed: $\vec{d}_1 = (1, 0, 0)$, $\vec{d}_2 = (-1, 0, 0)$, $\vec{d}_3 = (0, 1, 0)$, $\vec{d}_4 = (0, -1, 0)$, $\vec{d}_5 = (0, 0, 1)$, $\vec{d}_6 = (0, 0, -1)$, $\vec{d}_7 = (1, 1, 1)$, $\vec{d}_8 = (-1, -1, -1)$, $\vec{d}_9 = (1, 0.5, 0.2)$, and $\vec{d}_{10} = (-1, -0.5, -0.2)$. The relative errors in the interaction energies calculated for the same r along different directions $\vec{d}_i$ were averaged over all $i \in \{1, 2, \dots, 10\}$ to obtain the values $\overline{Err}^{MEAD}(r)$ and $\overline{Err}^{APBS}(r)$ for each distance r (for details see Eq. (3) in the SI).

To obtain the interaction energy with MEAD and APBS we computed the ESP at the position of the probe charge and multiplied it by q_2 . The configurations M2-1 and M2-2 for MEAD, and A2 for APBS used in the computations of Model-2 are given in Tables 3 and 4, respectively. All grids (coarse-grained and focused) for both MEAD and APBS are centered at $(0, 0, 0)$, i.e., slightly offset from the centre of the ion. In the case of MEAD, the calculations were performed for two different sets of SOR parameters (Table 3). APBS yields very similar results for calculations with `etol` $1e-06$ and `etol` $1e-09$. Therefore, we present only the latter results.

Table 3: Computational configurations, used in MEAD calculations for Model-2.

Configuration	Coarse-grained mesh number of nodes/spacing (Å)	Intermediate mesh number of nodes/spacing (Å)	Fine mesh number of nodes/spacing (Å)	SOR maxiter	SOR maxrmdiff
M2-1	101 / 3	161 / 1	161 / 0.25	default (1440 and 2400)	default $2 \cdot 10^{-8}/(N_d - 1)$
M2-2	101 / 3	161 / 1	161 / 0.25	for all meshes 12000	$2 \cdot 10^{-8}/(N_d - 1)$

Table 4: Configuration A2 of APBS for Model-2. The `etol` parameter is set to $1e-09$.

dime	cglen grid spacing (Å)	fglen grid spacing (Å)	Automatically placed intermediate focused mesh: size / grid spacing (Å)
385 385 385	900 900 900 / 2.344	96 96 96 / 0.25	293.94 293.94 293.94 / 0.765

In the ARGOS calculations, we used `GridScale` = 2 and 4 with solutions taken at MRL = 5 and MRL = 3, respectively. The computational domain Ω is a cube with an edge length of approximately 1 350 000 Å. Here, we only present the ARGOS results for a homogeneous Dirichlet boundary condition (this is possible since the boundary of Ω is very far from the ion and the probe charge). The results for a non-homogeneous boundary condition are presented in the SI,

Figure S11. In this case the boundary potential is exactly prescribed, since the Coulomb potential coincides with the exact one given by (53) on the boundary of Ω and the quality of the potential² depends only on the geometric approximation of the surface of the spherical cavity. This boundary condition, being exact, also ensures the validity of the natural symmetry $E_{1-2} = E_{2-1}$ for the large, but still bounded computational domain Ω .

An illustration of the mesh for *GridScale* = 2 before and after refinement is presented in Figure 8. Comparison of the initial and final mesh topologies shows significant refinement in the critical areas: around the ion surface, the points, where the ion charge and probe charge are located, and in between the two charges. The convergence of the relative errors with increasing refinement levels of ARGOS is shown in Figures S8 and S11 in the SI.

The analytical solution for the interaction energy between the ion and the probe charge follows from the well-known expression for the electrostatic potential φ :

$$\varphi(r) = \begin{cases} \frac{q_1}{\epsilon_m r} + \frac{q_1 (\epsilon_m - \epsilon_s)}{\epsilon_m \epsilon_s R}, & r \leq R, \\ \frac{q_1}{\epsilon_s r}, & r \geq R. \end{cases} \quad (53)$$

The relative errors of the calculations obtained with MEAD, APBS and ARGOS using different solver configurations are given in Figure 9. It is readily seen, that all configurations of MEAD and APBS produce a significant error at the interface ($r = 2$ Å) between the ion and the solvent. Whereas for the default SOR parameters the errors of MEAD increase with increasing r , for the adjusted SOR parameters, the relative errors drop for distances $r > 2$ Å and do not exceed 0.3 % for distances larger than 6 Å. The errors of ARGOS are substantially smaller for short distances. The largest relative error occurs at $r = 1$ Å and reaches 2% for *GridScale* = 2. At the interface, the relative error decreases to 0.4 and 0.2 % for *GridScale* = 2 and 4, respectively.

For an ion radius $R = 1$ Å, the relative errors of MEAD and APBS are approximately 60 % at the interface between the two dielectrics (See Figure S13 in the SI). This result shows that the presence of charges near sharp interfaces has a strong impact on the accuracy of the FD approximation of the electrostatic interaction. The relative error for ARGOS is also maximal at the interface, but it does not exceed 1 % for, both, *GridScale* = 2 and 4. The error for $r > 1$ Å is around 10^{-2} % for, both, *GridScale* = 2 and 4. The convergence of the relative error for ARGOS with respect to the mesh refinement levels is shown in the SI, Figures S8, S11, and S12.

We also recall that the results obtained by MEAD and APBS for a grid spacing $h = 0.25$ Å in the final focused mesh should be compared to the ARGOS results with *GridScale* = 4, since the resolution of the triangulated molecular surface mesh in this case is determined from the intersection of a cubic grid with the same grid spacing $h = 0.25$ Å and the analytical representation of the SES.

To further investigate the influence of the interface on the accuracy of MEAD, APBS, and ARGOS we consider interaction between a Born ion with increased radius and a probe charge, placed near the surface of the ion cavity (see Figures S14, S15, S16, and S17 in the SI). The results reveal

that the relative error decreases when the interface is further away from the charge of the ion (located in the center of the spherical cavity). The relative errors at the interface at $r = 16$ Å are approximately 3 % on average for, both, MEAD and APBS with a grid spacing $h = 0.25$ Å in the last focused grids. The accuracy does not improve much when the grid spacing h of the last focused grids is reduced to 0.15 Å. In this case, the relative errors at the interface are approximately 2 % for MEAD and APBS. Although, a relative error of 2-3 % is not large, one should keep in mind that this case corresponds to only one ion and one probe charge. Since the error is additive, the chance of getting large errors in the interaction energy with FD solvers increases for large molecular systems. This effect will be revealed in Section 3.2, where we investigate the error in the electrostatic interaction between chromophores in a dielectric solvent environment obtained by MEAD and APBS. Similarly to the results presented in Figures 9 and 10, the relative errors for ARGOS are more than two orders of magnitude smaller around the interface than those of MEAD and APBS (Figure S14).

3.1.3 Model-3: Presence of interfaces and ions

We supplemented Model-2 with a continuous charge density of ions outside the cavity, described by an ionic strength of 0.3 M and taking into account the ion exclusion layer (IEL). Here, the IEL is enclosed between the ion sphere and another concentric sphere with a radius enlarged by that of a probe ion's radius $R_{\text{ion}} = 2$ Å. For MEAD and APBS we used the configurations, defined in Tables 3 and 4, respectively. The APBS results for `etol 1e-06` and `etol 1e-09` again do not differ, and thus we report only the latter results. The computational domain for ARGOS is the same as for Model-2. We compare the solutions of ARGOS obtained at MRL = 3 and 4 for *GridScale* = 2 and 4, respectively, with those obtained by MEAD and APBS. The convergence of the ARGOS relative errors with increasing mesh refinement levels is shown in the SI, Figure S18. The analytical solution for the ESP φ in the present model is given as:

$$\varphi(r) = \begin{cases} \frac{q_1}{\epsilon_m r} + \frac{q_1 (\epsilon_m - \epsilon_s)}{\epsilon_m \epsilon_s R} - \frac{q_1 \frac{\bar{k}}{\sqrt{\epsilon_s}}}{\left(1 + a \frac{\bar{k}}{\sqrt{\epsilon_s}}\right) \epsilon_s}, & r \leq R, \\ \frac{q_1}{\epsilon_s r} - \frac{q_1 \frac{\bar{k}}{\sqrt{\epsilon_s}}}{\left(1 + a \frac{\bar{k}}{\sqrt{\epsilon_s}}\right) \epsilon_s}, & R \leq r \leq a, \\ \frac{q_1 \exp\left(a \frac{\bar{k}}{\sqrt{\epsilon_s}}\right) \exp\left(-r \frac{\bar{k}}{\sqrt{\epsilon_s}}\right)}{\left(1 + a \frac{\bar{k}}{\sqrt{\epsilon_s}}\right) \epsilon_s r}, & r \geq a, \end{cases} \quad (54)$$

where r is the distance of the probe charge from the center of the Born ion and $a = R + R_{\text{ion}}$.

The results of the calculations, shown in Figure 10, demonstrate a drastic change in the MEAD accuracy for M2-1 in comparison to Model-2 (Figure 9): the errors do not exceed 2% for all distances r , while for Model-2 the errors go beyond 10%, when r is increasing. Moreover, for $r > 20$ Å, the error depends more on the mesh spacing than on the SOR parameters, in contrast to Model-2. Obviously, the discretization error dominates the error of the iterative linear solver (for other MEAD configurations see also Figures S9 and S19 in the SI). It should also be mentioned that for configuration M2-2 the respective maximum iteration number `maxrmsdiff` is not reached in the iterative process.

Similarly to Model-2, the errors of MEAD and APBS are particularly large around the interface between the ion and the solution, reaching more than 30 %. Decreasing the distance between the singular charge density and the interface results in a further increase of those errors. For an ion with

²In the case of exactly prescribed Dirichlet boundary data, the numerical solution converges to the exact solution of the LPBE for the particular triangulation of the spherical cavity. This exact solution is different from the exact solution given by (53) for a cavity which is a perfect sphere, see Figures S8 and S11 in the SI for the convergence of the relative error with respect to the MRLs in the case of nonhomogeneous Dirichlet BC.

Figure 8: ARGOS ($GridScale = 2$) mesh refinement for the solution of the primal and adjoint problems for Model-2. Here, the ion sphere radius is 2 Å and the probe charge is at a distance $r = 15$ Å from the center of the spherical ion. The color code corresponds to the solution of the adjoint problem (28) for the ESP in units of $10^{-3}e_0/\text{Å}$.

Figure 9: Model-2. Relative errors (Eq. (52)) for MEAD and APBS averaged over 10 directions $\vec{d}_i$ between ion and probe charge (Eq. (3) in the SI) together with the relative errors for ARGOS as a function of the distance r between the ion and the probe charge. The ARGOS calculations were performed with MRL = 5 and MRL = 3 for $GridScale = 2$ and 4, respectively.

Figure 10: Same as in Figure 9, but for Model-3 and performing the ARGOS calculations with MRL = 3 and 4 for $GridScale = 2$ and 4, respectively.

a radius $R = 1$ Å, they are as high as 70 % and 50 % for MEAD and APBS, respectively, while ARGOS exhibits even smaller errors than for $R = 2$ Å (Figure S22 in the SI).

The results for all test models, studied above, reveal that the accuracy of MEAD is greatly compromised even for the simplest case of point charges in vacuum. The presence of interfaces with a high jump in the dielectric coefficient leads to a strong dependence of the accuracy on the mesh spacing and the SOR stopping criterion, especially for long-range interactions. Besides the stopping criterion issue, the use of single precision arithmetics can also lead to significant round-off errors. The results of APBS calculations exhibit much smaller relative errors than MEAD, but still do not reach the accuracy of ARGOS, which exhibits a predicted and uniform convergence. The test calculations also reveal,

that $GridScale = 2$ and refinement levels between 3 and 5 for ARGOS result in a high enough accuracy and should be sufficient for most cases.

In the next sections we will use ARGOS as a reference to explore the accuracy of the MEAD and APBS solvers in computations of interaction energies in real systems.

3.2 Interactions between chromophores in a dielectric environment

The interaction energy between two chromophores, described by molecule-shaped cavities with atomic partial charges, in a dielectric environment cannot be computed analytically. Therefore, we exploit the high accuracy of the ARGOS results and determine the deviations, D_{AM} and D_{AA} , of the MEAD and APBS calculations, respectively, relative to the solution E_{1-2}^{Adj} of the adjoint problem of ARGOS as:

$$D_{AM} = \frac{|E_{1-2}^{Adj} - \bar{E}^{MEAD}|}{|E_{1-2}^{Adj}|} \cdot 100. \quad (55)$$

$$D_{AA} = \frac{|E_{1-2}^{Adj} - \bar{E}^{APBS}|}{|E_{1-2}^{Adj}|} \cdot 100. \quad (56)$$

Refinement level $\ell = 4$ and $GridScale = 2$ are employed in the ARGOS calculations. The computational domain is a cube with a side length between 160 000 Å and 400 000 Å depending on the size of the smallest box containing the molecular system. For MEAD and APBS we used configurations M1-2 and A1 given in Tables 1 and 2, respectively.

We start with the coupling between the charge densities of the electronic ground states of the two chromophores (see Tables S3 and S4 for the respective partial charges) in equilibrium with the solvent. The orientational polarization of water molecules dominates the dielectric response of the solvent. The large dielectric coefficient of water results in a sharp interface at the molecular surface, where ϵ jumps from a low value for the chromophores or the protein ($\epsilon_m = 4$) to the large value $\epsilon_s = 80$ for the aqueous solvent. In this case, the accuracy of the numerical solution of the electrostatic problem critically depends on the approximation of the interfaces.

To study the impact of the description of interfaces, we investigate different MD frames for POLY6 and POLY20, as in Model-1 studied above, but using the following models with different types of interfaces:

- NO-INT, no interfaces: chromophores are represented by point charges placed at the positions of

atoms, immersed in a homogeneous environment with a dielectric coefficient ϵ_s ;

- DYES, dyes-solution interface: chromophores are represented by point charges inside molecule-shaped cavities with a dielectric coefficient ϵ_m inside, surrounded by a homogeneous dielectric with a coefficient ϵ_s , representing the aqueous solvent (see Figure 11a);
- DYES-POLYPRO, dyes-solution and protein-solution interfaces: chromophores are represented like in the previous model; polyproline is included as a molecule-shaped cavity with a dielectric coefficient ϵ_m inside (see Figure 11b);
- DYES-POLYPRO-IONS, dyes-solution and protein-solution interfaces: the previous model is complemented with ionic charge density in the aqueous solvent, characterized by an ionic strength of 0.3 M (see Figure 11c).

The Alexa chromophores in their electronic ground state are negatively charged by two elementary units ($-2e_0$).³⁷ In order to investigate, how the interchromophore charge density coupling changes, if no monopole moments are present, we artificially neutralize the chromophores by adjusting the charges within the SO_3 groups such that the overall charge of the chromophores is zero, as described in the SI (Tables S5 and S6). Please note that, this neutralization is not realistic for the present chromophores but represents a common situation for other chromophores, which would, of course, differ in the details of the charge density. Note also that, these charges have been employed in Model-1 above.

3.2.1 Charged dyes

The deviations D_{AM} and D_{AA} of the electrostatic interaction energies between the charged chromophores obtained with MEAD and APBS and the ARGOS reference solutions are presented in Figure 12 for three different models described above.

Examination of the MEAD results shows, that including the polyproline domain can notably change the error (for example, frames 5, 9 and 10 in Figure 12), but it still stays small. The maximum value of D_{AM} does not exceed 3%. Furthermore, for all frames, except frame 16, including the ionic strength of the solvent results in an increase of D_{AM} . For example, the relative deviations for frames 5 and 9 have increased 100 times. To understand these effects we investigate the reference interaction energies obtained with ARGOS (see Figure 13 and Figure 14 for logarithmic scale of the y -axis).

As evident from Figure 13, the difference between interaction energies, computed for different models, is more pronounced for POLY6 (frames 1 to 10). Introducing molecular cavities for the dyes and polyproline, increases the interaction energy in comparison to the simple model of a homogeneous dielectric. This effect is due to the lower dielectric coefficient inside the cavities. For larger interchromophore distances this effect gets smaller.

The results, shown in Figure 14, demonstrate a significant screening of the interaction by the ions, resulting in a decrease of the interaction energy for all frames. The magnitude of the screening effect for POLY6 can vary significantly, reducing the interaction energy by factors between 1.5 and 100, depending on the conformation of the system (see Figure 14). For the larger distances between the chromophores in POLY20, the magnitude of the screening effect varies stronger: between 10^2 to 10^6 times (see the results for

frames 11 and 12 in Figure 14). The large increase in screening for POLY20 with respect to POLY6 can be explained in the following way. When the charged chromophores are far apart, they can be approximated by Born ions, immersed in the dielectric continuum with ionic strength. In this case, the interaction energy can be obtained analytically from the potential in (54). It includes an exponential factor, which imposes a highly nonlinear decline with increasing distance. To estimate its influence, we compare the screening effect for frames 12 and 14, which correspond to interchromophore distances of 78 Å and 42 Å, respectively. The ratio between the interaction energies, calculated by ARGOS, is about 10^{-4} (see Figure 14 for the model DYES-POLYPRO-IONS). The estimation of this ratio, using (54) with $\bar{k}^2 = 2.529 \text{ Å}^{-2}$, gives $9 \cdot 10^{-4}$, providing a qualitative explanation of the effect of the exponential screening factor. Without exponential factor we obtain a ratio of 0.5. An examination of the results in Figure 14 shows a significant drop in the interaction energies due to ion screening for frames 12, 18 and 19. At the same time these frames exhibit the largest relative deviations for MEAD (Figure 12).

The significant increase of D_{AM} for the model DYES-POLYPRO-IONS has several reasons. First of all, the accuracy of MEAD for M1-2 is worse for systems with ions and long distances between interacting chromophores. This effect can be readily seen from comparison of the results for Model-2 and Model-3 in setup M2-2 (see Figures 9 and 10), which has the same parameters (grid spacing and SOR) as M1-2. The errors for Model-3 increase with distance. The accuracy for the model DYES-POLYPRO-IONS can probably be improved by applying a finer mesh, but this would increase the computational costs substantially.

The APBS results in Figure 12 deviate from the reference solution of ARGOS much less, compared to those obtained with MEAD. Only D_{AA} for frame 12 (in the model DYES-POLYPRO-IONS) approaches 20%. Similar to the MEAD case, the deteriorated accuracy can originate from the extremely small interaction energy. On the other hand, in this frame the system is in a maximally stretched conformation, see Figure 17d. The minimal box that contains the polyproline system has dimensions of $90.7 \times 25.1 \times 17.1 \text{ Å}^3$ and it is only slightly smaller, than the box of the finest focused grid (`fglen 100 100 100`) in configuration A1 of APBS. Hence, it probably does not provide enough margins between the sides of the fine grid box and the atoms of the system together with the IEL. To confirm this hypothesis, we computed the electrostatic energy with enlarged fine grid box (`fglen 110 110 110`), keeping all other parameters of configuration A1. Moreover, this configuration of APBS was employed to analyze the sensitivity of the results (for frame 12) with respect to the relative position of the molecular system and the grid. For this setup, electrostatic interaction energies, together with their mean values and standard deviations were calculated for a perturbed grid center, given by $(x_0 + \Delta x_i, y_0 + \Delta y_i, z_0 + \Delta z_i)$ for $i = 1, 2, \dots, 15$ where (x_0, y_0, z_0) is the center of the minimal box, containing all atoms' centers, with sides parallel to the coordinate axes, and $(\Delta x_i, \Delta y_i, \Delta z_i)$ takes the values in (10) from the SI with $\Delta h = 0.15 \text{ Å}$. The interaction energies were averaged over all 15 grid positions, giving average values $\bar{E}_{1-2}^{APBS}$ and $\bar{E}_{2-1}^{APBS}$, and their respective standard deviations σ_{1-2} and σ_{2-1} . The final value for the interaction energy, $\bar{E}^{APBS}$, was obtained, as usual, by averaging $\bar{E}_{1-2}^{APBS}$ and $\bar{E}_{2-1}^{APBS}$.

The results for frame 12 with the enlarged fine grid box in configuration A1 are presented in Table 5, demonstrating a pronounced asymmetry in the average values $\bar{E}_{1-2}^{APBS}$ and $\bar{E}_{2-1}^{APBS}$ for the larger `eto1` value (for details see Table S9 in the SI where one can also infer the absolute values of

Figure 11: Interfaces in different models explained in the text. The conformation of the to dyes and the shorter polyroline helix POLY6 corresponds to frame 10 from the MD simulation mentioned in the text. Figures generated using VisIt.⁵²

Figure 12: Relative deviations of the interaction energies of charged chromophores, obtained with MEAD (green symbols) and APBS (blue symbols) with respect to the reference ARGOS solutions. The electrostatic energies are calculated for MEAD configuration M1-2 from Table 1 and APBS configuration A1 from Table 2. The circles, squares, triangles correspond to the models DYES, DYES-POLYPRO, DYES-POLYPRO-IONS, respectively, that are explained in the text. The different frames correspond to different molecular geometries taken from MD simulations.

the errors). For comparison, the adjoint solutions E_{1-2}^{ARGOS} and E_{2-1}^{ARGOS} of ARGOS are equal to $-5.5959e-10$ and $-5.5967e-10$, respectively, for $GridScale = 2$ at $MRL = 4$. The average value of $\overline{E}_{1-2}^{APBS}$ and $\overline{E}_{2-1}^{APBS}$ comes closer to the above reference value of ARGOS than the single values (Table 5). Also, the average over different grid positions improves the accuracy, as seen from Table S9 and from the standard deviations in Table 5. Decreasing the value of the parameter `etol` improves the accuracy as well. All these effects reduce the relative deviation of the APBS result for frame 12 from the reference (ARGOS) value $-5.5959e-10$ $e_0^2/\text{\AA}$ in Figure 12 to 3 % in Table 5. Reduction of the grid spacing of APBS or equivalently an increase of the number of grid points is expected to further decrease the relative deviation D_{AA} . However, a limiting factor for this strategy would be primarily the memory capacity required to store the coefficients and the potential values at all grid points.

3.2.2 Neutral dyes

The relative deviations for the interaction energies between ground state charge densities of neutral chromophores, presented in Figure 15, reveal a drastic increase compared to the previous case of charged chromophores in Figure 12. For MEAD computations, the relative deviations exceed 10% for most of the models and frames. Moreover, for some frames

Table 5: Sensitivity of the APBS results to the `etol` parameter and the relative position of the grid and the molecular system for **frame 12** for model **DYES-POLYPRO-IONS** with **charged chromophores**. σ_{1-2} and σ_{2-1} are standard deviations of interaction energies obtained for different grid positions. All energies in units of $e_0^2/\text{\AA}$.

h (Å)	$\overline{E}_{1-2}^{APBS}$	σ_{1-2}	$\overline{E}_{2-1}^{APBS}$	σ_{2-1}	Avg. $\overline{E}^{APBS}$	$D_{AA}(\%)$
0.286	-4.82e-10	1.80e-11	-6.94e-10	1.23e-11	-5.88e-10	5.08
<code>etol</code> 1e-06						
0.286	-5.79e-10	1.79e-11	-5.79e-10	1.16e-11	-5.79e-10	3.45
<code>etol</code> 1e-09

Figure 13: Comparison of the interaction energy between charged chromophores for all models, calculated with ARGOS. The asterisks, circles, squares, triangles correspond to the models NO-INT, DYES, DYES-POLYPRO, and DYES-POLYPRO-IONS, respectively. The different frames correspond to different molecular geometries taken from MD simulations.

Figure 14: Same as in Figure 13, but with logarithmic scale of vertical axis. Note that, the interaction energies of the first three models cannot be distinguished in the present representation.

Figure 15: Relative deviations of interaction energies of chromophores in the charge-neutral ground states, obtained with MEAD (green) and APBS (blue) with respect to the ARGOS results. MEAD configuration M1-2 from Table 1 and APBS configuration A1 from Table 2 were used. The circles, squares, triangles correspond to the models DYES, DYES-POLYPRO, DYES-POLYPRO-IONS, respectively. The different frames correspond to different molecular geometries taken from MD simulations.

Figure 16: Comparison of the interaction energy between the chromophores in the charge-neutral ground states for all models, calculated with ARGOS. The asterisks, circles, squares, triangles correspond to the models NO-INT, DYES, DYES-POLYPRO, and DYES-POLYPRO-IONS, respectively. The different frames correspond to different molecular geometries taken from MD simulations.

MEAD produces a different sign of the interaction energies compared to ARGOS (see frames 15 and 10 for M1-1 in Table 6), and even a different sign of E_{1-2}^{MEAD} and E_{2-1}^{MEAD} (see frames 15 and 16 for M1-1 in Table 6). Obviously, a strong asymmetry ($E_{1-2}^{MEAD} \neq E_{2-1}^{MEAD}$) is an indicator of a poor accuracy. In contrast to the MEAD calculations, the energies E_{1-2}^{Adj} computed by ARGOS match the energies E_{2-1}^{Adj} very well. The respective values are even computed on two different sequences of adapted meshes, both starting from one and the same initial mesh, demonstrating again a very good symmetry, and thus the superior performance of the FE-solver. The ARGOS calculations in Figures 15 and 16 and in Table 6 were performed with $GridScale = 2$ at $MRL = 4$. Increasing the $GridScale$ to 4 in the calculations for frames 15 and 10 leads to very similar energies E_{1-2}^{Adj} and E_{2-1}^{Adj} of $1.9649e-07$ and $1.9648e-07$, and $2.7915e-07$ and $2.7911e-07$, respectively.

Similarly to the previous case, the APBS calculations deviate much less than the MEAD results from the ARGOS solutions: D_{AA} values do not exceed 10% except frame 6 (model DYES-POLYPRO-IONS) and frame 10 (model DYES). However, unlike frame 12 in the system

Table 6: Interaction energies between ground state charge densities of neutral dyes, calculated with MEAD and ARGOS (with $GridScale = 2$ at $MRL = 4$) for different frames and interface models. The table demonstrates an unphysical asymmetry of the MEAD interaction energies: $E_{1-2}^{MEAD} \neq E_{2-1}^{MEAD}$. All energies in units of $e_0^2/\text{\AA}$.

		E_{1-2}^{MEAD}	E_{2-1}^{MEAD}	$\bar{E}^{MEAD}$	E_{1-2}^{Adj}	E_{2-1}^{Adj}
Frame 15,	M1-1	1.20e-07	-1.87e-07	-3.36e-08	1.9797e-07	1.9796e-07
	DYES	M1-2	1.03e-07	9.15e-08		
Frame 10,	M1-1	-1.65e-07	-4.86e-08	-1.07e-07	2.8530e-07	2.8539e-07
	DYES	M1-2	1.36e-07	2.00e-07		
Frame 16,	M1-1	2.33e-07	-1.59e-07	3.66e-08	2.9233e-07	2.9232e-07
	DYES-POLYPRO	M1-2	2.75e-07	2.68e-07		

of charged dyes, the minimal boxes for frames 6 and 10 ($31.2 \times 20.7 \times 15.3 \text{ \AA}^3$ and $48.2 \times 24.6 \times 19.9 \text{ \AA}^3$, respectively) are well covered by the final focused grid box. To examine if the results can be improved by decreasing the spacing h of the finest focusing grid and decreasing the `etol` parameter, we compare the interaction energies, calculated for h equal to $0.25 \text{ \AA}$, $0.2 \text{ \AA}$, $0.15 \text{ \AA}$ and `etol` $1e-6$ and $1e-9$. For every case, the sensitivity of the results to the grid centre position was also analysed in the same way as in the previous case of charged dyes. For both frames, the results computed with `etol` values $1e-6$ and $1e-9$ do not differ significantly and so in what follows we will present only the latter results.

The interaction energies for frames 6 and 10 are analyzed in Tables 7 and 8, respectively (for more details see Tables S10 and S11 in the SI). Like in the case of charged dyes, the values of E_{1-2}^{APBS} and E_{2-1}^{APBS} for all the tested grid spacings do not exhibit the required symmetry $E_{1-2} = E_{2-1}$. The effect is even more pronounced for particular relative positions of the molecular system and the grid, when E_{1-2}^{APBS} and E_{2-1}^{APBS} can have different sign (see the SI, grid centering perturbation $i = 10$ in Table S10 for $h = 0.2 \text{ \AA}$ of frame 6). Moreover, the asymmetry of the solutions, as well as the standard deviations σ_{1-2} and σ_{2-1} for some frames, do not scale down as $h \rightarrow 0$. The relative deviations D_{AA} show a slow convergence to the ARGOS solution. Even at the smallest grid spacing the variation of computed energies with respect to small perturbation of the grid position remains substantial. On the other hand, the convergence also indicates that the ARGOS solution is indeed accurate. Similar to the case of charged dyes, our analysis demonstrates that one calculation per frame is in general not sufficient to obtain a reliable result. Instead, an averaged value, obtained from a series of calculations with variations of the grid centering and grid spacing h is required. This example also shows that if one wishes to use a finer final focused grid, it is required to considerably increase the values of the `dime` parameter of APBS (see Table S10 in the SI) which, on the other hand, increases the computational costs and even more so the memory requirements.

Both FD solvers do not ensure the natural symmetry condition. Reducing the grid spacing helps to improve the accuracy, but not to the level obtained for the charged chromophores. In contrast to the latter, in the present case, the leading term in the interactions is due to dipoles. The electrostatic interaction strongly depends on their mutual orientation. This determines the high variation in the interaction energies for the neutral dyes (in models DYES and DYES-POLYPRO, Figure 16), spanning a range of about 3 orders of magnitude, in contrast to the case of the charged dyes. The overall screening effect in the present case is drastically diminished, compared to the case of charged

Table 7: Sensitivity of the APBS results to the grid spacing h and relative position between grid and the molecular system for **frame 6** in model **DYES-POLYPRO-IONS** with **neutral ground state charge densities of the chromophores**. σ_{1-2} and σ_{2-1} are standard deviations of interaction energies obtained for different grid positions. All energies in units of $e_0^2/\text{\AA}$.

h (Å)	$\bar{E}_{1-2}^{APBS}$	σ_{1-2}	$\bar{E}_{2-1}^{APBS}$	σ_{2-1}	Avg. $\bar{E}^{APBS}$	$D_{AA}(\%)$
0.25	1.74e-06	6.55e-07	1.79e-06	6.76e-07	1.77e-06	83.47
0.2	1.17e-06	5.73e-07	1.27e-06	4.98e-07	1.22e-06	26.69
0.15	7.88e-07	2.32e-07	8.48e-07	2.21e-07	8.18e-07	15.00

Table 8: Same as in Table 7, but for **frame 10** in model **DYES**.

h (Å)	$\bar{E}_{1-2}^{APBS}$	σ_{1-2}	$\bar{E}_{2-1}^{APBS}$	σ_{2-1}	Avg. $\bar{E}^{APBS}$	$D_{AA}(\%)$
0.25	2.23e-07	3.40e-08	2.71e-07	3.66e-08	2.47e-07	11.50
0.2	2.62e-07	2.43e-08	2.65e-07	6.28e-08	2.63e-07	5.74

chromophores. The maximum screening for the present case corresponds to a factor of 10^{-3} (see frame 12, Figure 16), while for the charged chromophores the maximum factor is about 10^{-6} (see frame 12, Figure 14). In contrast to the case of charged dyes, where the DYES-POLYPRO-IONS model exhibits the largest errors for almost all frames, here, the highest D_{AA} and D_{AM} are obtained for other models for half of the frames.

Influence of the charge distribution and presence of the interface

Our goal here is to quantify the influence of the charge distribution and the jump in the dielectric coefficient across the solute-solvent interface on the accuracy of the FD solvers MEAD and APBS. While such an investigation was already performed on academic problems with known solutions (Models-1, -2, and -3 in Sections 3.1.1, 3.1.2, and 3.1.3), here we consider the interaction energies in the DYES model with different charge densities and different sharpness of the solute-solvent interface. The latter is characterized by the magnitude of the jump of the dielectric coefficient across the interface, i.e., $|\epsilon_s - \epsilon_m|$.

In Table 9, we summarize five setups for the DYES model with different dielectric coefficients ϵ_m and ϵ_s where the Alexa chromophores are prescribed neutral ground state or transition charges (see Tables S1 and S2 for the transition charges). Setups 1 and 4 were already employed in Sections 3.2.2 and 3.1.1, respectively, while Setups 2, 3, and 5 are new. In Setup 5 we use transition charges and dielectric coefficients $\epsilon_m = 1$ for the dyes and $\epsilon_s = 2$ for the solvent. This case is relevant for Förster resonance energy transfer (FRET) between the chromophores.^{37,59} In FRET, only the electronic part of the environmental polarization, described by an optical dielectric coefficient of 2 (typical for organic matter), needs to be taken into account.⁵⁹ While Setup 5 is relevant for FRET and Setup 1 is needed for an implicit solvent description, Setups 2-4 just serve as test cases for the numerical solution, without any direct physical application.

Since there are no analytical solutions available for this model except for Setup 4, to assess the accuracy of MEAD and APBS, we use the reference solutions obtained with ARGOS at $MRL = 4$ and a geometric approximation of the interface $GridScale = 4$ in NanoShaper.

The relative deviations of the interaction energies obtained with MEAD and APBS from the ARGOS reference

Table 9: Setups for calculations of interaction energies between chromophores in charge-neutral state, using different dielectric coefficients ϵ_m and ϵ_s and different atomic partial charges.

Setup	ϵ_m	ϵ_s	charge type
1	4	80	ground state
2	4	80	transition
3	1	2	ground state
4	1	1	ground state
5	1	2	transition

solutions for Setups 1-4 are presented in Figure 18. The MEAD results reveal that removing the interface, passing from Setup 1 to Setup 3 and Setup 4, gradually reduces the relative deviations D_{AM} from ARGOS for almost all frames, indicating a strong influence of the jump in the dielectric coefficient on the numerical accuracy.

The APBS results exhibit a higher accuracy than the MEAD calculations with D_{AA} values not exceeding 22%, but the same general dependence on the "sharpness" of the interface. The results for frames 14 and 19 depart from this behaviour, revealing a noticeable increase of the errors in Setup 3, compared to Setup 1: D_{AA} rises from 6% to 22% and from 7% to 9% for frames 14 and 19, respectively. The accuracy for those two frames in Setup-3 can be improved by reducing the grid spacing and averaging the interaction energy over different relative positions of the grid and the molecular system, like above. This procedure reduces the deviation D_{AA} to 3% for frame 14 and to 1% for frame 19 (see details in the SI, Tables S13 and S15), restoring the correct trend and demonstrating again the strong sensitivity of the APBS results on the relative position of the grid.

Application of the transition charges in Setup 2 leads to smaller relative deviations D_{AM} and D_{AA} in the interaction energies for all frames. However, the deviations can still exceed 10% for both MEAD and APBS. The one noteworthy exception here is frame 1 which shows the opposite trend: for Setup 2 with transition charges D_{AA} increases almost 10 times, compared to Setup 1 with ground-state charges. Calculations of averaged $\bar{E}^{APBS}$ on grids with smaller spacing h (see Table S16 in the SI) cut down the relative deviation D_{AA} between APBS and ARGOS to 0.5%, which is again an evidence that the ARGOS reference solutions are indeed accurate.

The smaller relative deviations obtained for the coupling between transition charges can be explained in the following way: in the case of transition charges the potential is created by a smaller number of nonzero charges (see Tables S3 and S4, and S1 and S2 for the ground-state and transition charges, respectively, for the two dyes) compared to the case of ground-state charges, and then multiplied again by a smaller number of nonzero charges³ to obtain the interaction energy.

In order to describe a situation relevant for FRET the dielectric coefficients ϵ_m and ϵ_s of the dyes with transition charges and the solvent, respectively, have to be decreased from $\epsilon_m = 4$ and $\epsilon_s = 80$ in Setup 2, considered so far, to $\epsilon_m = 1$ and $\epsilon_s = 2$ (Setup 5 in Table 9). The electrostatic interaction energies in this case have been assessed with MEAD in our previous publication 37. Here, we reinvestigate these results for a large number of frames of POLY6

³Here by a smaller number of nonzero charges we also refer to the fact that many transition charges in both dyes are either zero or negligible in magnitude relative to a few nonzero ones.

(Figure S23). For most frames of POLY6, the relative deviation of MEAD from ARGOS stays below 3%, while for some the error can reach up to 20%. The large deviations, however, occur only for small couplings and, therefore, do not affect the FRET efficiencies calculated previously.³⁷

Hence, it is not the calculation of the coupling between transition densities but that between the ground state charge densities of the chromophores that can be critically improved by ARGOS. The latter aspect will be crucial in an implicit description of the solvent in MD simulations that will allow one to simulate much larger time scales than by using an explicit description. In this way the sampling of chromophore configurations can be improved which is critical for the interpretation of FRET experiments. An important aspect of an implicit solvent description concerns the screening of the interchromophore coupling by the ions, which is investigated in the following.

Screening by ions

In order to investigate the ionic screening effects in more detail, we applied ARGOS to calculate the reference interaction energies for all interface models. Two additional models without dielectric boundaries ($\epsilon_m = \epsilon_s = 80$), but including the ion exclusion layer (IEL) around different parts of the molecular system, were introduced:

- DYES-IONS includes ionic charge density in the aqueous solvent, characterized by the same ionic strength of 0.3 M as before and the IEL around the chromophores (as in Figure 1);
- DYES-POLYPRO-IONS-2 includes ionic charge density in the solvent with ionic strength of 0.3 M and the IEL around the chromophores and polyproline helix.

Please note that, the ions are not allowed to penetrate into the region surrounded by the IEL.

The model NO-INT (no interface, $\epsilon_m = \epsilon_s = 80$, see Section 3.2) differs from the model DYES-IONS only by the absence of ions. In DYES-IONS, ions penetrate in the area between the dyes and this ionic screening decreases the electrostatic inter-chromophore coupling (as seen in Figure 19 by comparing the stars with the rombs). The screening is particularly effective for POLY20 (frames 11 to 19) because of the large inter-chromophore distances. For POLY6, in three cases (frames 1, 3 and 8) the dyes are so close that the ionic screening is practically zero. The molecular configuration of frame 1 in Figure 17a shows that the molecular cavities of the two dyes start to overlap and, hence, there is no space for the ions.

Exclusion of the ions from the polyproline domain, while passing from the model DYES-IONS to DYES-POLYPRO-IONS, can substantially change the interaction energy for both POLY6 and POLY20. The interaction energy is enhanced for frames 4, 5, 7-9, 12, 14, 17, 19, decreased for frames 2, 11, 13, 16, 18 and left practically unchanged for the remaining frames (as seen by comparing the crosses with the rombs in Figure 19). The effect depends mostly on the orientation of the dyes with respect to the polyproline helix. In the closed conformation, when the chromophore with a long linker (Alexa 488) is close to the helix (as in frames 2 and 11 shown in Figures 17b and 17c, respectively), the dipole of the chromophore is not completely surrounded by the ions. This configuration results in partial ionic screening of the charges that changes the magnitude and orientation of the effective dipole of the chromophores, modifying the interaction between the dyes. The rotation of the effective dipoles (frames 11, 13, 18) can even change the sign of the interaction. On the other hand, the open configuration (see Figure 17d) experiences a complete screening by the

ions. In this case, the effective dipole is not rotated but just decreased in magnitude, and excluding the ions, therefore, leads to an increase of the interaction.

Similar effects are observed while passing from the model DYES-POLYPRO to DYES-POLYPRO-IONS. Here, the presence of the ionic charge density in the solvent leads to a change of the sign of the interaction energy for some frames in comparison to the model DYES-POLYPRO (compare triangles with squares for frames 6, 12, 14, 16, 19 in Figure 19). As before, the presence of the ionic charge density in the solvent does not always lower the magnitude of the interaction energies. For example, the interaction energy is not changed for frames 1, 3, 8, 9 of POLY6, while frame 2 shows an enhanced interaction between the chromophores.

Interestingly, analysis of Figures 15 and 19 shows that the highest D_{AM} and D_{AA} (for the model DYES-POLYPRO-IONS) correspond to frame 6, in spite of the fact that the interaction energy is not the smallest in the set. Moreover, this coupling is of the same order of magnitude as the interaction energies obtained for frames 4 and 5. The latter, however, exhibit much lower relative deviations for both solvers MEAD and APBS. As it was shown above, the APBS relative deviation D_{AA} in the calculated energy for frame 6 can be reduced down to 15% by reducing the grid spacing h and taking averages over different relative positions of the grid and the molecular system. However, this percentage is still much larger than the relative deviation D_{AA} between the APBS calculations and the reference solutions of ARGOS for frames 4 and 5. The difference in the relative deviations between frames 6 and 4 (or 5) is most probably due to the particular geometry of the solute-solvent interface and the distribution of the charges near them, resulting in the poor performance of both FD solvers in the case of frame 6.

The above screening effects will be of particular interest for the development and applications of coarse-grained models of biomolecules, which use an implicit representation of the solvent (water and ions) mentioned already above. In the common approaches the charges are screened using either the Debye-Hückel theory or generalized Born models. The first method⁶⁰ describes ionic screening of electrostatic interactions, essentially using equation (54), though this expression is accurate only for a spherical charged particle, completely surrounded by the solvent. The authors of Ref. 61 propose to model partial screening in the generalized Born model with modified atomic radii. The latter are parameterized to characterize the extent of the screening. However, the parameterization in this approach was performed so far only for the dielectric screening without including the ionic charge density in the solvent. The present results demonstrate that both conditions - partial screening due to the particular geometry of the solute and the nonzero ionic strength of the solvent - are important for the interaction energy. With calculations as shown above a reliable parameterization of ionic screening is within reach.

3.3 CPU times

Although ARGOS³⁶ is currently in a prototype version (implemented in FreeFem⁴⁸) and is currently undergoing a substantial performance improvement, we present first CPU times in comparison to APBS obtained on a machine with Intel(R) Core(TM) i5-7500 CPU @ 3.40GHz and 16 GB of RAM. Since the primal and adjoint problems are discretized using the linear and quadratic Lagrange finite elements, respectively, the number of DOFs, N_{DOFs}^{Adj} , in 3D for the adjoint problem is approximately 8 times larger than that for the primal. Here, we note that for a Dirichlet BC the number of DOFs for the FE space V_h^1 , N_{DOFs}^{Prim} , is exactly equal to $N_V - N_{\partial\Omega}$, where $N_{\partial\Omega}$ is the number of mesh vertices on the boundary. Consequently, the major part of the CPU

Figure 17: Molecular geometries of POLY6 (frames 1 and 2) and POLY20 (frames 11 and 12), labelled with Alexa chromophores, obtained from snapshots of MD simulations.

Figure 18: Relative deviations between interaction energies of the chromophores in a charge-neutral ground state, obtained with APBS (blue symbols) and MEAD (green symbols). Calculations are done for the DYES model and different setups (Table 9), using APBS configuration A1 from Table 2 and MEAD configuration M1-2 from Table 1. The ARGOS reference interaction energies are obtained with *GridScale* = 4 at MRL = 4. The different frames correspond to different molecular geometries taken from MD simulations.

Figure 19: Interaction energies of the chromophores in a charge-neutral ground state, calculated by ARGOS with *GridScale* = 2 at MRL = 4 for different interface models. The blue and red colours of the symbols correspond to positive and negative interaction energies, respectively. The different frames correspond to different molecular geometries taken from MD simulations.

time per MRL for ARGOS is spent for the solution of the discrete adjoint problem and the refinement of the mesh for the next MRL. In all the presented examples, we use a basic Jacobi-preconditioned conjugate gradient (CG) method (with a stopping criterion ensuring a very high precision⁴)

⁴The reference solutions of ARGOS presented in the paper are obtained with an error tolerance between $1e-14$ and $1e-12$ for the residual in the linear system, measured in the norm generated by the inverse of the matrix C , where C is a diagonal matrix whose

as an iterative solver for the resulting linear systems for, both, the primal and adjoint problem. As it was mentioned above, the mesh refining is done using mmg3d v4.0,⁴⁷ but a newer version of it can also be used. Any other remeshing software can be utilized as well, which could drastically reduce the CPU time for the mesh refining steps. In Tables 10 and 11 we present the CPU times for some examples with an increasing number of vertices in the initial mesh. We note that the reference solutions of ARGOS given above are computed on meshes with uniform mesh size in the molecular regions, whereas the results below are obtained with coarser initial meshes of lower quality (the molecular surfaces are still the same). Nevertheless, the computed values for the interaction energy are practically the same. Below, t_{MRL} , t_{mat}^{Adj} , t_{CG}^{Adj} , t_{ref} denote the total CPU time per MRL, the CPU time to assemble the matrix for the adjoint problem, the CPU time to solve the adjoint problem, and the CPU time to refine the mesh, respectively.

As a first example, in Table 10, we consider Model-2 where the probe charge is placed at a distance $r = 10$ Å from the center of the Born ion. The total execution time for ARGOS with *GridScale* = 2 and a CG tolerance of $1e-09$ is 46.9 sec (the corresponding time for a CG tolerance of $1e-12$ is 51.14 sec., see footnote 4). In comparison, APBS with configuration A2 (see Table 4) for **etol** $1e-06$ takes 658 sec on average. Note that from this one computation with APBS one can extract the interaction energies for all distances $r \leq 48$ Å. However, the size of the final focused grid does not allow to place the test charge at a larger distance, whereas with ARGOS the size of the molecular system is not an issue. A further comparison of the CPU times for *GridScale* = 2 and 4 in ARGOS is given in Tables S17 and S18 in the SI.

In Table 11 we give the CPU times of ARGOS for frame 10 of model DYES (Figure 11a) with neutral ground-state charges (see Figure 18). The total execution time for *GridScale* = 2 and with a CG tolerance of $1e-09$ is 99 sec. compared to 109 and 143 sec. of APBS for configuration A1 (see Table 2) with **etol** $1e-06$ and **etol** $1e-09$, respectively. However, those CPU times of APBS are only for the computation of E_{1-2}^{APBS} , which alone is not enough to obtain a somehow reliable interaction energy and do not include the time of 17 sec. needed by the plugin *multival* to read each of the potential “dx” files and return the potential at the atom’s centres. Taking only E_{1-2}^{APBS} or E_{2-1}^{APBS} can result in a very large error (see Table S11 in the SI, together with the other tables concerning the inspection of problematic APBS frames). Therefore, to compute $\bar{E}^{APBS}$ (see (51)), it takes in total 218 and 286 sec. for **etol** $1e-06$ and **etol** $1e-09$, respectively.

Finally, we report the CPU times for frame 12 of model DYES-POLYPRO-IONS in the case of charged dyes (see Table 5 and Table S9 in the SI). ARGOS with a CG tolerance of $1e-09$ and three MRLs takes 616 sec. compared to 837 sec. necessary to compute E_{1-2}^{APBS} with configuration A1 for **etol** $1e-09$ (recall that $1e-06$ is not enough for this frame), and 747 sec. for **etol** $1e-06$. Again, those times do

diagonal is equal to the diagonal of the system’s matrix.

not include the time of around 32 sec. spent by *multivalue* to read the “.dx” file for each run and return the potential at the specified atom’s centres. If the CG tolerance in ARGOS is decreased to $1e-12$, then the time needed increases to 1114 sec., even though the resulting interaction energy is practically the same. However, we note that exchanging the Jacobi-preconditioned CG algorithm for an advanced iterative method, for example a CG with an appropriate preconditioner or a multigrid method, will dramatically improve the performance of ARGOS, which is already higher now than that of APBS.

Table 10: CPU times of ARGOS (including the generation of the initial mesh with NanoShaper and TetGen) for Model-2 obtained for different MRLs. The probe charge is placed at a distance $r = 10$ Å from the center of the Born ion. $GridScale = 2$, a CG tolerance of $1e-09$, and a homogeneous Dirichlet boundary condition were used. Err^{ARGOS} denotes the relative error (Eq. (52)) and t_{MRL} , t_{mat}^{Adj} , t_{CG}^{Adj} , and t_{ref} are the total CPU time per MRL, the CPU time to assemble the matrix for the adjoint problem, the CPU time to solve the adjoint problem, and the CPU time to refine the mesh, respectively, all times given in units of seconds.

MRL	N_V	N_{DOFs}^{Adj}	Err^{ARGOS} (%)	$t_{MRL}; \{t_{mat}^{Adj}/t_{CG}^{Adj}/t_{ref}\}$
0	2437	18201	2.5353	3.79; {0.19/0.24/3.27}
1	2900	21811	0.4458	4.10; {0.22/0.26/3.37}
2	4447	34038	0.1274	4.97; {0.38/0.46/3.75}
3	7879	61117	0.0424	7.32; {0.74/1.09/4.80}
4	15306	120060	0.0148	13.28; {1.63/2.99/7.21}
5	28711	226208	0.0067	13.44; {3.35/8.22/0.00}

Table 11: CPU times of ARGOS (including the generation of the initial mesh with NanoShaper and TetGen) in seconds (for definition of the different times see Table 10) for frame 10 of model DYES (Figure 11a) with neutral ground-state charges, $GridScale = 2$ and a CG tolerance of $1e-09$. The values of E_{1-2}^{Adj} for MRL = 4 and 5 are $2.78e-07$ and $2.79e-07$, respectively. The latter value coincides with the value reported in Table S11, obtained from a better quality initial mesh.

MRL	N_V	N_{DOFs}^{Adj}	E_{1-2}^{Adj} ($e_0^2/\text{\AA}$)	$t_{MRL}; \{t_{mat}^{Adj}/t_{CG}^{Adj}/t_{ref}\}$
0	14684	111448	2.75e-07	17.42; {1.64/3.29/5.89}
1	19829	152680	2.73e-07	16.49; {2.24/4.33/7.73}
2	31378	244342	2.78e-07	29.94; {3.74/8.37/14.14}
3	55058	431526	2.77e-07	35.17; {6.89/23.93/0.00}

4 Conclusions

A new finite element method, termed ARGOS, for the solution of the linearized Poisson-Boltzmann equation is formulated, implemented and tested. The advantages of the method include a correct handling of the interfaces and an appropriate error estimation technique for the target quantity (the electrostatic interaction energy), which leads to an efficient mesh adaptation and an efficient numerical solution.

Applications of ARGOS to model systems, for which analytical solutions exist, demonstrate the high accuracy of this solver. This high accuracy was used to evaluate the performance of FD solvers in an application to chromophore-labelled polyproline helices. A critical source of errors concerns complex geometries of interfaces, especially in the presence of ionic screening. An accurate description of such

scenarios will be important for the development of implicit-solvent models. Moreover, if the motion of solvent molecules is restrained, a spatially varying dielectric coefficient would result in the solvent domain. The present solver is ready to take this dependence into account. We note that in this case, the jump in the dielectric coefficient between solute and solvent is less than considered in the present models. Nevertheless, a quantitative evaluation of these effects requires an accurate solver as the present one.

The new method exhibits a uniform and fast convergence of the solution with respect to the number of degrees of freedom. Comparison with the well-known finite difference solvers MEAD and APBS demonstrates the superior performance of ARGOS and reveals critical inaccuracies of MEAD and APBS, originating from both numerical implementation and general problems of the FD methods. In particular, the accuracy of the MEAD results is much worse in the presence of interfaces with large changes in the dielectric coefficient, a common problem of FD approaches. In addition, we find that inappropriate stopping criteria in the iterative solver (successive over relaxation (SOR)) in combination with single precision arithmetics can significantly influence the accuracy of MEAD calculations.

Most of the time, APBS performs much better than MEAD. However, it still exhibits the common fundamental issues typical for FD based solvers: a trade off between grid spacing h and the dimensions of the final focused grid, a strong dependence of the results on the relative position of the focused grids and the molecular system, as well as of the grid spacing h , which in some cases (see, e.g., frames 6, 10, 12 in Sections 3.2.2 and 3.2.1) requires a thorough inspection of the calculated values to detect inaccurate results. A typical inspection includes the recalculation with perturbation in the focused grids’ centering and/or with slight changes in the grid spacing h , comparison of the two values E_{1-2}^{APBS} and E_{2-1}^{APBS} (or E_{1-2}^{MEAD} and E_{2-1}^{MEAD} for MEAD) and looking for more pronounced violation of the symmetry condition $E_{1-2} = E_{2-1}$. All this extra work required to ensure reliable results takes a lot of computational and post processing time (see, e.g., the inspection of frames 6, 10, and 12 in Sections 3.2.2 and 3.2.1), which also provokes uncertainty in users who are unexperienced in numerical methods.

Due to the use of an appropriate (goal-oriented) error estimation in combination with an adaptive mesh generation the present FE solver requires much less degrees of freedom than MEAD or APBS to obtain a comparable or better accuracy. This advantage, if complemented with parallelization techniques, will make ARGOS a very fast and reliable PBE solver. First benchmark studies show that ARGOS already in its present non-optimized version is faster and more accurate than the standard FD solvers MEAD and APBS.

Acknowledgement This work was supported by the Doctorate College program “Nano-Analytics of Cellular Systems (NanoCell)” of the Austrian Science Fund (FWF) (grant number: W 1250), project LIT-2017-4-SEE-004, project P 33154-B, funded by the FWF in conjunction with the district of Upper Austria and the grant LIT-2019-8-SEE-120.

References

- (1) Brovchenko, I.; Oleinikova, A. Which Properties of a Spanning Network of Hydration Water Enable Biological Functions? *Chem. Phys. Chem* **2008**, *9*, 2695–2702.
- (2) Chen, S.-H.; Liu, L.; Fratini, E.; Baglioni, P.; Faraone, A.; Mamontov, E. Observation of fragile-to-

- strong dynamic crossover in protein hydration water. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9012–9016.
- (3) De Simone, A.; Dodson, G.; Verma, C.; Zagari, A.; Fraternali, F. Prion and water: tight and dynamical hydration sites have a key role in structural stability. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7535–7540.
 - (4) Fogolari, F.; Brigo, A.; Molinari, H. The Poisson-Boltzmann Equation for biomolecular electrostatics: a tool for structural biology. *J. Mol. Recognit.* **2002**, *15*, 377–392.
 - (5) Vascon, F.; Gasparotto, M.; Giacomello, M.; Cendron, L.; Bergantino, E.; Filippini, F.; Righetto, I. Protein electrostatics: From computational and structural analysis to discovery of functional fingerprints and biotechnological design. *Computational and Structural Biotechnology Journal* **2020**, *18*, 1774 – 1789.
 - (6) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effects of ionic strength and amino-acid modification. *Proteins* **1986**, *1*, 47–59.
 - (7) Nicholls, A.; Sharp, K.; Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **1991**, *11*, 281–296.
 - (8) Bashford, D. An Object-Oriented Programming Suite for Electrostatic Effects in Biological Molecules. Proceedings of the Scientific Computing in Object-Oriented Parallel Environments. London, UK, 1997; pp 233–240.
 - (9) Madura, J.; Briggs, J.; Wade, R.; Davis, M.; Luty, B.; Ilin, A.; Antosiewicz, J.; Gilson, M.; Bagheri, B.; Scott, L.; McCammon, J. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Comp. Phys. Comm.* **1995**, *91*, 57 – 95.
 - (10) Jo, S.; Vargyas, M.; Vasko-Szedlar, J.; Roux, B.; Im, W. PBEQ-Solver for online visualization of electrostatic potential of biomolecules. *Nucleic Acids Res.* **2008**, *36*, W270–W275.
 - (11) Holst, M.; Baker, N.; Wang, F. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comput. Chem.* **2000**, *21*, 1319–1342.
 - (12) Gilson, M. K.; Sharp, K. A.; Honig, B. H. Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment. *J. Comput. Chem.* **1988**, *9*, 327–335.
 - (13) Baker, N.; Sept, D.; Joseph, S.; Holst, M.; McCammon, J. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041.
 - (14) Rogers, N.; Sternberg, M. Electrostatic interactions in globular proteins: Different dielectric models applied to the packing of α -helices. *Mol. Biol.* **1984**, *174*, 527 – 542.
 - (15) Sakalli, I.; Schöberl, J.; Knapp, E. W. mFES: A Robust Molecular Finite Element Solver for Electrostatic Energy Computations. *J. Chem. Theory and Comput.* **2014**, *10*, 5095–5112.
 - (16) Bashford, D. Macroscopic electrostatic models for protonation states in proteins. *Frontiers in Bioscience* **2004**, *9*, 1082–1099.
 - (17) Zhou, Z.; Payne, P.; Vasquez, M.; Kuhn, N.; Levitt, M. Finite-difference solution of the Poisson-Boltzmann equation: Complete elimination of self-energy. *J. Comput. Chem.* **1996**, *17*, 1344–1351.
 - (18) Niedermeier, C.; Schulten, K. Molecular dynamics simulations in heterogeneous dielectrics and Debye-Huckel media-application to the protein bovine pancreatic trypsin inhibitor. *Mol. Simul.* **1992**, *8*, 361–387.
 - (19) Froese, B.; Salvador, T. Higher-order adaptive finite difference methods for fully nonlinear elliptic equations. *J. Sci. Comput.* **2018**, *75*, 1282–1306.
 - (20) Oberman, A.; Zwiers, I. Adaptive finite difference methods for nonlinear elliptic and parabolic partial differential equations with free boundaries. *J. Sci. Comput.* **2016**, *68*, 231–251.
 - (21) Mirzadeh, M.; Theillard, M.; Gibou, F. A second-order discretization of the nonlinear Poisson-Boltzmann equation over irregular geometries using non-graded adaptive Cartesian grids. *J. Comput. Phys.* **2011**, *230*, 2125–2140.
 - (22) Harris, R.; Boschitsch, A.; Fenley, M. Numerical Difficulties Computing Electrostatic Potentials Near Interfaces with the Poisson-Boltzmann Equation. *J. Chem. Theory and Comput.* **2017**, *13*, 3945–3951, PMID: 28640608.
 - (23) Boschitsch, A.; Fenley, M. A Fast and Robust Poisson-Boltzmann Solver Based on Adaptive Cartesian Grids. *J. Chem. Theory and Comput.* **2011**, *7*, 1524–1540, PMID: 21984876.
 - (24) Zauhar, R.; Morgan, R. A new method for computing the macromolecular electric potential. *Mol. Biol.* **1985**, *186*, 815 – 820.
 - (25) Rashin, A.; Namboodiri, K. A simple method for the calculation of hydration enthalpies of polar molecules with arbitrary shapes. *J. Phys. Chem.* **1987**, *91*, 6003–6012.
 - (26) Zauhar, R.; Morgan, R. The rigorous computation of the molecular electric potential. *J. Comput. Chem.* **1988**, *9*, 171–187.
 - (27) Lu, B.; McCammon, J. Improved Boundary Element Methods for Poisson-Boltzmann Electrostatic Potential and Force Calculations. *J. Chem. Theory and Comput.* **2007**, *3*, 1134–1142, PMID: 26627432.
 - (28) Fenley, M.; Boschitsch, A.; Zhou, H.-X. Fast Boundary Element Method for the Linear Poisson-Boltzmann Equation. *J. Phys. Chem. B* **2002**, *106*, 2741–2754.
 - (29) Lu, B.; Cheng, X.; Huang, J.; McCammon, J. Order N algorithm for computation of electrostatic interactions in biomolecular systems. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19314–19319.
 - (30) Lu, B.; Cheng, X.; Huang, J.; McCammon, J. An Adaptive Fast Multipole Boundary Element Method for Poisson-Boltzmann Electrostatics. *J. Chem. Theory and Comput.* **2009**, *5*, 1692–1699.

- (31) Orttung, W. Direct solution of the Poisson equation for biomolecules of arbitrary shape, polarizability density, and charge distribution. *Ann. N.Y. Acad. Sci.* **1977**, *303*, 22–37.
- (32) Cortis, C.; Friesner, R. Numerical solution of the Poisson-Boltzmann equation using tetrahedral finite-element meshes. *J. Comput. Chem.* **1997**, *18*, 1591–1608.
- (33) Holst, M. *The Poisson-Boltzmann Equation: Analysis and Multilevel Numerical Solution*; Applied Mathematics and CRPC, California Institute of Technology; 1994.
- (34) Ji, N.; Liu, T.; Xu, J.; Shen, L. Q.; Lu, B. A Finite Element Solution of Lateral Periodic Poisson-Boltzmann Model for Membrane Channel Proteins. *Int. J. Mol. Sci.* **2018**, *19*.
- (35) Bank, R.; Holst, M. A New Paradigm for Parallel Adaptive Meshing Algorithms. *SIAM J. Sci. Comp.* **2000**, *22*, 1411–1443.
- (36) Nakov, S. ARGOS (Adaptive Refinement Goal-Oriented Solver). <https://doi.org/10.5281/zenodo.4269441>, 2020.
- (37) Sobakinskaya, E.; Schmidt am Busch, M.; Renger, T. Theory of FRET "Spectroscopic Ruler" for Short Distances: Application to Polyproline. *J. Phys. Chem. B* **2018**, *112*, 54–67.
- (38) Nakov, S. The Poisson-Boltzmann Equation: Analysis, A Posteriori Error Estimates and Applications. Ph.D. thesis, Johannes Kepler University, 2019; Available at <https://www.numa.uni-linz.ac.at/Teaching/PhD/Finished/nakov>.
- (39) Iglesias, J. A.; Nakov, S. Weak formulations of the nonlinear Poisson-Boltzmann equation in biomolecular electrostatics. 2020; Preprint, arXiv:2012.06437.
- (40) Kraus, J.; Nakov, S.; Repin, S. Reliable Computer Simulation Methods for electrostatic Biomolecular Models Based on the Poisson-Boltzmann Equation. *Comput. Methods Appl. Math.* **2020**, *20*, 643–676.
- (41) Chen, L.; Holst, M. J.; Xu, J. Adaptive Finite Element Modeling Techniques for the Poisson-Boltzmann Equation. *Siam J. Numer. Anal.* **2007**, *45*, 2298–2320.
- (42) Holst, M.; McCammon, J.; Yu, Z.; Zhou, Y. C.; Zhu, Y. Adaptive Finite Element Modeling Techniques for the Poisson-Boltzmann Equation. *Commun. Comput. Phys.* **2012**, *11*, 179–214.
- (43) Becker, R.; Rannacher, R. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.* **2001**, *10*, 1–102.
- (44) Richter, T.; Wick, T. Variational localizations of the dual weighted residual estimator. *J. Comput. Appl. Math.* **2015**, *279*, 192–208.
- (45) Decherchi, S.; Rocchia, W. A general and Robust Ray-Casting-Based Algorithm for Triangulating Surfaces at the Nanoscale. *PLOS ONE* **2013**, *8*, 1–15.
- (46) Si, H. TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator. *ACM Transactions on Mathematical Software (TOMS)* **2015**, *41*.
- (47) Dobrzynski, C. *MMG3D: User Guide*; Technical Report RT-0422, 2012.
- (48) Hecht, F. New development in FreeFem++. *J. Numer. Math.* **2012**, *20*, 251–265.
- (49) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (50) Brooks, B. R.; Brooks III, C. L.; Mackerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kucsera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (51) Jorgensen, W. L.; Madura, J. D. Quantum and statistical mechanical studies of liquids. 25. Solvation and conformation of methanol in water. *J. Am. Chem. Soc.* **1983**, *105*, 1407–1413.
- (52) Childs, H.; Brugger, E.; Whitlock, B.; Meredith, J.; Ahern, S.; Pugmire, D.; Biagas, K.; Miller, M.; Harrison, C.; Weber, G. H.; Krishnan, H.; Fogal, T.; Sanderson, A.; Garth, C.; Bethel, E. W.; Camp, D.; Rübel, O.; Durant, M.; Favre, J. M.; Navrátil, P. *High Performance Visualization—Enabling Extreme-Scale Scientific Insight*; 2012; pp 357–372.
- (53) Andreussi, O.; Hörmann, N. G.; Nattino, F.; Fisi-caro, G.; Goedecker, S.; Marzari, N. Solvent-Aware Interfaces in Continuum Solvation. *J. Chem. Theory and Comput.* **2019**, *15*, 1996–2009.
- (54) Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallog.* **1983**, *16*, 548–558.
- (55) Richards, F. M. Areas, volumes, packing, and protein structure. *Ann. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176.
- (56) Greer, J.; Bush, B. L. Macromolecular shape and surface maps by solvent exclusion. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 303–307.
- (57) Sanner, M. F.; Olson, A. J.; Spehner, J.-C. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **1996**, *38*, 305–320.
- (58) Engel, T. In *Cheminformatics. Basic Concepts and Methods*; Engel, T., Gasteiger, J., Eds.; Wiley-VCH, 2018; Chapter 3, pp 103–106.
- (59) Renger, T.; Müh, F. Theory of excitonic couplings in dielectric media. *Photosynth. Res.* **2012**, *111*, 47–52.
- (60) Knotts, T.; Rathore, N.; Schwartz, D.; de Pablo, J. A coarse grain model for DNA. *J. Chem. Phys.* **2007**, *126*, 084901.
- (61) Marenich, A.; Cramer, C.; Truhlar, D. Universal Solvation Model Based on the Generalized Born Approximation with Asymmetric Descreening. *J. Chem. Theory and Comput.* **2009**, *5*, 2447–2464.