

Energetic Decomposition with the Generalized-Born and Poisson–Boltzmann Solvent Models: Lessons from Association of G-Protein Components

Noel Carrascal[†] and David F. Green^{*,†,‡}

Department of Applied Mathematics and Statistics and Graduate Program in Biochemistry and Structural Biology, Stony Brook University, Stony Brook, New York 11794-3600

Received: November 4, 2009; Revised Manuscript Received: March 8, 2010

Continuum electrostatic models have been shown to be powerful tools in providing insight into the energetics of biomolecular processes. While the Poisson–Boltzmann (PB) equation provides a theoretically rigorous approach to computing electrostatic free energies of solution in such a model, computational cost makes its use for large ensembles of states impractical. The generalized-Born (GB) approximation provides a much faster alternative, although with a weaker theoretical framework. While much attention has been given to how GB recapitulates PB energetics for the overall stability of a biomolecule or the affinity of a complex, little attention has been given to how the contributions of individual functional groups are captured by the two methods. Accurately capturing these individual electrostatic components is essential both for the development of a mechanistic understanding of biomolecular processes and for the design of variant sequences and structures with desired properties. Here, we present a detailed comparison of the group-wise decomposition of both PB and GB electrostatic free energies of binding, using association of various components of the heterotrimeric-G-protein complex as a model. We find that, while net binding free energies are strongly correlated in the two models, the correlations of individual group contributions are highly variable; in some cases, strong correlation is seen, while in others, there is essentially none. Structurally, the GB model seems to capture the magnitude of direct, short-range electrostatic interactions quite well but performs more poorly with moderate-range “action-at-a-distance” interactions—GB has a tendency to overestimate solvent screening over moderate distances, and to underestimate the costs of desolvating charged groups somewhat removed from the binding interface. Despite this, however, GB does seem to be quite effective as a predictor of those groups that will be computed to be most significant in a PB-based model.

Introduction

The interactions between biological molecules lie at the heart of the majority of fundamental biological processes. As a result, a deep understanding of these interactions is an essential component of understanding biology as a whole. The mechanisms by which extracellular signals are transduced into intracellular changes in cellular behavior are a perfect example of this, being driven by cascades of interactions between various signaling molecules.

One of the prototypical models for signal transduction is the heterotrimeric-G-protein signal transduction pathway.^{1–3} Helical, integral membrane proteins (the G-protein-coupled receptors, GPCRs) interact with a diverse range of extracellular ligands. Intracellularly, the GPCRs interact with heterotrimeric G-proteins, which consist of three subunits (α , β , and γ). In the inactive state, the G-proteins are in a trimeric form, with the α -subunit additionally associated with guanosine-5'-diphosphate (GDP). Upon receptor activation (by ligand binding), the receptor promotes exchange of the GDP bound to G_α with guanosine-5'-triphosphate (GTP); this leads to a conformational change in G_α that results in dissociation from the β - and γ -subunits (which remain complexed). Both G_α and $G_{\beta\gamma}$ are then free to interact with additional cellular-signaling proteins;

different variants of the G-protein subunits have different secondary targets.

In mammalian genomes, there are multiple variants of each subunit—at least 20 α (16 well-characterized, and several putative forms that are less or uncharacterized), 7 β (including multiple isoforms from individual genes), and 12 γ variants;⁴ it is the appropriate combination of these which leads to the correct coupling of a specific GPCR to the cognate cellular response. While a great deal of work has been done to decode the specific interactions made between various G-protein subunits, there remain many open questions regarding which combinations of subunits are possible, as well as which are physiologically relevant.^{5–9} A detailed understanding of the energetics underlying subunit association could provide important insight into this problem.

The association of biological molecules typically occurs in an aqueous environment of moderate ionic strength; thus, solute–solvent interactions must be appropriately treated in any energetic model. Fully explicit, all-atom simulations with the biomolecule embedded in a box of water molecules and ions, with periodic-boundary conditions to reduce boundary effects, have been very successful in providing a deep understanding of biomolecular structure and dynamics in solution.^{10–12} However, while well-suited to determining the dynamic, configurational ensembles of a single species (or complex), many challenges arise in computing free-energy differences between the bound and unbound states of a complex. In particular, accurate estimations of free energies in explicit solvent require adequate sampling of solvent degrees of freedom; additional

* To whom correspondence should be addressed. Phone: (631) 632-9344. Fax: (631) 632-8490. E-mail: david.green@stonybrook.edu.

[†] Department of Applied Mathematics and Statistics.

[‡] Graduate Program in Biochemistry and Structural Biology.

difficulties arise in comparing two independent simulations where the number of explicit solvent molecules may vary. Accurate calculations of free-energy differences in solution can be achieved through free-energy perturbation techniques, but these approaches are tremendously costly.^{13,14}

It has long been recognized that a powerful alternative to explicit-solvent free-energy perturbation approaches to computing solute–solvent interaction free energies is the use of continuum solvent models.^{15,16} Representing the interior of a biomolecule as a region of low dielectric (typically between 1 and 8) and solvent as a region of high dielectric (roughly 80 for water at room temperature) with a Debye–Hückel-like treatment of mobile ions, the electrostatic potential of a molecule, and thus its free energy, can be computed by solution of the Poisson–Boltzmann (PB) equation.^{17,18} PB-based continuum solvent models have been used to provide important insights into many aspects of molecular recognition: understanding the optimum of enzymes for substrate or transition-state binding;^{19–21} dissecting the contribution of various groups to binding affinity;^{22,23} predicting the relative affinities of ligands;^{24–26} understanding protein–membrane association;²⁷ and applications to the design of protein complexes.^{28–30}

Over the past decade, there has been an increasing interest in faster alternatives to PB-based models for use in applications that require many repeated evaluations, such as high-throughput virtual screening of potential drug molecules and molecular dynamics simulations. One of the most commonly used of these is the generalized-Born (GB) model originally proposed by Still and co-workers.^{31–33} Several implementations of the GB model have been applied to many problems with a great deal of success, but some concerns have also been raised.^{34–39} For example, Brooks and co-workers have noted an imbalance between solute–solvent and intersolute interactions,⁴⁰ and Onufriev et al. have found that GB-based methods can be sensitive to the method by which effective Born radii are computed.⁴¹

One of the great advantages of Poisson–Boltzmann-based models is that the free-energies are, in the context of the linearized-PB equation, pairwise decomposable in terms of individual partial atomic charges. This allows for a rigorous decomposition of contributions from individual chemical groups, such as the amino-acid side chains of a protein; such a detailed understanding of the role of particular residues in contributing to protein stability and complex affinity has been shown to be an important tool in the engineering of proteins with enhanced properties.^{42–44} Additionally, the pairwise decomposability of PB-based electrostatics has led to an elegant optimization theory, through which the electrostatic interactions in a complex may be tuned to provide a provably optimal contribution to binding.^{45–47}

Generalized-Born models are equally pairwise decomposable, and thus should be applicable in the same framework that has made PB-based analysis so powerful. However, the ability of GB-based models to recapitulate PB-based energetics at this level of detail has not been extensively explored. Here, we present results from the application of both PB- and GB-based component analysis to the interactions between various components of the heterotrimeric G-proteins. The results highlight important differences that must be taken into consideration in understanding any energetic decomposition with GB models. We begin with a review of the underlying theory, and then present the computational results along with a detailed discussion.

Theory

The Linearized-Poisson–Boltzmann Model Is Pairwise Decomposable. Implicit solvent models replace an atomistic description of solvent with a continuum model; typically, this

solvent continuum is described by macroscopic solvent properties such as a dielectric constant and surface tension. While debate remains as to the most appropriate treatment of nonpolar effects, the physical theory for interactions of charged particles in a (nonuniform) dielectric continuum is well-defined.⁴⁸

Given an arbitrary density of charge, $\rho(\vec{r})$, in a medium of spatially varying dielectric, $\epsilon(\vec{r})$, the potential in all space, $\phi(\vec{r})$, is determined by the solution to the Poisson equation, $\nabla \cdot \epsilon(\vec{r}) \nabla \phi(\vec{r}) = -4\pi\rho(\vec{r})$. When monovalent mobile ions are additionally considered (with a Debye–Hückel-like treatment), this leads to the Poisson–Boltzmann (PB) equation:

$$\nabla \cdot \epsilon(\vec{r}) \nabla \phi(\vec{r}) + \kappa^2(\vec{r}) \sinh \phi(\vec{r}) = -4\pi\rho(\vec{r}) \quad (1)$$

The Poisson equation simply lacks the second term, as $\kappa^2(\vec{r}) = 0, \forall \vec{r}$ if no mobile ions are present, and a well-defined extension to multivalent ions exists.⁴⁹ When the electrostatic potential in solution is low, the hyperbolic sine term may be approximated by the leading (linear) term in the power series expansion, giving the linearized-Poisson–Boltzmann (LPB) equation:

$$\nabla \cdot \epsilon(\vec{r}) \nabla \phi(\vec{r}) + \kappa^2(\vec{r}) \phi(\vec{r}) = -4\pi\rho(\vec{r}) \quad (2)$$

The PB equation, both in its linearized and nonlinear forms, can be solved numerically through many methods, including finite-difference,^{17,50} finite element,⁵¹ and boundary element⁵² implementations.

Decompositions of LPB energetics have been applied to many systems, but the theory underlying the decomposition is worth reviewing. Given the potential in all space, the free energy of a state in the LPB model is given by

$$G = \frac{1}{2} \int_V \rho(\vec{r}) \phi(\vec{r}) dV \quad (3)$$

with the integral taken over all space. This is indeed a free energy, as the cost of creating a reaction field in a dielectric medium (including entropic terms) is included. In a typical application involving biological macromolecules, the molecular charge distribution is approximated by a set of point charges at atomic centers. Thus, the term $\rho(\vec{r})$ vanishes at all points other than atomic centers, and the integral in eq 3 becomes a simple sum over atoms:

$$G = \frac{1}{2} \sum_{i=1}^N q_i \phi_i \quad (4)$$

where q_i is the partial atomic charge on atom i and ϕ_i is the total electrostatic potential at atom i . Now, since the LPB equation describes a linear-response model, the potential at a given point can be described by a linear combination of the potentials due to all atoms in the system. Furthermore, the potential due to a given atom is the potential due to a unit charge at that position, ϕ'_{ij} , scaled by the partial atomic charge, q_j :

$$\phi_i = \sum_{j=1}^N q_j \phi'_{ij} \quad (5)$$

Combining eqs 4 and 5, we get the result

$$G = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N q_i \phi'_{ij} q_j \quad (6)$$

Thus, we see that the electrostatic free energy of a system described by the LPB model is pairwise decomposable by atom. This feature allows the energy to be similarly decomposed into pairwise terms involving various sets of atoms, such as chemical functionalities or individual amino acids in a protein. We may simplify this expression by defining the electrostatic potential matrix, $\Phi = \frac{1}{2}[\phi'_{ij}]$, and a partial atomic charge vector, $\vec{Q} = [q_i]$, which gives the matrix expression for the electrostatic free energy:

$$G = \vec{Q}^T \Phi \vec{Q} \quad (7)$$

It is worthwhile to briefly discuss the origin of the factor of one half in the free-energy expression, which has two distinct sources. The off-diagonal terms of the Φ matrix (pairwise interactions) are halved in order to prevent double counting; in the double-sum form of eq 6, the factor of half may be dropped if the second sum is taken over $j > i$ rather than over all j . These terms involve both a Coulombic term and a term due to dielectric screening of the Coulombic interaction. The diagonal terms of Φ (self-energies) are halved for a very different reason. These terms, which correspond to the interaction of a single charge with its own reaction field, are not counted twice in the summation. However, the energetic cost of creating the reaction field is equal to half of the total interaction energy of the charge with the reaction field but of opposing sign. Thus, the factor of one half for the diagonal elements accounts for the cost of reorganizing the dielectric medium; as this includes both entropic and enthalpic terms, this leads to a free energy, and not simply a potential energy.

Note that, in the full nonlinear PB model, the energy is not simply given by eq 3, as an additional integral over all space is required to account for the energy of the mobile ions. Additionally, the potential in the nonlinear form is not a linear superposition of potentials from unit charges. Thus, this decomposition strictly applies only to the linearized form.

Free-Energy Differences in the Linearized-Poisson–Boltzmann Model. The atom-by-atom pairwise decomposability of the LPB model for the free energy of a single state leads directly to similar pairwise decomposability for differences in free energy, including the free energies of solvation, of molecular association, and of conformational changes (such as protein folding). Given two states, A and B, with corresponding potential matrices Φ_A and Φ_B , the electrostatic free-energy change going from A to B is

$$\begin{aligned} \Delta G &= \vec{Q}^T \Phi_B \vec{Q} - \vec{Q}^T \Phi_A \vec{Q} \\ &= \vec{Q}^T \Delta \Phi_{AB} \vec{Q} \end{aligned} \quad (8)$$

where $\Delta \Phi_{AB} = (\Phi_B - \Phi_A)$ is the difference of the two potential matrices. Note that this presumes a nonpolarizable description of the molecular system, where partial atomic charges do not vary with state. In a polarizable model, the pairwise decomposability remains, but the representation of eq 8 does not hold.

For a solvation free energy, the two states correspond to the same molecular configuration in two different dielectric environments; most often, the external dielectric environment is that of aqueous solution in one state and that of vacuum ($\epsilon = 1$) in

the other. As a result, the Coulombic terms (in a vacuum) are identical in the two states, and the difference matrix consists only of differences in interactions with the reaction field; in moving from a higher to lower external dielectric, the difference matrix will then be positive definite.⁴⁶ For a conformational free-energy change, the molecular composition and general dielectric environment remain fixed, but the positions of both partial atomic charges and the boundary between dielectric regions vary. Thus, the difference matrix will contain both Coulombic and reaction-field components, and no general statements can be made about the properties of the matrix. When the conformational change of interest is that of protein folding, often an approximation to the unfolded state is made in which short, contiguous sequences of amino acids (1–3) are modeled as isolated molecules. In this model, the same formalism holds, and the potential matrix for the unfolded state will be block diagonal, with zero elements for all atoms not present in the same reference molecule and each diagonal block corresponding to the potential matrix for the model. Thus, in the difference matrix, only the diagonal blocks will differ from the potential matrix of the folded state. For a binding free energy, the two states consist of (A) a complex and (B) each component of the complex in isolation. The unbound-state matrix, similar to the unfolded-state matrix discussed above, will be block diagonal, with all intermolecular elements zero; each diagonal block will be the potential matrix of the isolated molecule. Again, this leads to the result that the off-diagonal blocks of the difference matrix are unchanged from that of the complex potential matrix. In a rigid-body binding model, where the unbound states are identical in conformation to the bound state, the Coulombic entries of the diagonal blocks are identical in both states, and thus the diagonal blocks of the difference matrix contain only the difference of the reaction field terms.

Energetic Decomposition in the Linearized-Poisson–Boltzmann Model. The form of eq 8 makes it clear that the differences in electrostatic free energies are strictly pairwise decomposable over the atoms of the system, given Φ_A and Φ_B . These matrices do not depend on the magnitudes of the charges in the system but do depend both on the position of the atomic charges and on the boundary separating high from low dielectric (and nonzero ionic strength from zero). The effect of the dielectric boundary on the potential matrices is fundamentally not pairwise decomposable, and this must be considered in the definition of an appropriate reference state. Thus, we raise the concept of a hydrophobic isostere—a hypothetical molecule with an identical shape as that of interest but with some subset of partial atomic charges set to zero. Electrostatic free-energy differences in such a system will be represented by an identical $\Delta \Phi$ matrix, but the charge vector will have the appropriate subset of charges zeroed. We define zero energy to be that of the fully hydrophobic system ($\vec{Q} = \vec{0}$).

The atomic-level pairwise decomposability of the model allows for pairwise decomposability across any groupings of atoms. Consider a set of m groups, each of which may have its charges defined by a vector \vec{q}_i . Then,

$$\vec{Q}^T = [\vec{q}_1^T \ \vec{q}_2^T \ \cdots \ \vec{q}_m^T] \quad (9)$$

where we recall that the ordering of the indices for the charge vector and the potential matrix is arbitrary, as long as it is consistent. We may similarly consider $\Delta \Phi$ as a block matrix, divided in the same way:

$$\Delta\Phi = \begin{bmatrix} \Delta\phi_{11} & \Delta\phi_{12} & \cdots & \Delta\phi_{1m} \\ \Delta\phi_{21} & \Delta\phi_{22} & \cdots & \Delta\phi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta\phi_{m1} & \Delta\phi_{m2} & \cdots & \Delta\phi_{mm} \end{bmatrix} \quad (10)$$

The free-energy expression then becomes

$$\Delta G = \sum_{i=1}^m \sum_{j=1}^m \vec{q}_i^T \Delta\Phi_{ij} \vec{q}_j \quad (11)$$

which is a sum over pairs of groups, just as desired.

The meaning of each term in eq 11 depends specifically on how the groups were defined and on the physical process being considered. Consider, for example, a binding reaction between two species (L and R), where the groups are defined simply as all the atoms in each species. In this case, we have

$$\Delta G = \vec{q}_L^T \Delta\Phi_{LL} \vec{q}_L + \vec{q}_L^T \Delta\Phi_{LR} \vec{q}_R + \vec{q}_R^T \Delta\Phi_{RL} \vec{q}_L + \vec{q}_R^T \Delta\Phi_{RR} \vec{q}_R \quad (12)$$

The terms $\vec{q}_L^T \Delta\Phi_{LL} \vec{q}_L$ and $\vec{q}_R^T \Delta\Phi_{RR} \vec{q}_R$ correspond to the difference in the interactions of an individual species with its own reaction field in the bound and unbound states, which may be termed the desolvation energy for each species. As $\Delta\Phi$ is symmetric, the two off-diagonal terms are equal, and their sum ($2\vec{q}_L^T \Delta\Phi_{LR} \vec{q}_R$) corresponds to the solvent-screened Coulombic interaction between the two species in the bound state; the factor of 2 simply accounts for the factor of one half that was included in our definition of $\Delta\Phi$. Thus, we have the result

$$\Delta G = \Delta G_L^{\text{des}} + \Delta G_R^{\text{des}} + \Delta G_{LR}^{\text{inter}} \quad (13)$$

where the desolvation terms are quadratic in the charges of each species and the interaction term involves a linear contribution from each. This formalization allows for an elegant technique for finding a charge distribution on one species that optimizes the binding free energy to its partner; details of this procedure and applications have been extensively discussed in the literature.^{19–21,45–47,53–55}

Another common decomposition is into the individual chemical functionalities of a molecule. For a protein, each amino acid is easily divided into three groups: one containing all atoms of the side chain, another containing the backbone amino group, and a third containing the backbone carbonyl. Again considering a binding reaction, we obtain terms of three types (unrelated to the division of each amino acid into three groups). These are the following:

- The diagonal terms, $\vec{q}_i^T \Delta\Phi_{ii} \vec{q}_i$, which correspond to the difference (between the bound and unbound state) of the interactions that the i th group makes with its own reaction field. This is termed the *desolvation* energy of group i , analogously to the desolvation of species A and B, above.
- The off-diagonal terms, $\vec{q}_i^T \Delta\Phi_{ij} \vec{q}_j$, where group i and group j belong to different binding partners. These are easily interpreted as the solvent-screened Coulombic interactions between the two groups in the bound state, analogous to the interaction term above, and are termed *direct interactions*.

• The off-diagonal terms, $\vec{q}_i^T \Delta\Phi_{ij} \vec{q}_j$, where group i and group j belong to the same binding partner. In this case, there are nonzero contributions from both the bound and unbound states, and in the rigid-body binding model, the Coulombic terms in the two states perfectly cancel. Thus, these terms correspond to the difference in the solvent screening that intramolecular interactions experience in the bound and unbound states, referred to as *indirect interactions*.

To compute the $\Delta\Phi$ matrix, a separate solution to the LPB equation must be found for each atom in the system, in each of the two states. For each state, considering a system in which a single atom is charged to 1 (and all others are replaced with their hydrophobic isosteres) gives the potential at all atoms due to a unit charge at one specific atom. The vector of these potentials forms a single row of the Φ matrix for that state.

However, while the full $\Delta\Phi$ matrix provides the motivation for energetic decomposition, it is generally not computed; rather, a single calculation can be done for each group in the block-matrix representation. A system in which a single group is charged (with appropriate partial atomic charges) and all others are replaced with hydrophobic isosteres provides the potential due to that single group; this provides each of the terms from eq 11 ($\vec{q}_i^T \Delta\Phi_{ij} \vec{q}_j$) for a particular group, i . Details and applications of these decompositions have also been discussed in the literature.^{22,42,56}

Energetic Decomposition in the Generalized-Born Model. As discussed above, in the linearized-PB model, a separate calculation must be done for each group. For a full residue-by-residue decomposition of a large protein complex, this can involve several thousand individual calculations. While this is entirely feasible for a single structure, it becomes computationally costly to take this approach to ensembles of structures of any significant size. For example, to perform a complete analysis over an ensemble of 1000 snapshots from a constant temperature molecular dynamics simulation could require millions of solutions of the LPB equation.

The generalized-Born model has gained a great deal of favor in providing a fast alternative to Poisson–Boltzmann electrostatics in situations where speed is a key factor, such as implicit-solvent molecular dynamics and high-throughput virtual screening. Thus, the question naturally arises as to whether GB may provide a useful approximation to PB-based energetic decomposition.

In the GB model, each atom in the system is described by an effective Born radius, which is the radius for which a sphere with a centrally located charge would have the same solvation free energy as the charge located at the atomic center in the actual system of interest; for unit charges, these values correspond to the diagonal elements of the Φ matrices discussed above. Given the effective Born radii, the free energy is approximated by the Still equation:³¹

$$G = \frac{1}{8\pi\epsilon_0} \left[\sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{\epsilon_{\text{int}} ||\vec{r}_{ij}||} - \left(\frac{1}{\epsilon_{\text{int}}} - \frac{1}{\epsilon_{\text{ext}}} \right) \frac{q_i q_j}{\sqrt{||\vec{r}_{ij}||^2 + R_i^2 R_j^2} e^{-||\vec{r}_{ij}||^2/(4R_i R_j)}} \right] \quad (14)$$

where the first term (skipped for $i=j$) corresponds to Coulomb's law in uniform dielectric, ϵ_{int} , and the second term accounts for solvent polarization (solvent dielectric, ϵ_{ext}). Clearly, this expression can be expressed in an analogous fashion to eq 7, with $\Phi_{ij} = (1/\epsilon_{\text{int}} ||\vec{r}_{ij}||) - [(1/\epsilon_{\text{int}}) - (1/\epsilon_{\text{ext}})] \{1/||\vec{r}_{ij}||^2 + R_i^2 R_j^2\}^{-1/2}$.

$R_i R_j e^{-\|\vec{r}_{ij}\|^2/(4R_i R_j)}\}^{\frac{1}{2}}$. Thus, the decompositions with respect to groups of charges apply equally well to GB-based electrostatics.

For the self-energy of individual charges, the Still equation is rigorously correct when provided exact effective Born radii; the pairwise interactions converge to the energy of a dipole in a sphere in the small distance limit and to the interaction of two isolated point charges in the far distance limit. In practice, however, exact effective Born radii are not known but rather are approximated by a number of different approaches.^{32,34,57,58} Case and co-workers have also derived an extension to the GB model that allows for a treatment of mobile ions;⁵⁹ the extension simply involves a substitution in eq 14 of

$$\left(\frac{1}{\varepsilon_{\text{int}}} - \frac{1}{\varepsilon_{\text{ext}}} \right) \rightarrow \left(\frac{1}{\varepsilon_{\text{int}}} - \frac{e^{-\kappa\sqrt{\|\vec{r}_{ij}\|^2 + R_i R_j}} e^{-\|\vec{r}_{ij}\|^2/(4R_i R_j)}}{\varepsilon_{\text{ext}}} \right) \quad (15)$$

Note that, in the GB formulation of the potential matrix Φ , only the effective Born radii are needed; this is a result of the fact that eq 7 separates contributions from geometry and from partial atomic charge. However, whereas in a PB-based model a full solution of the Poisson–Boltzmann equation is required for each row in the potential matrix, a single evaluation of the GB energy (in any implementation) yields all effective Born radii, and thus one energetic evaluation provides the full potential matrix. This could potentially result in tremendous savings in computational cost; while an individual GB computation may be 10- to 100-fold faster than a single solution of the PB equation, when upward of a thousand components are involved, the savings becomes 4 or 5 orders of magnitude.

However, while the potential savings in computation cost are exceptional, it is important to consider accuracy as well. While many studies have considered the overall ability of generalized-Born methods to recapitulate Poisson–Boltzmann-based solvation and binding free energies, there have been few studies of how the two methods compare when energetics are decomposed into individual contributions. The remaining sections of the manuscript discuss these issues.

Methods

Structure Preparation. Initial heavy-atom coordinates were obtained from the Protein Data Bank (1gia, 1gdd, and 1gp2).⁶⁰ The positions of heavy atoms with missing electron density were modeled using coordinates from the other structures (where these atoms were observed). Protonation states were chosen with the REDUCE program, as were the orientations of amides and imidazole rings;⁶¹ this resulted in a single protonated histidine in the β subunit of the trimer (1gp2), with all other histidines in a neutral state, all aspartates and glutamates negatively charged, and all lysines and arginines positive. Hydrogen atom positions were determined with the HBUILD module⁶² of the CHARMM computer program.

Molecular Dynamics. All-atom, explicit solvent molecular dynamics simulations were performed using the CHARMM⁶³ and NAMD⁶⁴ programs with PARAM22/27 parameters;^{65,66} system setup and postprocessing were done with CHARMM, while production dynamics simulations were run using NAMD. Each complex was placed in a pre-equilibrated box of TIP3P waters,⁶⁷ with a minimum of 10 Å between any solute atom and the box edge in all directions. Random water positions were replaced with enough sodium and chloride ions to yield physiological (145 mM) ionic strength; the sodium to chloride ratio was adjusted from unity to provide a system with zero net

charge. Periodic boundary conditions were applied to minimize edge artifacts, and particle-mesh Ewald (PME) summation was used for long-range electrostatic interactions; short-range interactions were cut off at 12 Å.

Poisson–Boltzmann Calculations. All solutions of the linearized-Poisson–Boltzmann equation were obtained using a multigrid finite-difference solver distributed with the ICE (Integrated Continuum Electrostatics) package.^{68,69} The atomic radii used were those optimized by Roux and co-workers for use with continuum electrostatic models;^{70,71} partial atomic charges were taken from the PARAM22/27 parameter set.^{65,66} For consistency with the all-atom molecular dynamics simulations, an internal dielectric constant of 1 was used, along with an external dielectric constant of 80. Two sets of calculations were performed, one with an ionic strength of 0.145 M, using a 2.0 Å ion exclusion (Stern) layer, and the other with 0.0 M salt. Binding free energies were computed as the difference between the bound complex and rigidly separated unbound components; the bound and unbound states were positioned identically on the finite-difference grid so as to cancel artifactual grid energy. Focusing boundary conditions were used: an initial calculation was done with the largest dimension of the system occupying 23% of one side of the grid, and Debye–Hückel boundary conditions were used; a second calculation was then done with the system occupying 92% of the grid, and using the potentials from the previous calculation at the boundary. For individual components, but not for net binding energies, an additional overfocused calculation was done with the molecule occupying 184% of the grid (centered on the component of interest). Net binding energies were computed with a 257³ unit cubic grid, and individual components with a 129³ unit grid; due to the overfocusing used for components, the grid spacing was identical in both of these cases.

Generalized-Born Calculations. Generalized-Born calculations were done using the GBSW module⁵⁸ of a version of the CHARMM computer program modified to output the effective Born radii. As for the PB calculations, the atomic radii used were those optimized by Roux and co-workers.^{70,71} Binding was considered as the difference between the bound state and a state in which one component was rigidly translated 500 Å. The scaling coefficients were set to standard values, of $a_0 = 1.2045$ and $a_1 = 0.1866$, the molecular surface was used, and a smoothing length of 0.2 Å was applied; no cutoffs were used. GBMV calculations (Supporting Information) were also performed using CHARMM,³⁴ using Lebedev integration with a grid of 38.

Component Analysis. PB-based component analysis was done with the ICE software package, using a standard approach.^{68,72} Each amino acid was divided into three groups: side chain, amino (including C_α and attached hydrogens), and carbonyl. For each group, the LPB equation was solved (in both the bound and unbound states) for a system in which only that group is charged; the difference in the potentials from this pair of calculations was used to compute the self-desolvation for the group and all interactions (direct and indirect) involving the group.

GB-based component analysis was done with the same partitioning of groups. Effective Born radii for each atom in the bound and unbound states were computed with the CHARMM computer program as outlined above. These radii were then used within a separate, locally written program to compute the pairwise energies between all groups. As every term is a difference between the bound and unbound states, intramolecular terms (desolvation and indirect interactions) include only the

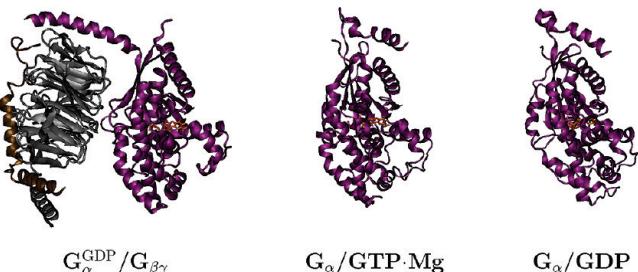


Figure 1. Overview of complex structures. For each system studied, a representative structure is displayed, with G_α in purple cartoon, G_β in gray cartoon, G_γ in bronze cartoon, and the nucleotide (GTP or GDP) in tan licorice: (left) the G-protein heterotrimer, $\text{G}_\alpha^{\text{GDP}}/\text{G}_\beta\gamma$ (PDB ID 1gp2); (center) the activated G_α monomer, $\text{G}_\alpha/\text{GTP}\cdot\text{Mg}$ (PDB ID 1gia); (right) the G_α monomer in the inactive state, $\text{G}_\alpha/\text{GDP}$ (PDB ID 1gdd). Figures generated with VMD.⁷⁸

polarization term of eq 14, while intermolecular (direct) interactions additionally include the Coulombic term for the bound state.

Results

The heterotrimeric G-proteins can exist in both a bound complex or an unbound state; additionally, the α -subunit may be bound to guanosine-5'-di- or triphosphate. In the unbound state, G_α has been structurally characterized bound to both GDP and GTP (or more precisely, a nonhydrolyzable GTP analogue), while, in the bound state, only the GDP-bound state is experimentally accessible—nucleotide exchange is coupled with trimer dissociation as a fundamental step in signaling. We have thus begun our analysis of this system with three states: isolated G_α bound to GDP; isolated G_α bound to GTP; and G_α bound to GDP, in complex with $\text{G}_\beta\gamma$.

In each of these systems, we have considered a single binding equilibrium; for the unbound α -subunits this is nucleotide association, while for the trimeric complex it is the association of $\text{G}_\beta\gamma$ with $\text{G}_\alpha\text{-GDP}$. Each of these systems has distinct features at the binding interface, and thus, the set of three presents a range of interaction chemistries. In particular, $\text{G}_\alpha^{\text{GDP}}\text{-G}_\beta\gamma$ binding is characteristic of the interactions of large proteins, with a large amount of surface area buried on binding, and a diverse range of side-chain–side-chain, side-chain–backbone, and backbone–backbone contacts. Nucleotide binding, on the other hand, involves the near complete burial of a small organic molecule in a deeper binding pocket. The nucleotides are both highly polar molecules carrying a net negative charge ($-3e$ for GDP and $-4e$ for GTP). GTP, however, is associated in the binding site with a magnesium ion. In solution, the preferred state of nucleotide triphosphates is in complex with magnesium,⁷³ and thus, considering the ligand as a $\text{GTP}\cdot\text{Mg}$ complex with net charge $-2e$ is a reasonable approximation. These structures are displayed in Figure 1.

Electrostatic Binding Free Energies Computed with Generalized Born and with Linearized Poisson–Boltzmann are Well-Correlated. Molecular dynamic simulations of each system were performed for 100 ns in explicit solvent with periodic boundary conditions. Analysis of these trajectories indicated that they were all stable and well-behaved (see the Supporting Information). 101 evenly spaced snapshots were then extracted for the computation of binding free energies with both the generalized-Born and Poisson–Boltzmann methods. These data are shown in Figure 2 and Table 1. There is quite strong correlation between GB- and PB-computed electrostatic binding free energies in all cases, with R^2 values of greater than 80%;

these correlations persist over a range of over 100 kcal/mol in all three systems. However, there are important differences as well.

Generalized Born Overestimates Net Electrostatic Binding Free Energies. While for some snapshots the GB-computed binding free energy is lower than that computed with PB, overall there is a strong trend for lower binding free energies with PB. The ensemble-averaged binding free energy for the trimer is 2 kcal/mol lower when calculated with PB as compared to the same value computed with GB; the nucleotides show an even greater tendency for GB to give higher values, with a 10 kcal/mol difference for $\text{GTP}\cdot\text{Mg}$ binding to G_α and a 22 kcal/mol difference for GDP. While these differences are within the standard deviations of the ensembles (see Table 1), they are not within the standard errors of the mean.

Generalized Born Underestimates Differences in Net Electrostatic Binding Free Energies between Configurations. Considering the best-fit lines relating the two values makes it clear that the reason for these discrepancies is not simply a difference of reference state. In all cases, the least-squares fit describing the variation of GB-computed binding energies to those computed with PB has a slope less than unity—slightly above 0.8 for association of the trimer and for the binding of $\text{GTP}\cdot\text{Mg}$ to G_α and below 0.6 for GDP– G_α binding. A slope of below unity indicates that GB will tend to underestimate the magnitude of free-energy differences between configurations, as compared to PB—both highly favorable (negative) and unfavorable (positive) values will be less so with GB. While this general trend holds in all three cases, the offset from zero (intercept) is very different in all cases. For the association of the protein components of the trimer, where the net electrostatic binding energy is always positive, the intercept is also significantly positive (+15 kcal/mol); it is this positive offset that makes the mean GB result greater than that from PB, despite a slope of correlation below $y = x$. For the binding of the nucleotides, however, the intercept is negative (−15 kcal/mol for $\text{GTP}\cdot\text{Mg}$ and −34 kcal/mol for GDP). Despite this, the highly negative values of the binding free energies make the less than unit slope dominate the difference in the means. Although there are significant variations between the equations of the best-fit lines in the different systems, fitting the data from all three components gives a strong correlation ($R^2 = 0.98$) with a least-squares fit having a slope of 0.91 and intercept of 5.3 (see the Supporting Information). Thus, gross differences between systems seem to be more closely captured by the two models than are the details of energetic variation within a system.

Energetic Components Show Highly Variable Correlations between GB and LPB Models. As discussed in the Theory section, one of the great benefits of linear-response models such as generalized Born and linearized Poisson–Boltzmann is that computed energies may be rigorously decomposed into contributions from individual groups and pairs of groups. In order to better understand the observed differences in binding free energies computed with the two models, we performed such a decomposition with both approaches. Each amino acid was partitioned into three groups: side chain, backbone amino, and backbone carbonyl. For each group, the difference in the self-energy in the bound and unbound states (due to differences in interactions with solvent) is termed the desolvation energy; for each pair of groups within a molecule, the difference in solvent screening of the interaction (in the bound and unbound states) is termed the indirect interaction; and for each pair of groups on opposite binding partners, the solvent-screened Coulombic interaction is termed the direct interaction. Finally, the sum of

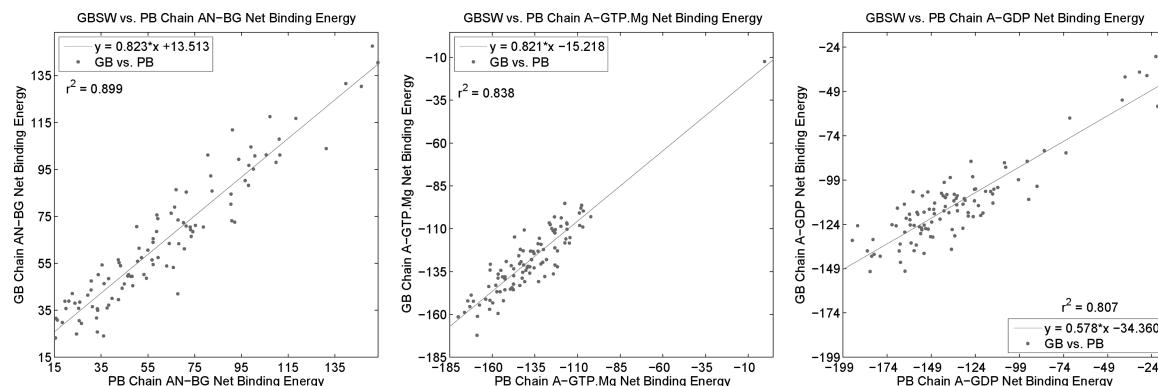


Figure 2. Variation in single-snapshot net binding energies computed by GB and PB. Rigid-body electrostatic binding free energies (in kcal/mol) computed for each snapshot by both GB (y-axis) and PB (x-axis) are shown for all three systems studied: (left) $G_{\alpha}^{\text{GDP}}/G_{\beta\gamma}$; (center) $G_{\alpha}/\text{GTP}\cdot\text{Mg}$; (right) G_{α}/GDP . Note that the single outlier of high energy for $G_{\alpha}/\text{GTP}\cdot\text{Mg}$ corresponds to the first frame of the simulation; when this point is excluded, R^2 is 0.761, and the best-fit line is $y = 0.808x - 17.078$.

TABLE 1: Net Electrostatic Binding Free Energies^a

system	mean		std. dev.		R^2	rms
	PB	GB	PB	GB		
$G_{\alpha}^{\text{GDP}}/G_{\beta\gamma}$	62.16	64.70	32.26	28.01	0.90	10.82
$G_{\alpha}/\text{GTP}\cdot\text{Mg}$	-138.38	-128.81	22.98	20.61	0.84 ^b	13.30
G_{α}/GDP	-135.33	-112.63	36.11	23.25	0.81	29.13

^a All energies are in kcal/mol; all values are computed over a set of 101 evenly spaced snapshots from a 100 ns MD trajectory. ^b R^2 is 0.761 when the first snapshot in the trajectory is excluded.

all terms involving a single group (desolvation and all interactions, both direct and indirect) is termed the mutation energy; the mutation energy is the energetic difference in binding free energy between the natural system and a hypothetical system with that group (and that group alone) replaced with a hydrophobic isostere. Figure 3 and Table 2 summarize these data, excluding components for which both GB and PB are within ± 1.0 kcal/mol of zero; these consist primarily of components not close to the binding interface.

Direct, Solvent-Screened Coulombic Interactions Show Bimodal Correlations. In considering the mutation energies, a very interesting result appears. The great majority of points show strong correlation in the values computed by each method, with the data largely falling along a line of near unit slope. However, in all systems, there appears to be a subpopulation with a decidedly lower slope. This is most apparent for the nucleotide-binding systems but is also seen for a few groups in trimer association.

The mutation energy includes both indirect (intramolecular) and direct (intermolecular) terms, as well as self-desolvation penalties, and it is informative to consider how each of these vary when computed in the two models. Individual group desolvation energies computed with the two models are only loosely correlated. On average, generalized Born underestimates individual desolvation energies; this is by a small amount for the components of the trimer but a larger amount for the protein components of the nucleotide-binding systems. However, visual analysis of these data clearly reveals that the deviations are not randomly distributed. Indirect interactions, corresponding to the differential solvent screening of intramolecular interactions in the bound and unbound state, show relatively weak correlations, although the degree of correlation is system-dependent. The direct interactions, on the other hand, show the greatest degree of correlation. The direct interactions also most clearly exhibit the existence of two populations, with different slopes of correlation; this seems to be the primary but not exclusive source of the patterns seen in the mutation term.

Inconsistencies between GB and PB Are Group-Specific.

That certain groups show qualitatively different behaviors raises the possibility that the distinct properties of particular amino acids may play a role. To consider this, the contributions of each group were further broken down by group type. Figures 4 and 5 show these data for the association of G_{α}^{GDP} and $G_{\beta\gamma}$ to form the G-protein heterotrimer.

Individual group solvation free energies (Figure 4) show a wide range of behaviors, with some groups showing strong correlation between the GB- and PB-computed values (10 of the 39 most significant groups have R^2 values of greater than 0.80) and others showing minimal correlation (R^2 is below 0.5 for 13 of the 39 groups with significant desolvation energies by PB). For example, essentially all glutamate and lysine points are distributed about $y = x$, albeit with some degree of variation. Glutamine, tyrosine, and arginine also display relatively well-distributed linear correlations but with slopes somewhat below unity; the backbone groups (both amino and carbonyl) also tend to follow this pattern. However, some groups show much more structure in the variation between the two methods. The data for asparagine, for example, fall into distinct clusters, largely identified as particular residues; the data for Asn B88, for example, are well-correlated ($R^2 = 0.88$) and approximately follow $y = x$ (with a slight offset), while the data for Asn B119 are more weakly correlated ($R^2 = 0.40$) with a slope significantly below 1. For serine and threonine, most positions show very little correlation but in a manner that varies; Ser A16 and A206, as well as Thr A182 and B153, all have GB-computed values largely confined to about a 1 kcal/mol range near 3 kcal/mol, while their PB-computed values span a wider range (1–6 kcal/mol). In other cases, however, some degree of correlation is seen, although generally with a slope much less than 1. The most profound deviation between the GB- and PB-computed values is for the protonated histidine at B54, with a correlation coefficient of only 0.18; by Poisson–Boltzmann, the desolvation of this group ranges from about 5 to 25 kcal/mol, while, with GB, the desolvation never exceeds 5 kcal/mol.

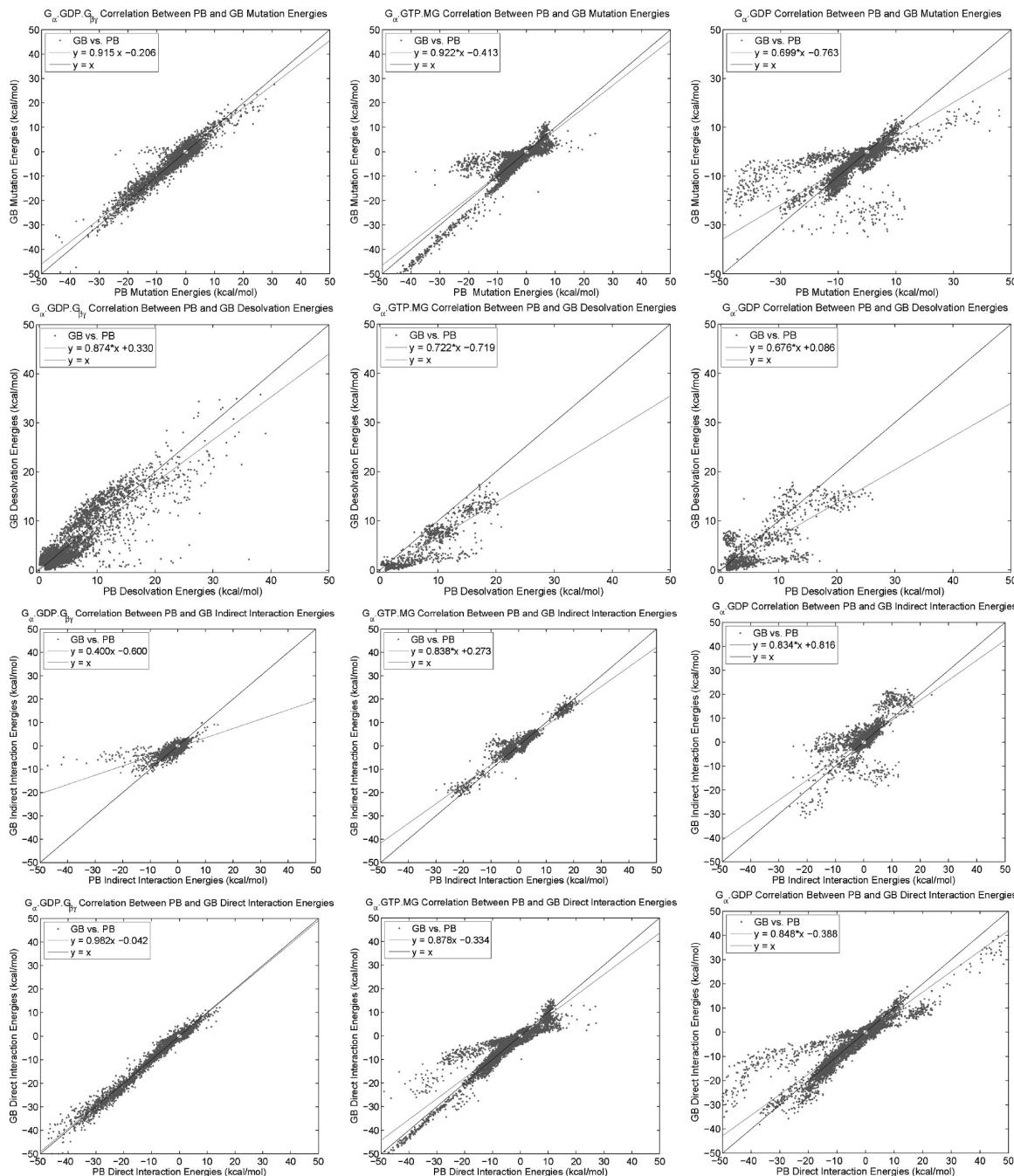


Figure 3. Variation in group-based energetic terms computed by GB and PB. The individual group-based energetic terms computed for every snapshot of each system are displayed. Columns correspond to the three different systems: (left) G_α^{GDP}/G_{βγ}; (center) G_α/GTP·Mg; (right) G_α/GDP. Rows correspond to different energetic terms: (top) mutation energies (relative to hydrophobic isostere); (second) desolvation penalties; (third) indirect (intramolecular) interactions; (bottom) direct (intermolecular) interactions. All energies are in kcal/mol, and for clarity, components with energies computed to be within ± 1.0 kcal/mol by both methods have been excluded, and the range plotted is restricted to ± 50.0 kcal/mol.

Contributions of a Buried Salt Bridge Are Missed by GB.

As noted earlier, group mutation energies in this system show a remarkable level of correlation given the many inconsistencies in the desolvation terms, and this is true for essentially all group types (Figure 5); 27 of the 39 most significant groups (by PB) have correlation coefficients between the two models of greater than 0.80. There are, however, two notable exceptions to this: Asp 76 and His⁺ 54 on the β-subunit. Both of these groups are computed in the Poisson–Boltzmann model to make significantly favorable contributions in a subset of snapshots, of as much as -25 kcal/mol for Asp 76 and -15 kcal/mol for His 54, while the terms computed with GB are uniformly low in

magnitude. As seen in Figure 6, these two groups are located proximal to one another, near—but not directly involved in—the interface with G_α. In the crystal structure, they are involved in an intramolecular salt bridge with one another; through the course of the simulation, this salt bridge breaks and then reforms. In all cases, however, both groups make additional intramolecular interactions that keep the groups localized beneath the protein surface. Under the Poisson–Boltzmann model, the desolvation penalty paid by both groups is substantial for certain snapshots, while the GB model suggests a small desolvation, uniform across all snapshots; this leads to a correlation coefficient between the two models of only 0.24 for His and 0.05

TABLE 2: Group-Based Electrostatic Binding Energy Components^a

system	N	mean		std. dev.		R^2	rms
		PB	GB	PB	GB		
Mutation Energies							
$G_{\alpha}^{GDP}/G_{\beta\gamma}$	9485	-1.70	-1.76	5.81	5.57	0.91	1.75
$G_{\alpha}/GTP \cdot Mg$	7328	-6.30	-6.22	18.18	17.35	0.93	4.71
G_{α}/GDP	9621	-4.21	-3.71	14.85	11.30	0.84	6.31
Desolvation Penalties							
$G_{\alpha}^{GDP}/G_{\beta\gamma}$	5505	4.67	4.41	5.17	5.08	0.79	2.43
$G_{\alpha}/GTP \cdot Mg$	1341	5.75	3.43	5.23	4.14	0.83	3.22
G_{α}/GDP	1439	5.33	3.69	5.34	4.51	0.64	3.60
Indirect Interactions							
$G_{\alpha}^{GDP}/G_{\beta\gamma}$	5294	-1.23	-1.09	3.25	1.87	0.48	2.38
$G_{\alpha}/GTP \cdot Mg$	5710	0.68	0.84	5.02	4.44	0.90	1.65
G_{α}/GDP	4765	0.31	1.08	4.59	5.36	0.51	3.89
Direct Interactions							
$G_{\alpha}^{GDP}/G_{\beta\gamma}$	9490	-3.87	-3.84	8.67	8.60	0.98	1.26
$G_{\alpha}/GTP \cdot Mg$	10342	-5.74	-5.38	18.99	17.07	0.96	4.30
G_{α}/GDP	10812	-4.68	-4.36	16.85	14.72	0.94	4.41

^a All energies are in kcal/mol. Values are computed over the set of components from all 101 snapshots with either a GB- or PB-computed energy of greater than 1.00 kcal/mol in absolute value; N is the size of this set.

for Asp, and the best fit lines have a slope of only 0.20 and 0.04, respectively. In the PB model, the indirect interactions made between these groups also are very large in some cases (particularly when the salt bridge is formed); these correspond to the enhancement of the favorable electrostatic interactions between the two groups upon removal of solvent from the volume of the α -subunit, and the GB model is unable to capture these effects.

Deviations between PB and GB Are Larger for Nucleotide Binding. The differences seen between the two methods in assessing the contributions of individual groups to the association of the G_{α} and $G_{\beta\gamma}$ subunits are generally moderate, particularly due to the cancellation of error in the mutation term. However, this is much less so in the case of nucleotide binding. Figures 7 and 8 show these data for the association of G_{α} with GTP·Mg. In this system, self-desolvation energies computed with GB are greatly underestimated (with respect to the values from PB) for the majority of groups with significant desolvation. This is seen for serine and threonine, as well as for cysteine, asparagine, and tyrosine, although no terms are particularly large for the latter three. Similar behavior is seen for the backbone groups. For individual groups, even when there is significant correlation, the best-fit line has a very low slope; the GB- and PB-computed results for Ser 151 are correlated with an R^2 value of 0.50, but the best-fit line has a slope of 0.15, and for the backbone NH of position 45 the R^2 value is 0.36, but the best-fit slope is 0.20 (correlation coefficients and best-fit lines for all significant groups are given in the Supporting Information). For both aspartate and arginine, reasonable agreement between the two methods is seen, although with significant variance—some positions are highly correlated with near unit slope (for example, Asp 272 with $R^2 = 0.96$, slope 0.84, and intercept -0.1), while others are less so (such as Asp 200, with $R^2 = 0.30$, slope 0.38, and intercept 3.3). For lysine, while a (weak) overall positive correlation is seen when all positions are considered, when individual positions are considered, the GB-computed values span a very small range as compared to the PB-computed values; with Lys 51, for example, the two methods give a correlation coefficient of only 0.27, and the best-fit line has a slope of 0.24 and an intercept of 8.54. Finally, the two methods give results with very little correlation for the single

glutamate that is desolvated; the data for Glu 43 have an R^2 value of 0.27.

Again, the mutation terms agree much better when computed in the two different models. For the backbone groups, as well as serine, threonine, aspartate, arginine, and lysine, the GB-computed mutation energies are linearly correlated to those computed with PB with quite a tight distribution. For example, while the desolvation energies computed with the two methods for Ser 47 were completely uncorrelated ($R^2 = 0.04$), the mutation terms are correlated with an R^2 value of 0.98; the data for the backbone of position 45, discussed earlier, shifts from a weak correlation with slope 0.2 for desolvation to a strong correlation ($R^2 = 0.79$) with a near perfect $y = x$ correlation (best-fit line, $y = 1.02x - 0.09$). However, while in the case of the $\alpha-\beta\gamma$ subunit association was largely due to a cancellation of errors in the computation of desolvation and interaction terms, in this case, the primary effect seems to be the dominance of the interaction terms, which are highly correlated in the two models. This can be seen clearly in the magnitude of the terms; for example, the (poorly correlated) desolvation energies for serine range from 0 to 4.5 kcal/mol, while the (highly correlated) mutation energies span a range of -40 to +10 kcal/mol. As seen in Figure 9, Ser 47 makes a hydrogen bond with the α phosphate of GTP; in the bound state, this position is completely buried from solvent. In a buried site, the direct interaction energies will approach the limit of Coulomb's law in a uniform dielectric constant equal to that of the protein interior (in this case, 1.0). Both GB and PB approach this limit in this case, and thus agree well. These terms are large in magnitude because of the highly charged nature of the ligand, and the low internal dielectric constant. Adding to this, the desolvation penalties of many of the groups lining the binding pocket are rather low, due to significant burial in the unbound state.

Moderate-Range, Action-at-a-Distance Interactions Are Poorly Captured by GB. In other cases, however, greater deviations between the GB- and PB-computed mutation energies are seen. Gln 204 pays essentially zero desolvation, and thus, the mutation energies are again dominated by interactions. However, the correlation of the values computed with the two models, while linear, has a slope much lower than 1; similar results are seen for certain asparagine, arginine, and lysine

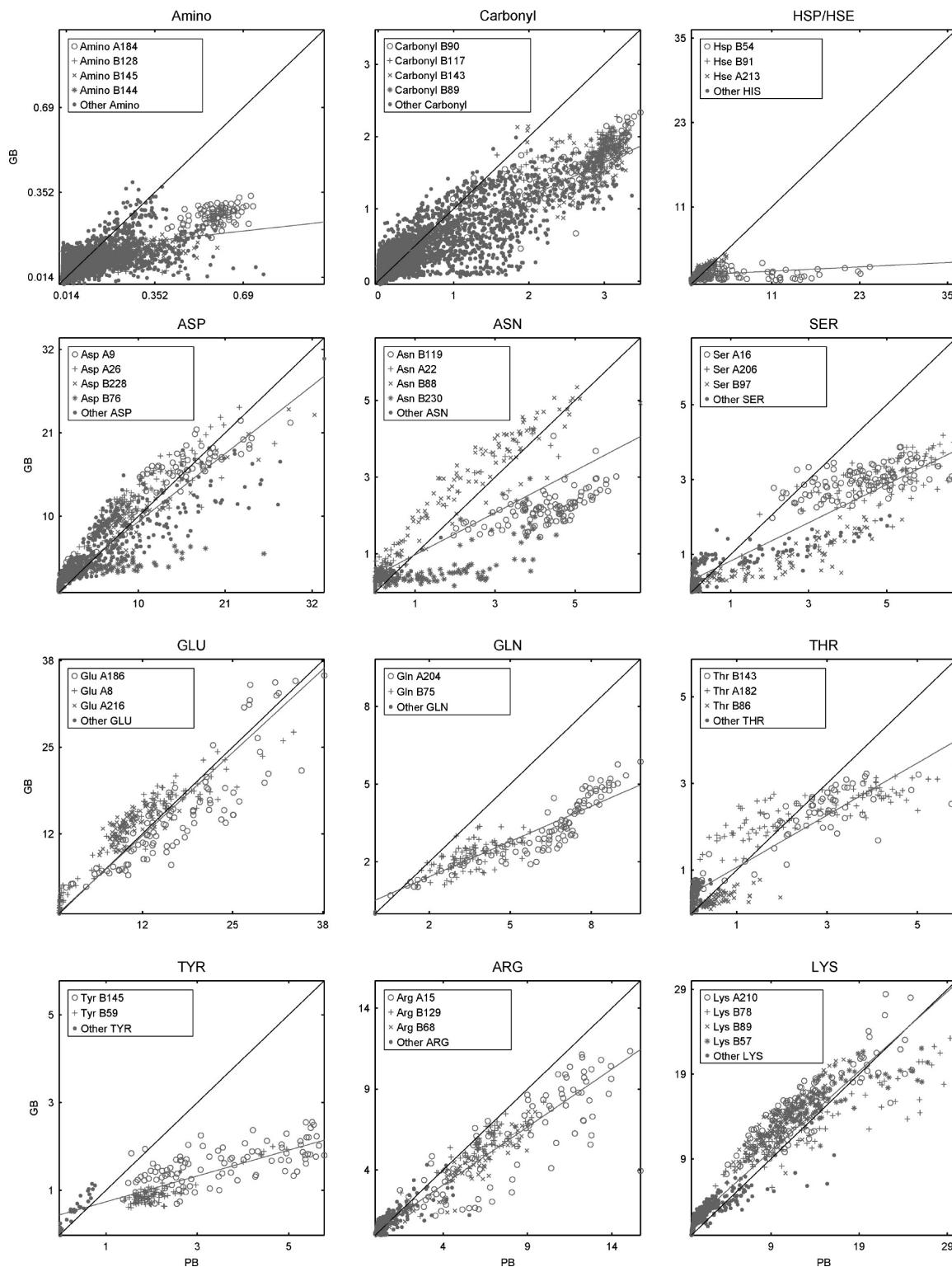


Figure 4. $G_u^{\text{GDP}}/G_{\beta\gamma}$ group desolvation penalties, by type. The data displayed in Figure 3 (second row, left) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no cysteines make interactions across the interface, these are also not shown.

residues. As can be seen in Figure 9, Gln 204 does not make direct contact with the GTP ligand but does make moderate-range, action-at-a-distance electrostatic interactions with the triphosphate. The GB model overestimates the solvent screening of this interaction, as compared to the PB-computed values. The results for Glu 43, which showed no correlation in desolvation, remain essentially uncorrelated in the mutation term. The

desolvation of this group is quite small for a charged group, and the interactions are unfavorable. This group is near the phosphate of the nucleotide but not involved in any direct interaction; the unfavorable action-at-a-distance effects of this group are not well-captured by GB. Lys 51, noted above as being weakly correlated with a slope much below unity for desolvation, behaves similarly in mutation energies, with only

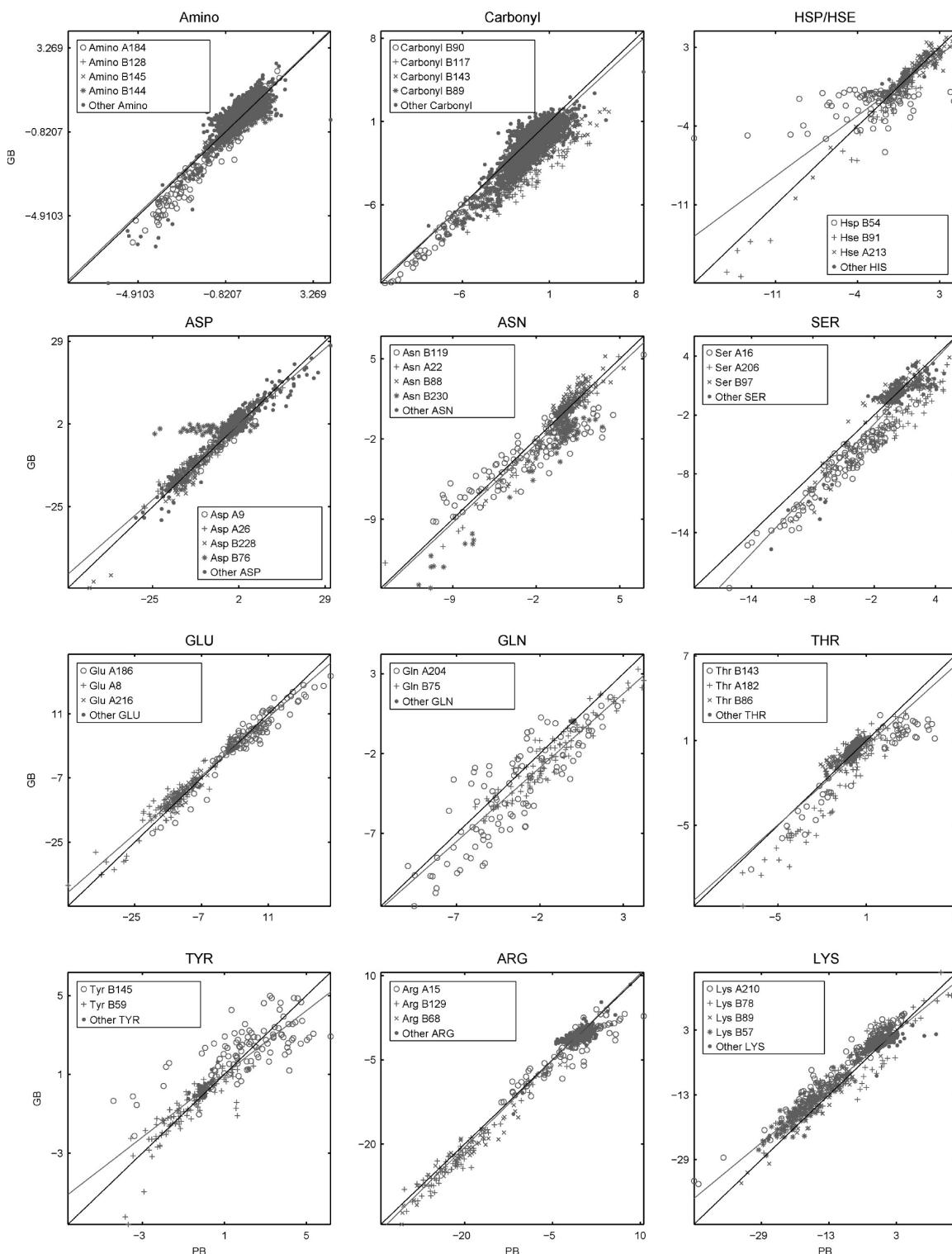


Figure 5. $G_{\alpha}^{\text{GDP}}/G_{\beta}^{\text{GDP}}$ group mutation energies, by type. The data displayed in Figure 3 (first row, left) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no cysteines make interactions across the interface, these are also not shown.

a slightly improved correlation coefficient of 0.50; this position is proximal to the phosphate tail of the nucleotide but does not make direct hydrogen-bonding interactions.

Structurally Similar Interactions Are Reproduced between Systems with Both Methods. Structurally and chemically, the binding of GDP to G_{α} is very similar to that of GTP•Mg, and thus, it may be expected that the behavior of the two models in these systems would also be similar. The data for G_{α} -GDP

binding are shown in Figures 10 and 11, and while many similarities are seen, there are notable differences. The desolvation energies of Glu 43, showing very little correlation with GTP•Mg binding, are strongly correlated ($R^2 = 0.90$) with a reasonable slope (0.68) with GDP; while a similar correlation remains in the mutation terms ($R^2 = 0.85$), the slope is significantly reduced (0.36). In this state, Glu 43 makes a hydrogen-bonded salt bridge with Arg 178; this interaction

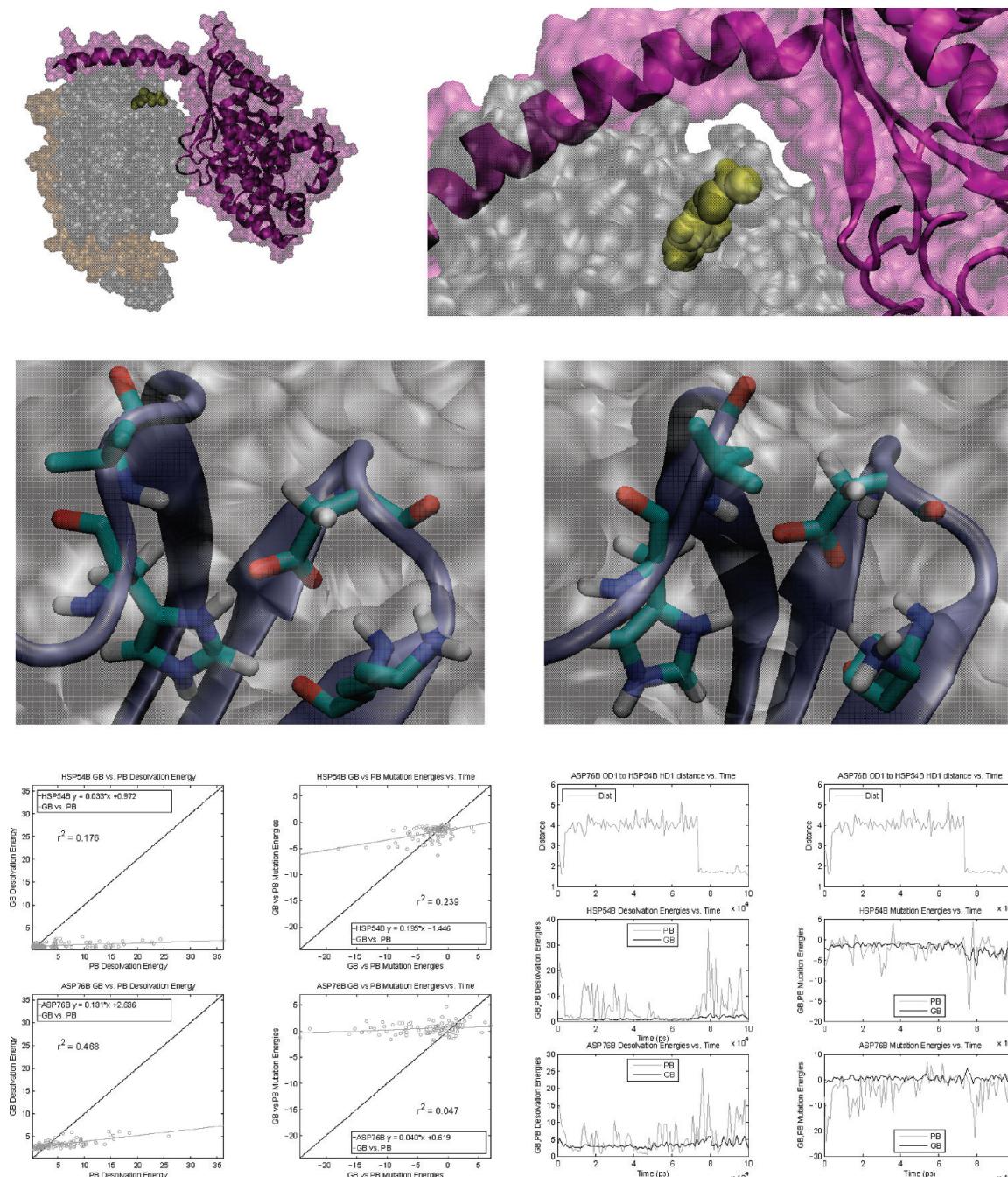


Figure 6. Structure of His⁺ 54 and Asp 76 of G β in the G α GDP/G β complex. Top: This pair of residues is located proximal to the G α binding interface (left) but is oriented so as not to make direct interfacial contacts, nor to be solvent exposed (right). Center: In some snapshots, the two groups make a hydrogen-bonded salt bridge, with a representative structure shown on the left; in others, the aspartate shifts to interact with the backbone amino group of Leu 55. Bottom: The energetics of each component are shown, with the left-most panels displaying the correlation of GB- and PB-computed desolvation and mutation terms for each residue (His⁺ 54 top, Asp 76 bottom) and the right-most panels showing the variation of these terms during the coarse of the simulation; the top graph of the left-most panels shows a metric of salt bridge formation. Structural figures generated with VMD.⁷⁸

covers the nucleotide diphosphate and thus brings the aspartate closer, making correspondingly stronger repulsions. These more intimate interactions seem to be better captured by GB, although there are still significant discrepancies.

The desolvation energies of lysines and arginines computed by the two methods appear, on visual inspection, to be less correlated with GDP binding. However, more careful analysis shows that the dominant observation is simply that the various groups have more similar mean desolvation penalties; there is little correlation for the data of a single position in either case, but with GTP•Mg binding, the wider range of contributions from different positions results in a better apparent correlation.

Quantification of Energetic Differences between Systems Can Be Very Model Dependent. When the interactions made in the two systems are similar, comparisons between the model generally give consistent results within an individual system, as seen in the case of Ser 47. Thr 48 also behaves similarly with both GTP•Mg and GDP binding, and is worthy of note as the details of the interaction change in the two systems. While in the GTP•Mg-bound state, the side-chain oxygen coordinates the metal ion, in the GDP-bound form, the side-chain hydroxyl donates a hydrogen bond directly to the terminal (β) phosphate (Figure 9). These interactions differ in important ways, but both are direct, short-range interactions, where the correlation of the

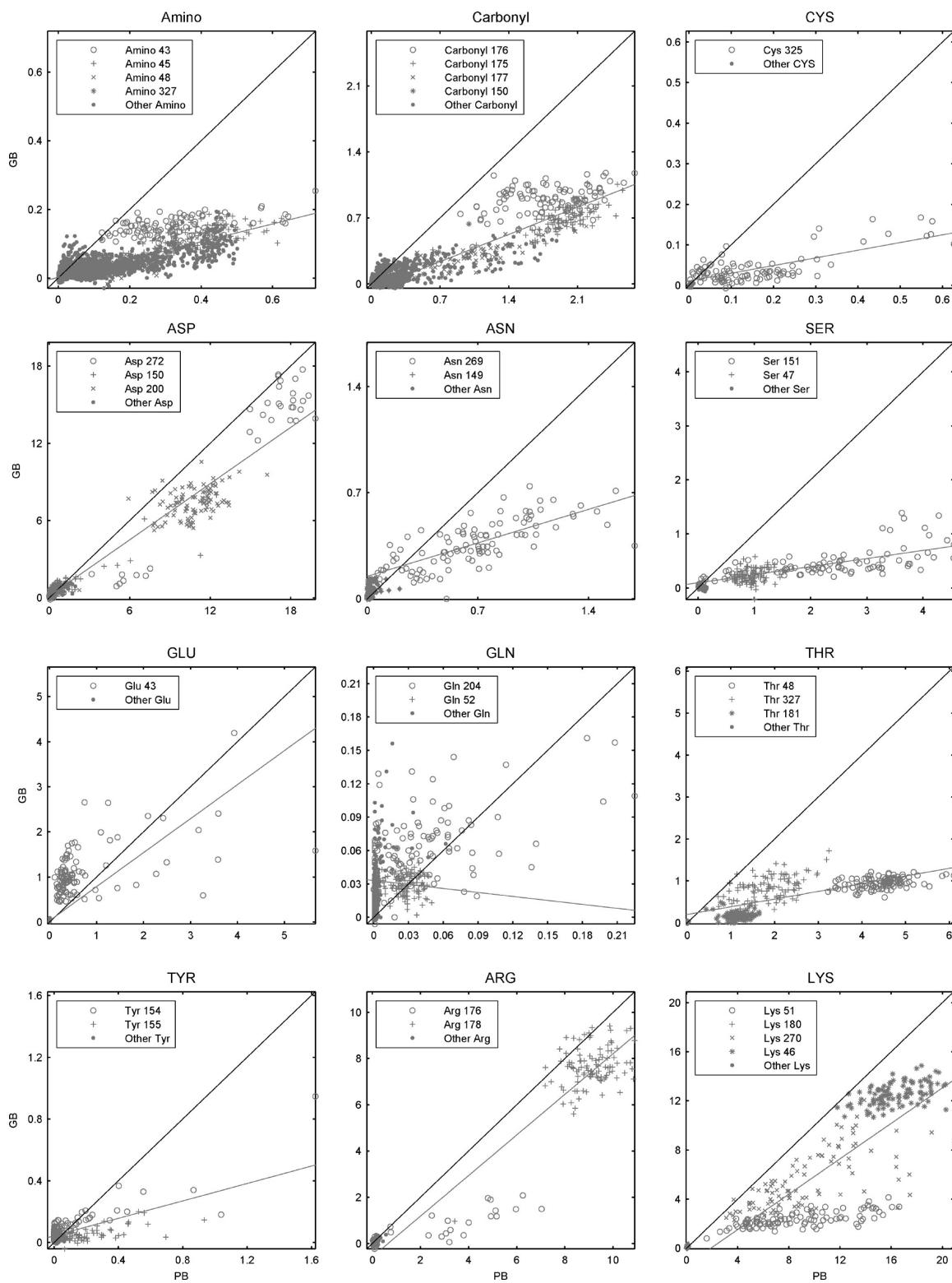


Figure 7. $G_d/\text{GTP}\cdot\text{Mg}$ group desolvation penalties, by type. The data displayed in Figure 3 (second row, center) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no histidines make interactions with the nucleotide, these are also not shown.

two models seems to be greatest. However, even when energetic differences between structures of the same system are captured similarly, there can be striking differences between systems; the difference in the ensemble-averaged mutation energy of Ser 47 for GTP·Mg binding versus GDP binding is computed to be +4.3 kcal/mol by PB (favoring GTP·Mg), while the same difference computed by GB is +13.2 kcal/mol. For Thr 48, on

the other hand, the same difference is -0.6 kcal/mol by PB and -1.8 kcal/mol by GB, a much closer agreement.

In the correlations of GB- and PB-computed mutation energies, there are additional differences between the two nucleotide-binding systems. Arg 178 and Asp 200 both make vastly reduced contributions to GDP binding than to GTP·Mg. When binding GTP·Mg, Arg 178 makes a bidentate interaction

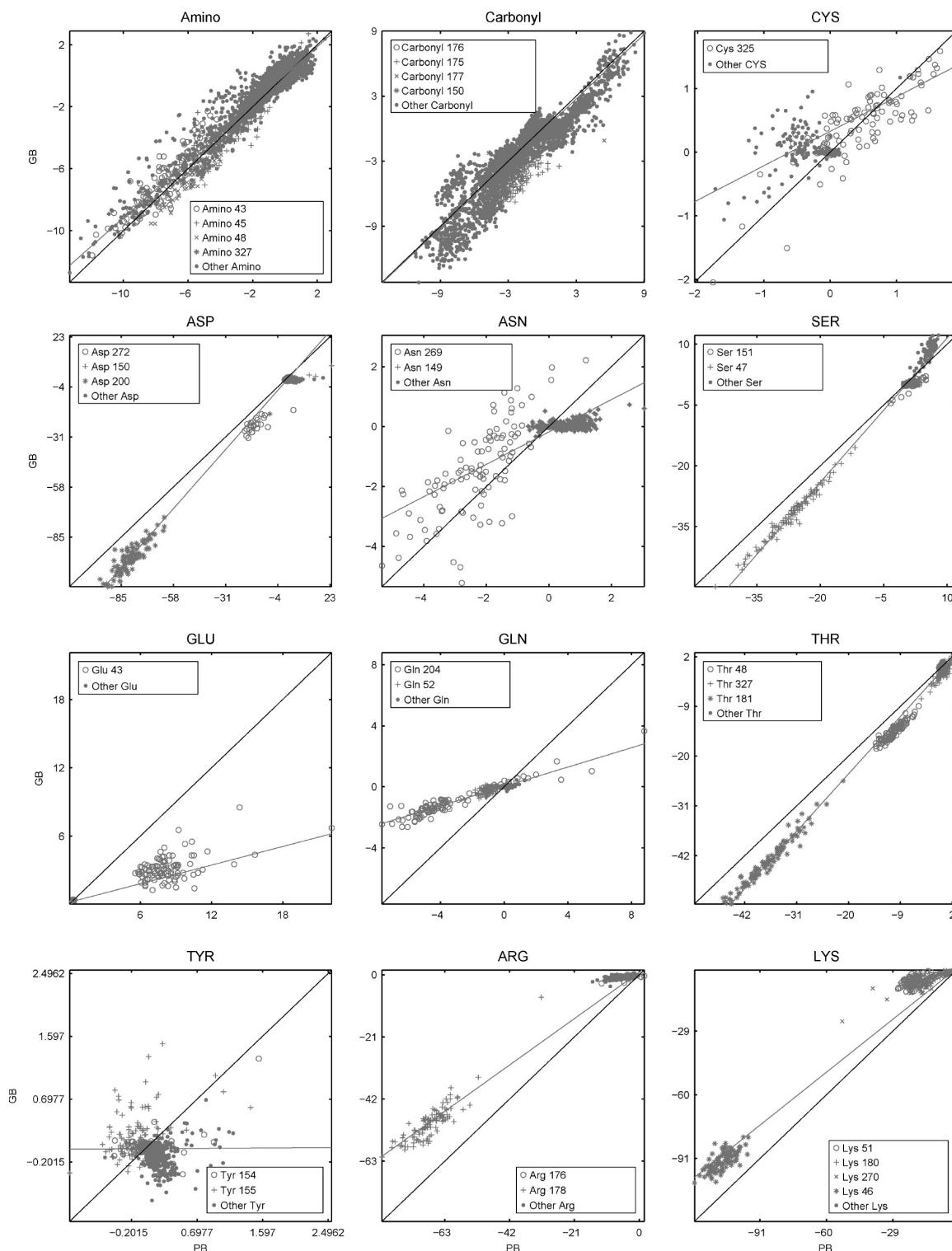


Figure 8. G_α /GTP•Mg group mutation energies, by type. The data displayed in Figure 3 (first row, center) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no histidines make interactions with the nucleotide, these are also not shown.

with the β and γ phosphates and Asp 200 coordinates the magnesium, but with the loss of both the γ phosphate and the associated Mg^{2+} , these interactions are lost (Figure 9). While Arg 178 becomes involved in an intramolecular interaction with Glu 43 (discussed above), Asp 200 becomes partially solvent exposed. As it was these groups that dominated the agreement between GB- and PB-computed mutation energies, the correlations drop greatly. Again, these unfavorable action-at-a-distance

interactions are not well captured. However, while the differences within a structure do not agree well, the gross differences between the states are in general agreement. PB gives an ensemble-averaged mutation energy for GTP•Mg binding of -79.5 kcal/mol for Asp 200 and -67.5 kcal/mol for Arg 178, and corresponding terms for GDP binding of $+5.6$ and -19.5 kcal/mol; with GB, these values are -95.6 and $+1.7$ kcal/mol for Asp 200 and -50.0 and -4.8 kcal/mol for Arg 178

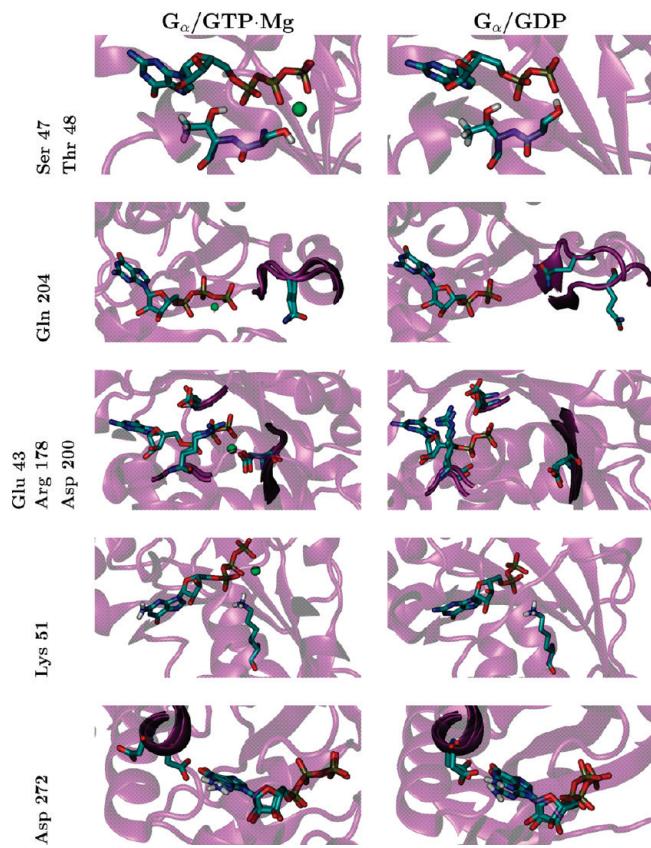


Figure 9. Structure of key G_{α} residues in the $G_{\alpha}/GTP \cdot Mg$ and G_{α}/GDP complexes. Ser 47 makes consistent interactions with the α phosphate in both complexes; Thr 48 coordinates the magnesium ion through the oxygen of the side chain in the complex with $GTP \cdot Mg$ but directly donates a hydrogen bond to the β phosphate in the GDP complex. Gln 204 makes consistent action-at-a-distance electrostatic interactions with $GTP \cdot Mg$ but is structurally variable in the GDP-bound state. Arg 178 shifts from making a bidentate interaction with the β and γ phosphates of GTP to forming an intramolecular salt bridge with Glu 43 when GDP is bound; Asp 200 coordinates the magnesium of $GTP \cdot Mg$ but makes only repulsive, action-at-a-distance interactions with GDP. Lys 51 makes no direct nucleotide contacts but makes favorable action-at-a-distance interactions with both GTP and GDP. While in the GTP -bound state, Asp 272 occupies two structural clusters (one interacting with the nucleotide base and the other not); in the complex with GDP, it makes consistent hydrogen-bonded interactions. Figures generated with VMD.⁷⁸

(complete data in the Supporting Information). Thus, the differences agree to within about 13% for Asp 200 and 6% for Arg 178. Asp 272 shows similar behavior but in the opposing direction and somewhat less pronounced. This group makes direct hydrogen-bonding interactions with the nucleotide base, which persist throughout the simulation with GDP bound, but which is broken for some of the $GTP \cdot Mg$ -bound simulation. The rough differences in desolvation energy between these two states are captured by GB (ensemble averages of 4.3 kcal/mol for GTP binding and 14.1 kcal/mol for GDP binding) in a somewhat similar fashion as by PB (5.2 and 19.3 kcal/mol, respectively), but while PB shows decided variations within each cluster, GB does not. Additionally, while average mutation terms agree reasonably well for GTP binding (-3.5 kcal/mol with PB, -6.4 kcal/mol with GB), there are dramatic differences with GDP binding—a weakly favorable contribution of -1.2 kcal/mol is computed with PB, while GB predicts a strongly favorable contribution of -25.0 kcal/mol. Many of the points for all three cases fall into the cluster of the majority of other

arginines and aspartates lying along a line of very low slope. This low-slope behavior has been discussed previously in the text.

Generalized Born Is an Effective Filter for the Selection of Important Groups. While the results discussed above indicate that there are clearly concerns with using GB as a substitute for PB in a detailed component-based decomposition, an another application is to use GB as a screen for those components expected to be most significant. These terms could then be considered in detail with PB-based calculations, while eliminating the need to perform these expensive calculations on those groups that do not make significant interactions. To test the utility of this, the ensemble-averaged mutation energy as computed by each method was compared for each group (Figure 12, top). The mean energies are highly correlated in all systems, although individual residues show differences of as high as 10 kcal/mol or more.

As a direct measure of the utility of the GB-computed energies as a screen, the enrichment in the most significant components (in mutation energy) computed by PB was measured, using a GB-ranked list of components (Figure 12, middle and bottom). GB performs remarkably well; when a cutoff of 1.36 kcal/mol is used to determine significance (corresponding to a 10-fold difference in affinity relative to the hydrophobic isostere), 100% of the set is found within the top third of GB-ranked components in all systems. Perhaps more remarkably, 60–70% of the set can be found with very little wasted effort; among these top GB-ranked components, no more than three components (out of between 30 and 60) are not in fact significant.

Salt Effects Are Well Correlated between LPB and GB Models, When Significant. The data presented above was all from computations done in the context of 0.0 M ionic strength; in this case, the Poisson–Boltzmann equation reduces to the Poisson equation. To evaluate the effect of this, net binding free energies for all snapshots were additionally computed with a 0.145 M concentration of monovalent ions (NaCl) using both the GB and linearized-PB models; the data are summarized in Table 3 and Figure 13. Inclusion of salt has a relatively small effect on the net binding free energies; in all three systems, the correlation coefficient between the energies computed with and without salt is greater than 99.9%. For nucleotide binding, the rms difference is roughly 0.4 kcal/mol, and the maximum absolute difference is 1.0 kcal/mol or less. In these cases, where the salt effects are relatively small, the salt effects computed by the two methods are only weakly correlated. The effect of salt in the computed protein–protein binding free energy for trimer formation is much more significant; the rms difference over all snapshots is almost 3 kcal/mol, and individual differences as great at 4.5 kcal/mol are seen. However, in this case, the effects are very similar when computed with the LPB and GB models; the differences are correlated with an R^2 value of 0.88, an rms difference of 0.3 kcal/mol, and a maximum difference of only 0.7 kcal/mol. As salt effects are either relatively weak (nucleotide binding) or highly correlated between the two models (trimer formation), the conclusions drawn from the analysis of the salt-free systems may be expected to transfer well to salt-containing calculations.

Alternative Approaches and Directions of Study. The results presented here were computed with a single implementation of the generalized-Born method (GBSW).⁵⁸ In this method, the dielectric boundary between solute and solvent is smoothed, while a sharp boundary is used in the Poisson–Boltzmann approach. To evaluate whether this difference contributes

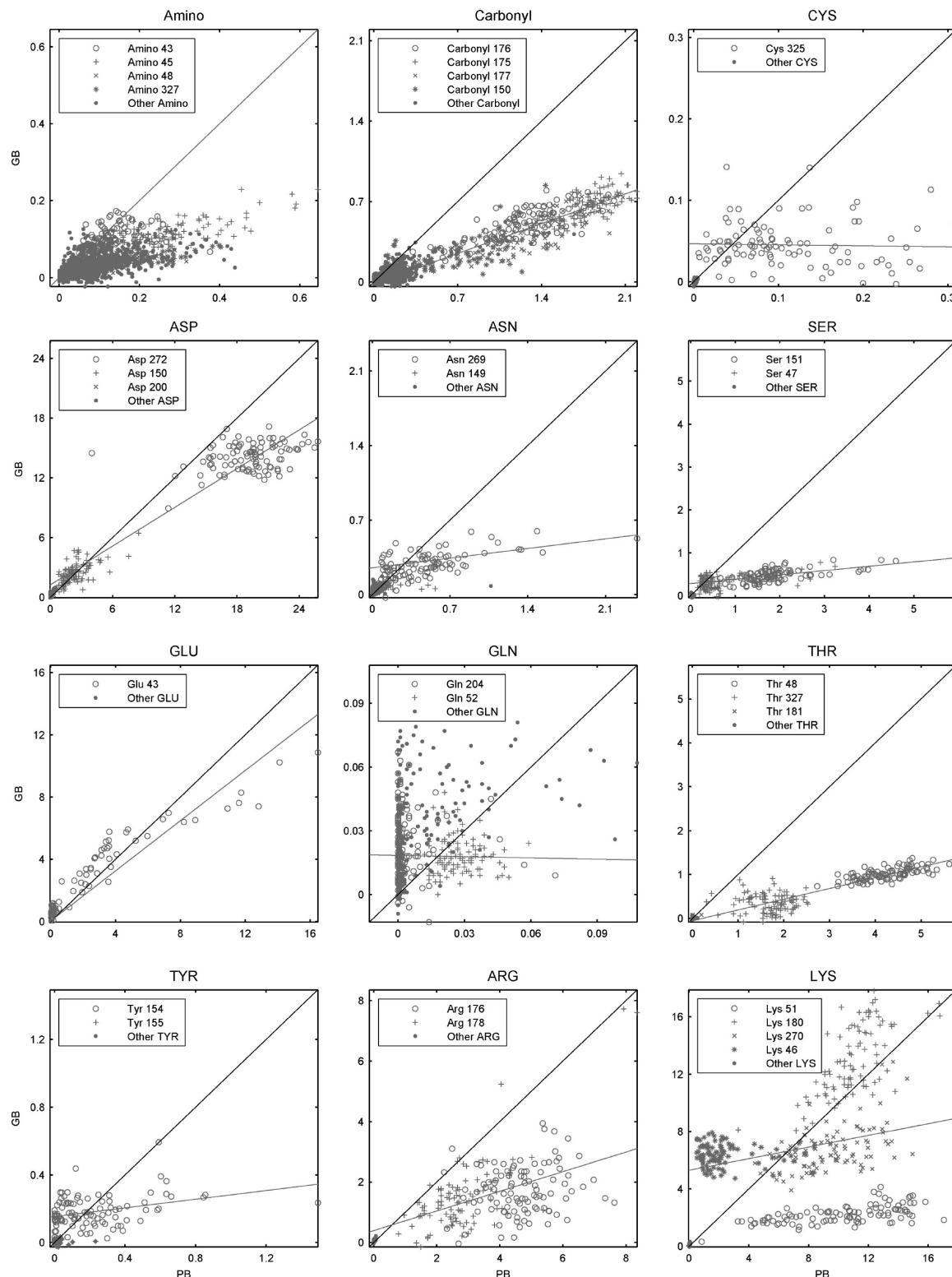


Figure 10. G_d/GDP group desolvation penalties, by type. The data displayed in Figure 3 (second row, right) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no histidines make interactions with the nucleotide, these are also not shown.

significantly to the observed deviations, we repeated all calculations using the GBMV model,³⁴ also within the CHARMM software package. While the results (presented in the Supporting Information) differ in the details, the deviations seen between GBMV and PB are qualitatively the same as those seen between GBSW and PB (it is worth noting, however, that GBMV gave notably poor agreement of net binding free energies). GBMV

differs from GBSW in a number of ways in addition to the use of an abrupt dielectric boundary, and thus, these results suggest that these deviations are not a peculiarity of one GB implementation. A number of alternatives are also available,^{32,57} and these approaches may have somewhat different behaviors. However, the *overall* energetics computed with the two models agree fairly well, and only specific interactions are poorly

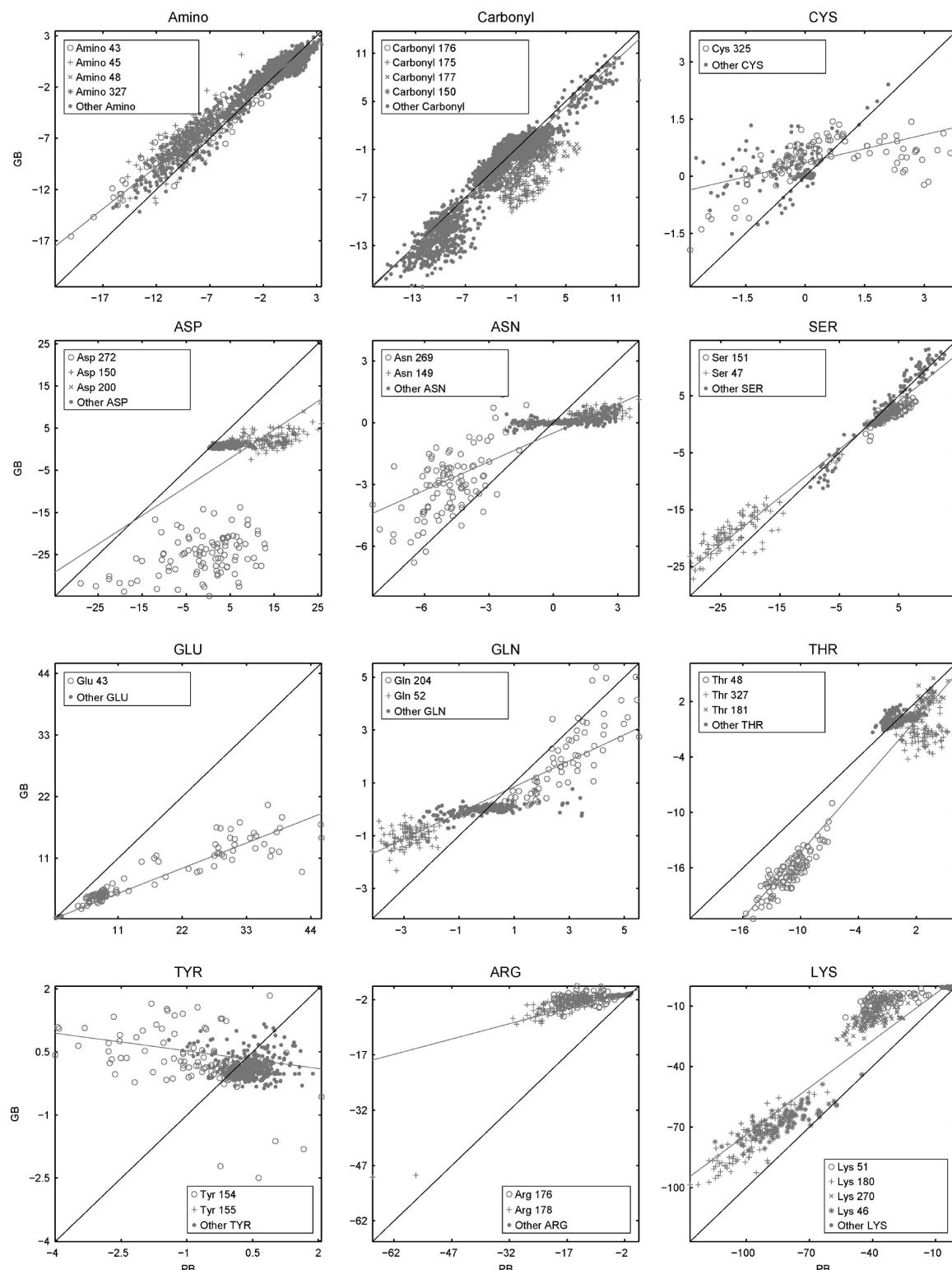


Figure 11. G_0/GDP group mutation energies, by type. The data displayed in Figure 3 (first row, right) are shown, further broken down by group type. In each case, individual residues making significant energetic contributions are uniquely marked. Nonpolar groups (aliphatic and aromatic amino-acid side chains) do not make significant electrostatic contributions and thus are not shown; as no histidines make interactions with the nucleotide, these are also not shown.

captured. Thus, it seems unlikely that other approaches would not show similar deviations, if not exactly the same differences.

Deficiencies in GB-based methods have been noted before. For example, Onufriev et al. have noted the sensitivity of GB-based methods to the method by which effective Born radii are computed, and demonstrated that the use of the so-called “perfect” radii computed from solution of the PB equation results

in much better performance.⁴¹ Unfortunately, the use of these perfect radii involves a solution of the PB equation for each atom in the system; in effect, calculation of the full Φ matrix, which is more than an order of magnitude more costly than performing a residue-by-residue decomposition with PB. While the use of a single reference structure to compute the effective Born radii may reduce costs, the results presented here clearly

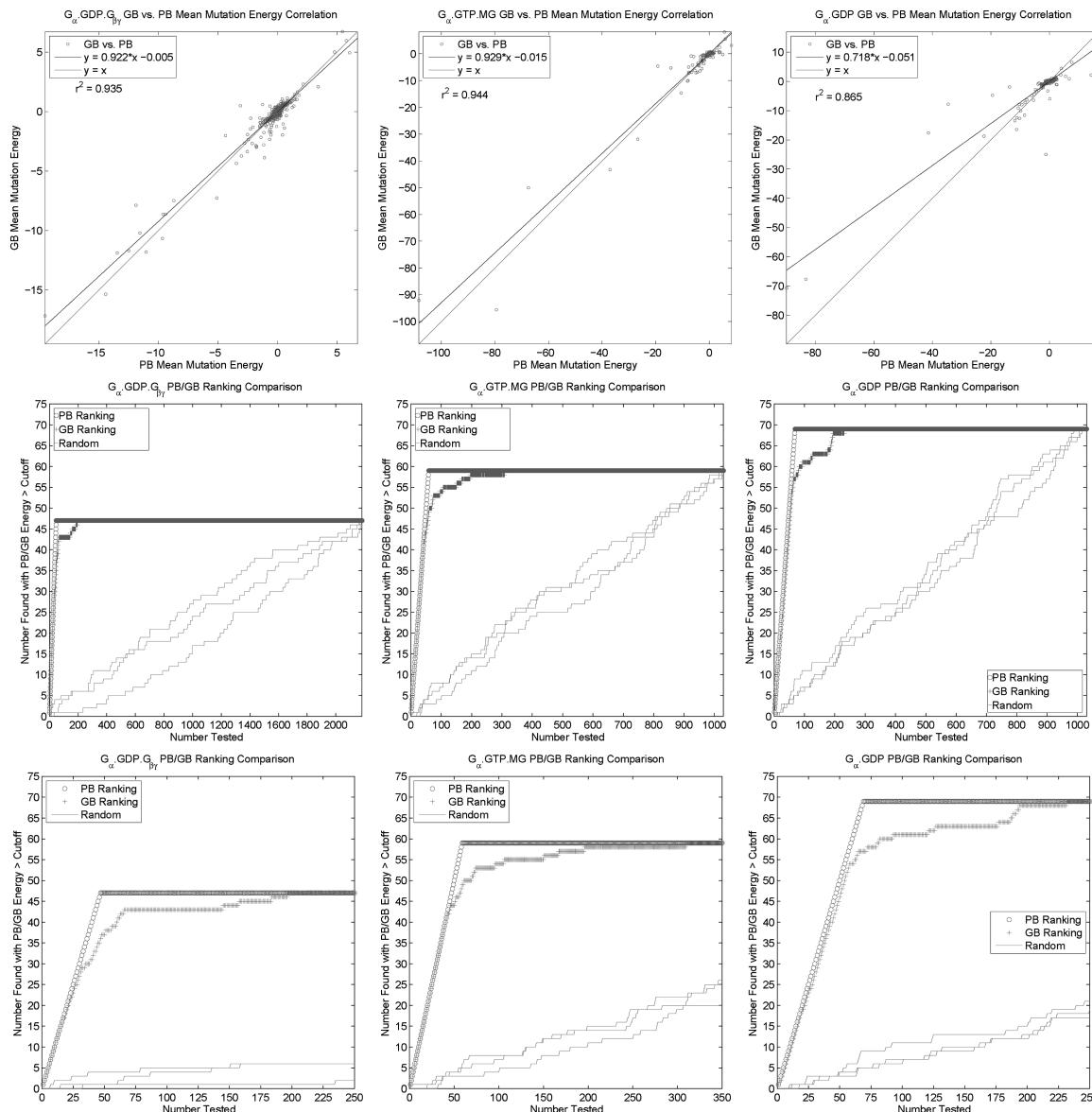


Figure 12. Performance of GB in selecting most significant residues by PB. The utility of GB in screening for those components that are most significant by PB is displayed in two ways for each system: (left) $G_{\alpha}^{GDP}/G_{\beta\gamma}$; (center) $G_{\alpha}/GTP \cdot Mg$; (right) G_{α}/GDP . Top: The ensemble-averaged mutation energies for each component computed with GB are plotted against the same terms computed with PB. Middle and Bottom: The cumulative number of components with PB-computed mutation energies of greater than 1.36 kcal/mol in magnitude are plotted against the number of components tested, with components ranked by PB, GB, or randomly. The bottom panels show the same data as the middle panels, with the x -axis scaled for increased resolution.

TABLE 3: Salt Effects on Binding Energies

system	PB ^a		GB ^a		GB vs PB ^b		
	0.145 vs 0.0 M		0.145 vs 0.0 M		0.145 M – 0.0 M		
	rms	max	rms	max	R^2	rms	max
$G_{\alpha}^{GDP}/G_{\beta\gamma}$	2.94	4.54	2.78	3.97	0.88	0.29	0.71
$G_{\alpha}/GTP \cdot Mg$	0.37	0.81	0.45	0.80	0.42	0.27	0.62
G_{α}/GDP	0.46	1.07	0.71	1.12	0.16	0.67	1.53

^a Root-mean-square (rms) and maximum absolute difference in net binding energy, in kcal/mol, computed with 0.145 M or with 0.0 M NaCl. ^b Correlation coefficient, rms, and maximum absolute difference (in kcal/mol) in the effect of salt ($\Delta G_{0.145M} - \Delta G_{0.0M}$) computed with GB or PB.

demonstrate large variations in group solvation energies with conformation; as these are dominated by individual atomic solvation terms, effective Born radii computed from a single structure would not be able to capture this variation.

Brooks and co-workers have noted that there can be an imbalance between solute–solvent and intersolute interactions with the generalized-Born model,⁴⁰ and as methods are improved to better capture this, it will be worthwhile examining their ability to reproduce individual component energies. Bardhan has also recently described an alternative approach based on a boundary-integral formulation.^{74,75} Interestingly, this method was demonstrated to better reproduce the eigenvectors of the solvation matrix (Φ) than did a generalized-Born approach, and thus, the performance of this method in a system-wide component analysis will also be interesting to evaluate.

Conclusions

A detailed comparison of the use of both generalized-Born and linearized-Poisson–Boltzmann models in performing a component analysis of the contribution of various groups and energetic terms to binding free energies has been performed.

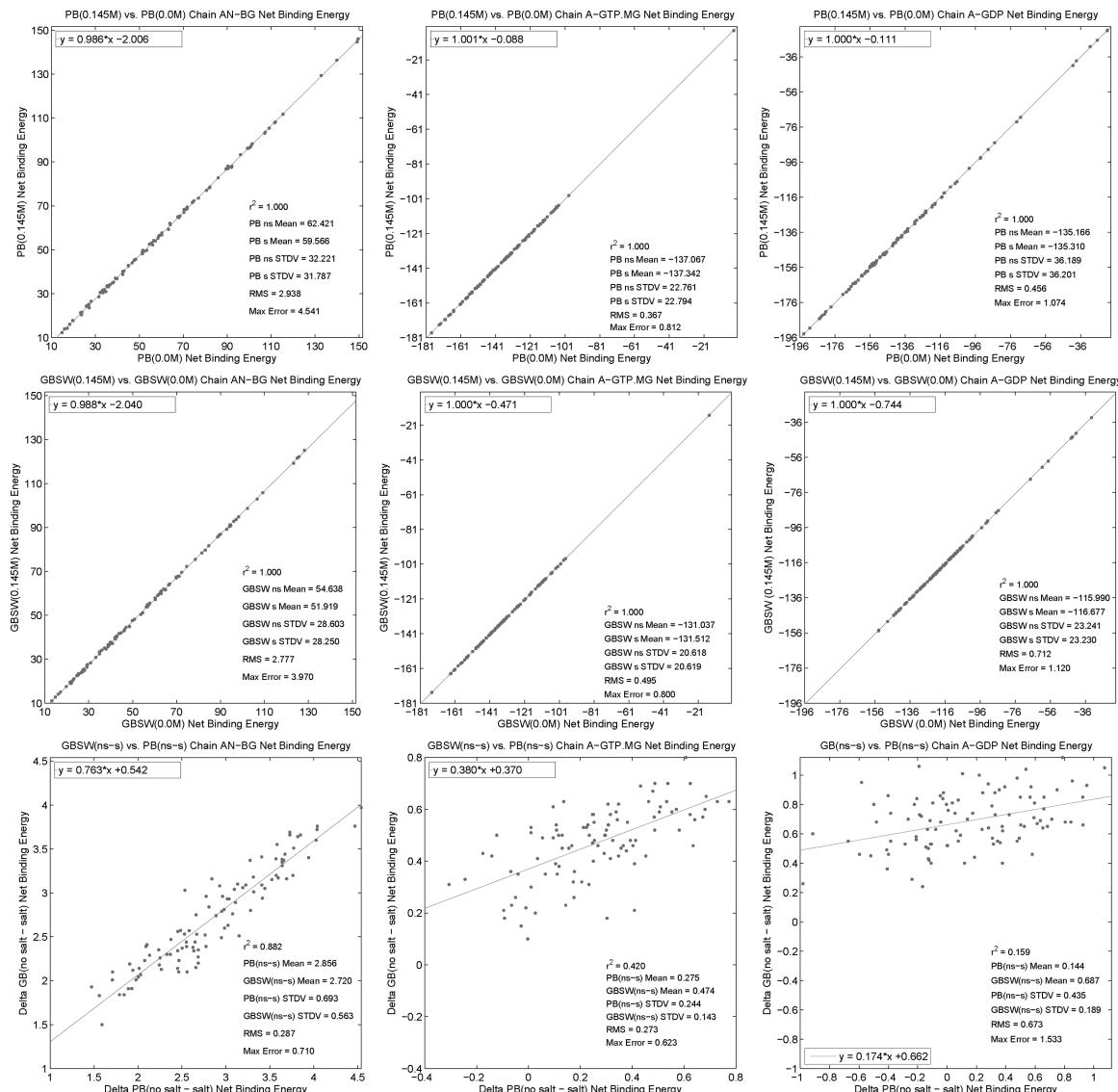


Figure 13. Effects of salt on net electrostatic binding free energies. The effects of including a physiological salt concentration of 0.145 M on the net binding free energies computed with both the LPB and GB models are shown for all three systems studied: (left) $G_{\alpha}^{GDP}/G_{\beta}$; (center) $G_{\alpha}/GTP \cdot Mg$; (right) G_{α}/GDP . Top: The linearized-Poisson–Boltzmann model with 0.145 M NaCl is compared to the Poisson model with 0.0 M ionic strength. Middle: The GB model with salt is compared to the same without. Bottom: The differences in net binding energies computed with and without salt are compared between the two models.

Association of the components of the heterotrimeric G-protein were used as a test case; three binding reactions were considered, including the formation of a protein–protein complex (the G-protein heterotrimer) as well as the binding of guanosine-5'-di- and triphosphate to the monomeric G_{α} subunit.

The results indicate that in many cases the relative net contributions of individual groups are captured well by the GB model, as measured by the difference in free energy of binding relative to a hydrophobic isostere—what we term mutation free energy. However, this agreement is far from universal, and a significant subset of groups show linear correlations between GB- and PB-computed mutation free energies with slopes much lower than unity. In a few cases, the differences are much more dramatic, with essentially no correlation between the results of the two methods. Individual energetic terms, such as the cost of desolvating a particular group upon binding, tend to be captured much more poorly by the GB-model, as compared to the results from PB. Cancellation of some of this error leads to better results for net contributions, as for total binding free energies. Furthermore, even when relative energetics are

captured relatively well (as indicated by a strong correlation of GB- and PB-computed results), individual values computed with GB often deviate by over 10 kcal/mol.

While universal rules are not clearly present, several observations can be made. First, in general, GB seems to capture PB-based energetics reasonably well for hydrogen-bonded interactions, particularly at buried sites. In contrast, for interactions that act over moderate distances, partially screened by solvent, GB-computed results are often lower in magnitude than those computed by PB; as “action-at-a-distance” interactions of this type have recently been identified as an attractive class of interactions to target in design,^{76,77} the poor performance of GB in this regime should be noted. Finally, GB does not seem to capture the same sensitivity to structural variation of the energetics of buried charged groups as does the PB model. These data clearly indicate a need to take care in interpreting the results of energetic decompositions with a generalized-Born model.

In many cases, however, the primary goal may not be to describe the contribution of every group in a quantitative manner but rather to provide a more qualitative profile of those groups

that contribute most to the affinity. In this application, GB is found to be remarkably accurate, and when used to select groups most likely to be significant under a PB-based model, the GB model performs very well. Given the great difference in computational cost of a full system-wide component analysis performed with GB and with PB, the use of GB as an initial screen to identify important components may thus be well merited. This may be followed by Poisson–Boltzmann-based analysis on this focused set of components, providing a reasonable trade-off between accuracy and speed.

Acknowledgment. D.F.G. would like to thank Bruce Tidor for making the Integrated Continuum Electrostatic (ICE) software package available. D.F.G. also thanks the State University of New York, the College of Engineering and Applied Sciences, and the Department of Applied Mathematics and Statistics for financial support. N.C. was partially supported by a Bridges to the Doctorate fellowship from the NSF-SUNY Alliance for Graduate Education and the Professoriate.

Supporting Information Available: An assessment of the structural stability of the simulations, details of correlations among additional partitionings of the data, and the ensemble-averaged desolvation and mutation free energies for the most significant groups in all systems. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Neves, S. R.; Ram, P. T.; Iyengar, R. *Science* **2002**, *296*, 1636–1639.
- (2) Offermanns, S. *Prog. Biophys. Mol. Biol.* **2003**, *83*, 101–130.
- (3) Cabrera-Vera, T. M.; Vanhaeue, J.; Thomas, T. O.; Medkova, M.; Preininger, A.; Mazzoni, M. R.; Hamm, H. E. *Endocr. Rev.* **2003**, *24*, 765–781.
- (4) Swiss Institute of Bioinformatics (SIB) and European Bioinformatics Institute (EBI). Swiss-Prot. <http://www.ebi.ac.uk/swissprot/>.
- (5) Schmidt, C. J.; Thomas, T. C.; Levine, M. A.; Neer, E. J. *J. Biol. Chem.* **1992**, *267*, 13807–13810.
- (6) Pronin, A. N.; Gautam, N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6220–6224.
- (7) Yan, K.; Kalyanaraman, V.; Gautam, N. *J. Biol. Chem.* **1996**, *271*, 7141–7146.
- (8) Robillard, L.; Ethier, N.; Lachance, M.; Hébert, T. E. *Cell Signalling* **2000**, *12*, 673–682.
- (9) Dingus, J.; Wells, C. A.; Campbell, L.; Cleator, J. H.; Robinson, K.; Hildebrandt, J. D. *Biochemistry* **2005**, *44*, 11882–11890.
- (10) Kuczera, K.; Gao, J.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8481–8485.
- (11) de Bakker, P. I. W.; Hunenberger, P. H.; McCammon, J. A. *J. Mol. Biol.* **1999**, *285*, 1811–1830.
- (12) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 915–920.
- (13) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407–435.
- (14) Gilson, M. A.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047–1069.
- (15) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671–679.
- (16) Gilson, M. K.; Honig, B. H. *Nature* **1987**, *330*, 84–86.
- (17) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1988**, *9*, 327–335.
- (18) Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. *J. Phys. Chem.* **1992**, *96*, 6428–6431.
- (19) Green, D. F.; Tidor, B. *J. Mol. Biol.* **2004**, *342*, 435–452.
- (20) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880–888.
- (21) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Comput. Chem.* **2004**, *25*, 1416–1429.
- (22) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1999**, *8*, 1381–1392.
- (23) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307–327.
- (24) Kuhn, B.; Kollman, P. A. *J. Am. Chem. Soc.* **2000**, *122*, 3909–3916.
- (25) Mardis, K. L.; Luo, R.; Gilson, M. G. *J. Mol. Biol.* **2001**, *309*, 507–517.
- (26) Fujimoto, Y. K.; TerBush, R. N.; Patsalo, V.; Green, D. F. *Protein Sci.* **2008**, *17*, 2008–2014.
- (27) Murray, D.; McLaughlin, S.; Honig, B. *J. Biol. Chem.* **2001**, *276*, 45153–45159.
- (28) Green, D. F.; Tidor, B. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 644–657.
- (29) Sarkar, C. A.; Lowenhaupt, K.; Horan, T.; Boone, T. C.; Tidor, B.; Lauffenburger, D. A. *Nat. Biotechnol.* **2002**, *20*, 908–913.
- (30) Green, D. F.; Dennis, A. T.; Fam, P. S.; Tidor, B.; Jasanooff, A. *Biochemistry* **2006**, *45*, 12547–12559.
- (31) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
- (32) Qui, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005–3014.
- (33) Reddy, M. R.; Erion, M. D.; Agarwal, A.; Viswanadhan, V. N.; McDonald, D. Q.; Still, W. C. *J. Comput. Chem.* **1998**, *19*, 769–780.
- (34) Lee, M. S.; Salsbury, F. R., Jr.; Brooks, C. L., III. *J. Chem. Phys.* **2002**, *116*, 10606–10614.
- (35) Zhu, J.; Shi, Y.; Liu, H. *J. Phys. Chem. B* **2002**, *106*, 4844–4853.
- (36) Sorin, E. J.; Engelhardt, M. A.; Herschlag, D.; Pande, V. S. *J. Mol. Biol.* **2002**, *317*, 493–506.
- (37) Liu, H.-Y.; Kuntz, I. D.; Zou, X. *J. Phys. Chem. B* **2004**, *108*, 5453–5462.
- (38) Chocholoušová, J.; Feig, M. *J. Phys. Chem. B* **2006**, *110*, 17240–17251.
- (39) Lee, M. R.; Sun, Y. *J. Chem. Theory Comput.* **2007**, *3*, 1106–1119.
- (40) Chen, J.-H.; Im, W.; Brooks, C. L., III. *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.
- (41) Onufriev, A.; Case, D. A.; Bashford, D. *J. Comput. Chem.* **2002**, *23*, 1297–1304.
- (42) Hendsch, Z. S.; Jonsson, T.; Sauer, R. T.; Tidor, B. *Biochemistry* **1996**, *35*, 7621–7625.
- (43) Hendsch, Z. S.; Nohale, M. J.; Sauer, R. T.; Tidor, B. *J. Am. Chem. Soc.* **2001**, *123*, 1264–1265.
- (44) Nohale, M. J.; Hendsch, Z. S.; Tidor, B.; Sauer, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3109–3114.
- (45) Lee, L.-P.; Tidor, B. *J. Chem. Phys.* **1997**, *106*, 8681–8690.
- (46) Kangas, E.; Tidor, B. *J. Chem. Phys.* **1998**, *109*, 7522–7545.
- (47) Sulea, T.; Purisima, E. O. *J. Phys. Chem. B* **2001**, *105*, 889–899.
- (48) Jackson, J. D. *Classical Electrodynamics*, 3rd ed.; John Wiley & Sons, Inc.: New York, 1998.
- (49) Roccia, W.; Alexov, E.; Honig, B. *J. Phys. Chem. B* **2001**, *105*, 6507–6514.
- (50) Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1989**, *10*, 386–391.
- (51) Holst, M.; Baker, N.; Wang, F. *J. Comput. Chem.* **2000**, *21*, 1319–1342.
- (52) Altman, M. D.; Bardhan, J. P.; White, J. K.; Tidor, B. *J. Comput. Chem.* **2009**, *30*, 132–153.
- (53) Chong, L. T.; Dempster, S. E.; Hendsch, Z. S.; Lee, L.-P.; Tidor, B. *Protein Sci.* **1998**, *7*, 206–210.
- (54) Kangas, E.; Tidor, B. *Phys. Rev. E* **1999**, *59*, 5958–5961.
- (55) Lee, L.-P.; Tidor, B. *Nat. Struct. Biol.* **2001**, *8*, 73–76.
- (56) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211–226.
- (57) Schaefer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100*, 1578–1599.
- (58) Im, W.; Lee, M. S.; Brooks, C. L., III. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
- (59) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. *Theor. Chem. Acc.* **1999**, *101*, 426–434.
- (60) Research Collaboratory for Structural Bioinformatics (RCSB). Protein Data Bank. <http://www.rcsb.org/pdb/>.
- (61) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735–1747.
- (62) Brünger, A. T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 148–156.
- (63) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (64) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulter, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (65) MacKerell, A. D.; Wiorkiewicz-Kuczera, J.; M.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.
- (66) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (68) Green, D. F.; Kangas, E.; Hendsch, Z. S.; Tidor, B. ICE - Integrated Continuum Electrostatics. MIT, 2000.
- (69) Altman, M. D.; Tidor, B. MultigridPBE - Software for computation and display of electrostatic potentials. MIT, 2003.

- (70) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.
- (71) Banavali, N. K.; Roux, B. *J. Phys. Chem. B* **2002**, *106*, 11026–11035.
- (72) Green, D. F.; Tidor, B. In *Current Protocols in Bioinformatics*; Petsko, G. E., Ed.; John Wiley & Sons, Inc.: New York, 2003; Chapter 8.3.
- (73) Sigel, I. H. *Eur. J. Biochem.* **1987**, *165*, 65–72.
- (74) Bardhan, J. P. *J. Chem. Phys.* **2008**, *129*, 144105.
- (75) Bardhan, J. P.; Knepley, M. G.; Anitescu, M. *J. Comput. Chem.* **2009**, *30*, 104108.
- (76) Joughin, B. A.; Green, D. F.; Tidor, B. *Protein Sci.* **2005**, *14*, 1363–1369.
- (77) Selzer, T.; Albeck, S.; Schreiber, G. *Nat. Struct. Biol.* **2000**, *7*, 537–541.
- (78) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

JP910540Z