Effects of High Temperature on Desolvation Costs of Salt Bridges Across Protein Binding Interfaces: Similarities and Differences between Implicit and Explicit Solvent Models

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · MARCH 2012

Impact Factor: 3.3 · DOI: 10.1021/jp210172b · Source: PubMed

2 AUTHORS:

Reza Salari
Rutgers, The State University of New Jersey
9 PUBLICATIONS **44** CITATIONS
SEE PROFILE

Lillian T Chong
University of Pittsburgh
24 PUBLICATIONS **2,466** CITATIONS
SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL

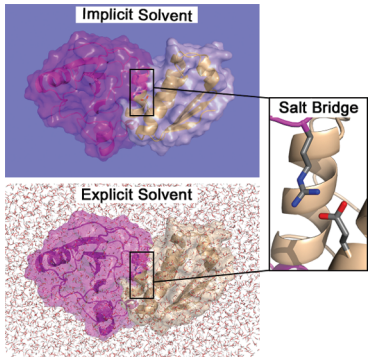
Desolvation Costs of Salt Bridges across Protein Binding Interfaces: Similarities and Differences between Implicit and Explicit Solvent Models

Reza Salari and Lillian T. Chong*

Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

ABSTRACT The prevalence of salt bridges across protein binding interfaces is surprising given the significant costs of desolvating the two charged groups upon binding. These desolvation costs, which are difficult to examine using laboratory experiments, have been computed in previous studies using the Poisson−Boltzmann (PB) implicit solvent model. Here, for the first time, we directly compare the PB implicit solvent model with several explicit water models in computing the desolvation penalties of salt bridges across protein−protein interfaces. We report both overall agreement as well as significant differences between the implicit and explicit solvent results. These differences highlight challenges to be faced in the application of implicit solvent methods.

SECTION Biophysical Chemistry



Protein binding interactions often involve salt bridges, that is, pairs of oppositely charged residues that are within hydrogen-bonding distance. On the basis of theoretical studies, salt bridges are thought to make surprisingly little (or even no) favorable contribution to protein folding or binding due to the significant cost of desolvating the two charged salt bridge partners.[1−5] For efficient computations, these previous studies all used a dielectric continuum solvent model based on the Poisson−Boltzmann (PB) equation. This model, which is the gold standard of implicit solvent models, has been successfully parametrized to reproduce solvation free energies of small molecules determined by either experiment[6] or simulations[7,8] with explicit water molecules. However, the PB model lacks important molecular details of the first solvation shell and a description of nonpolar contributions to solvation.[9] Valuable insights about modeling solvation can therefore be obtained by comparing explicit and implicit solvent calculations.[9−13]

Here, for the first time, we directly compare the PB implicit solvent model with several explicit water models in computing the desolvation penalties of salt bridges across protein−protein binding interfaces. We performed both implicit and explicit solvent calculations on all 14 salt bridges across the binding interfaces of four protein−protein complexes (Figure S1, Supporting Information) that were identified by a previous study as having a wide range of desolvation penalties.[5] We computed the desolvation penalty for each salt bridge upon binding relative to its hydrophobic isostere, that is, a hypothetical mutant version that has all partial charges on the salt bridge side chains set to 0; this desolvation penalty is reported as $\Delta\Delta G_{solv}$. In the explicit solvent calculations, this desolvation penalty was computed using thermodynamic integration techniques (see Methods). As done in previous theoretical studies, we focused on rigid binding, with the unbound conformations of the proteins being identical to the corresponding bound conformations. To circumvent convergence problems associated with net-charged systems in explicit solvent calculations,[14] we represented the unbound state, in both the implicit and explicit solvent calculations, with proteins separated by 30 Å (between their centers of mass) and simultaneously turned off the charges of the oppositely charged side chains of the salt bridge; this was done in both the unbound and bound states of the proteins.

In order to directly compare the solvation thermodynamics of the implicit and explicit water models, it was essential to keep the proteins completely rigid, even in the explicit solvent molecular dynamics (MD) simulations. A direct comparison also required that we fix all other parameters common to the two approaches to ensure that they remained absolutely identical, that is, protein coordinates, atomic charges and radii (OPLS-AA/L force field),[15] box volume, and temperature. MD simulations were performed with periodic boundary conditions and a PME treatment of long-range electrostatics.[16] Periodic boundary conditions were also employed in the PB calculations, implicitly including long-range electrostatic interactions with all periodic images. Implicit and explicit solvent calculations were performed using the DelPhi[17] and GROMACS[18] software packages, respectively. Three different water models were explored in the explicit solvent calculations, TIP3P,[19] TIP4P,[19] and SPC/E.[20] To represent the boundary between the low-dielectric protein region and high-dielectric solvent region in the implicit solvent calculations, we focused primarily on the molecular

Received Date: August 3, 2010
Accepted Date: September 5, 2010
Published on Web Date: September 13, 2010

DOI: 10.1021/jz1010863 | J. Phys. Chem. Lett. 2010, 1, 2844–2848

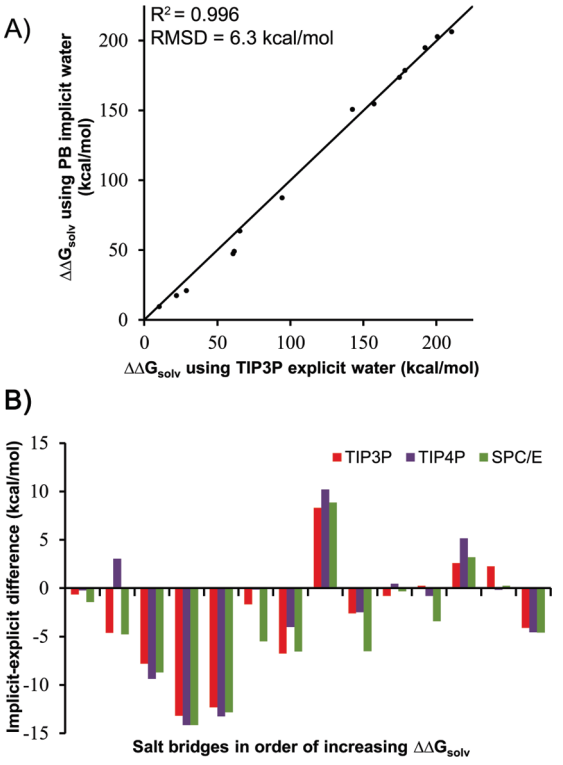


Figure 1. (A) Comparison of implicit and explicit solvent models for computing desolvation penalties of salt bridges upon protein binding ($\Delta\Delta G_{solv}$). The dielectric boundary in the implicit solvent calculations was represented by the molecular surface of the protein. The diagonal line represents perfect agreement. Error bars are included but difficult to see since they are small (< 2 kcal/mol). (B) Implicit−explicit differences for each salt bridge.

surface of the protein,[21] which is the standard representation; calculations were also performed using the van der Waals surface, which has been proposed as an alternative[22,23] but led to comparable results (see below).

As shown in Figure 1A, the desolvation penalties estimated by implicit solvent calculations are strongly correlated with those from explicit solvent calculations with the TIP3P water model (R^2 = 0.996). An equally strong correlation results when the TIP4P and SPC/E water models are used (R^2 of 0.993 and 0.992, respectively; Figure S2, Supporting Information). The overall agreement between the results from implicit and explicit water models is surprisingly good, given the dramatic differences in their representations of solvent and given the large range of the desolvation penalties (~10 to ~210 kcal/mol). These results provide important reinforcement, therefore, of the widely appreciated utility of Poisson-based calculations for modeling solvation effects in charged, biomolecular systems.

That said, a closer examination of the results reveals significant discrepancies between the implicit and explicit solvent predictions; the rms deviations between the predictions for all salt bridges are 6.3, 6.8, and 7.1 kcal/mol for the TIP3P, TIP4P, and SPC/E water models, respectively, which correspond to relative rms deviations of 5.5, 6.0, and 6.2%, respectively (absolute rms deviation divided by the average $\Delta\Delta G_{solv}$ of the explicit water model).

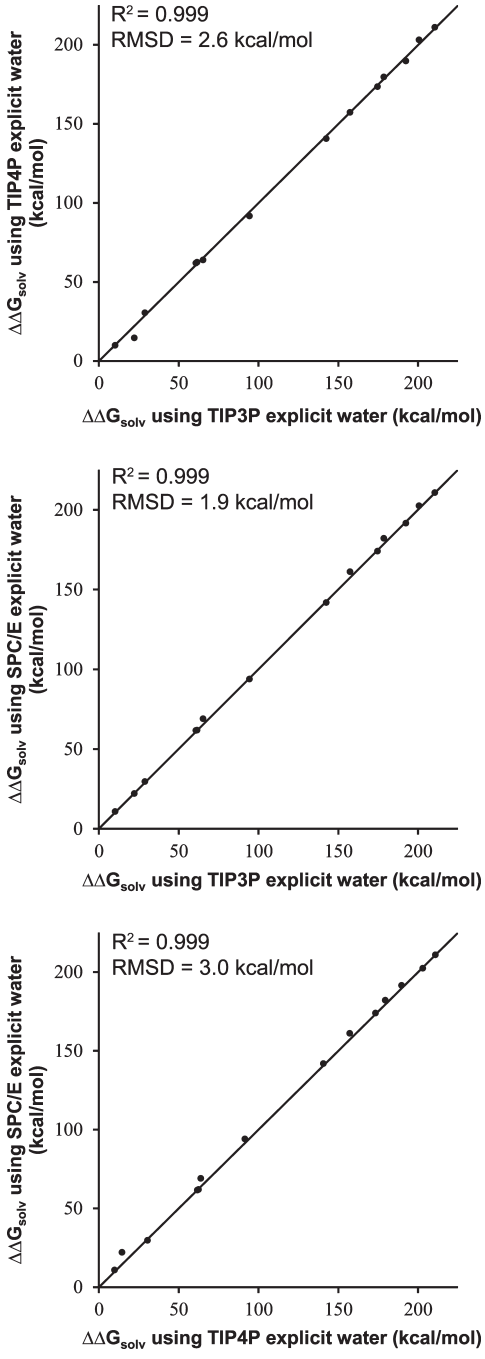


Figure 2. Comparison of explicit solvent models for computing desolvation penalties of salt bridges upon protein binding ($\Delta\Delta G_{solv}$). Diagonal lines represent perfect agreement. Error bars are included but difficult to see since they are small.

Notably, the implicit−explicit discrepancies for individual salt bridges are largely independent of the explicit water model (Figure 1B). Results among the three explicit solvent models are comparable, with rms deviations of 2.6, 1.9, and 3.0 kcal/mol for TIP4P versus TIP3P, SPC/E versus TIP3P, and SPC/E versus TIP4P, respectively (Figure 2). These findings not only provide further confidence in the explicit solvent calculations, they also strongly suggest that the discrepancies

DOI: 10.1021/jz1010863 | J. Phys. Chem. Lett. 2010, 1, 2844–2848

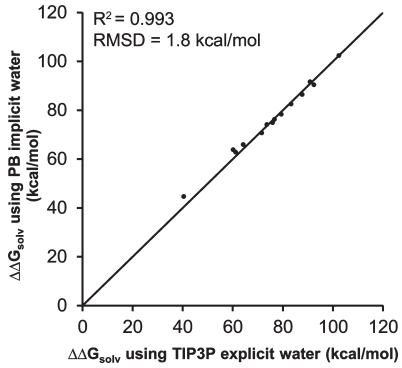


Figure 3. Comparison of implicit and explicit solvent models for computing desolvation penalties of salt bridges upon association in the absence of the protein environment ($\Delta\Delta G_{solv}$). The dielectric boundary in the implicit solvent calculations was represented by the molecular surface of the protein. The diagonal line represents perfect agreement. Error bars are included but difficult to see since they are small.

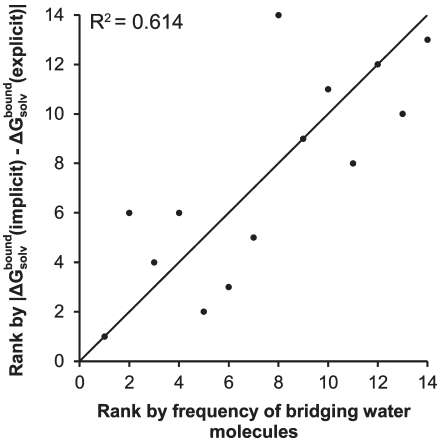


Figure 4. Correlation of the magnitude of implicit−explicit differences in ΔG_{solv}^{bound} versus the probability of observing bridging water molecules during simulations in TIP3P explicit water when the dielectric boundary in the implicit solvent calculations is represented by the molecular surface of the protein. Probabilities were computed from conformations sampled every ps during the 1 ns simulations. Water molecules were defined as bridging if they form hydrogen bonds with both salt bridge partners in their bound state. A hydrogen bond was defined as having a hydrogen−acceptor length of 2.5 Å and a donor−hydrogen−acceptor angle of more than 90°. The diagonal line represents perfect agreement.

reflect key differences between implicit and explicit models of solvation.

To investigate the source of the discrepancies between the implicit and explicit solvent results, we also performed calculations on the same 14 salt bridges in the absence of the protein environment, that is, with the same geometries, but in solution and with the residues capped with acetyl and *N*-methyl groups at the N- and C-termini, respectively. In this second set of explicit solvent calculations, only the TIP3P water model was used. The rms deviation between the implicit and explicit solvent results is significantly reduced from 6.3 to 1.8 kcal/mol when the protein environment is replaced by solvent (Figure 3). It appears, therefore, that the protein environment, and the solvent's response to it, is the primary source of the deviations observed between implicit and explicit solvent calculations.

Representation of the protein environment in the implicit solvent calculations is influenced not only by the protein dielectric constant but also by the dielectric boundary between the protein and solvent regions. In addition to using the molecular surface of the protein to represent the dielectric boundary, which is traced out by a spherical "water" probe with a radius of 1.4 Å, we also tested the van der Waals surface. However, the resulting (implicit) desolvation penalties were found to be comparable to those associated with the molecular surface, with rms deviations of 5.9 kcal/mol from the TIP3P explicit solvent calculations, for example (Figure S2, Supporting Information). Interestingly, although the molecular surface with the current set of atomic radii underestimates the solvation free energies of the salt bridges relative to their hydrophobic isosteres ($\Delta G_{solv}^{(un)bound}$) in the unbound and bound states (Figure S3, Supporting Information), the differences between the two states ($\Delta\Delta G_{solv}$) are underestimated for some of the salt bridges and overestimated for others when compared to explicit solvent calculations (Figure 1A).

To determine why certain salt bridges have larger implicit−explicit differences than others, we plotted these differences versus (a) $\Delta\Delta G_{solv}$ and (b) the percent burial upon binding; both van der Waals and molecular surface implicit solvent results were considered. Only the plot involving the van der Waals surface and percent burial resulted in any correlation (R^2 of 0.320; Figure S4, Supporting Information). We also looked for a correlation between implicit−explicit differences and the involvement of the salt bridge in a "network" where at least one of the charged partners forms another salt bridge;[5,24] however, no correlation was found (Table S2, Supporting Information). We did find a significant correlation between the magnitude of implicit−explicit differences in the solvation free energy of the salt bridge in its bound state relative to its hydrophobic isostere, ΔG_{solv}^{bound}, and the probability of observing "bridging" water molecules in the explicit solvent simulations when the molecular surface was used (Figure 4); no correlation was found when the van der Waals surface was used (Figure S5, Supporting Information). We tried reducing implicit−explicit differences by using a lower solvent dielectric constant for the implicit solvent calculations that is more representative of the explicit solvent models (i.e., 52, the dielectric constant of TIP4P, since this value is the lowest among the explicit solvent models tested),[25] but this lower value had no effect (Table S3, Supporting Information). Finally, we considered reducing implicit−explicit differences by either increasing the protein dielectric constant or scaling the atomic radii in the implicit solvent calculations. However, these approaches would underestimate the desolvation penalties for some salt bridges and overestimate those for others, even in the absence of their protein environments (as is evident in Figure 3).

In closing, we have performed the first direct comparison of implicit and explicit solvent models for use in evaluating free-energy contributions of salt bridges to protein−protein binding. We have demonstrated that the desolvation penalties of salt bridges upon protein binding are of similar magnitudes

DOI: 10.1021/jz1010863 | *J. Phys. Chem. Lett.* **2010**, *1*, 2844–2848

when estimated using implicit and explicit solvent models. Nonetheless, significant discrepancies exist for particular salt bridges. Given that bridging water molecules have been shown to be a source of discrepancies in other studies,[10,11] hybrid implicit−explicit solvent models might be an attractive alternative approach.[26] Since the set of salt bridges[5] studied here highlights challenges to be faced in the application of implicit solvent methods, it might also provide valuable test cases for the development of improved fast solvation models.

METHODS

To directly compare the solvation thermodynamics of the implicit and explicit solvent approaches, we kept the proteins rigid and fixed all parameters common to the approaches to be identical, that is, protein coordinates, atomic charges and radii (OPLS/AA-L force field),[15] box volume, and temperature. To enable a consistent treatment of long-range electrostatics, periodic boundary conditions were employed in both approaches, enabling the use of the PME method[16] for the explicit solvent calculations; all systems were electrically neutral. Details of the protein models are provided in Supporting Information. Desolvation penalties of salt bridges upon protein binding ($\Delta\Delta G_{solv}$) were computed according to the thermodynamic cycle shown in Figure S6 (Supporting Information).

Implicit Solvent Calculations. Implicit solvent calculations were performed using finite difference methods, as implemented in the DelPhi 4.0 software package,[17] to solve the linearized form of the PB equation (which reduces to the Poisson equation in the absence of salt, as in our calculations). In particular, electrostatic contributions to solvation free energies were computed for each wild-type salt bridge and its hydrophobic isostere in the unbound and bound protein states; these contributions were determined by first directly calculating the induced polarization charges and then calculating the interaction between the protein charges and the reaction field due to the polarization charges.[23] The electrostatic contribution to the solvation free energy of the salt bridge relative to its hydrophobic isostere in the unbound or bound state yielded the solvation free energies $\Delta G_{solv}^{unbound}$ or ΔG_{solv}^{bound}, respectively. The desolvation penalty of each salt bridge upon protein binding relative to its hydrophobic isostere was computed using $\Delta\Delta G_{solv} = \Delta G_{solv}^{unbound} - \Delta G_{solv}^{bound}$ (see Figure S6, Supporting Information).

Calculations of each state of the system were performed 14 separate times, with systematic molecular translations on the grid at 25 °C. Results reported are averages of 14 calculations, with uncertainties represented by the standard deviation. Each calculation was carried out for 10000 steps to satisfy a convergence criterion of 0.001 kT/e in the potential. To avoid errors in the dielectric boundary, the OPLS/AA-L radii of polar hydrogen atoms were converted from 0 to the default value of 1.0 Å. To represent the dielectric boundary, we tested both the molecular (default)[21] and van der Waals surfaces of the protein. Consistent with keeping the proteins rigid, a dielectric constant of 1 was used for the protein region; to represent the dielectric properties of water at 25 °C, a dielectric constant of 78.4 was used for the solvent region. A grid resolution of 0.33 Å/(grid units) was used for all protein systems, except for the neuraminidase−antibody complex, which was limited to a slightly lower resolution of 0.37 Å/(grid units) due to its large size. Grid dimensions for the barnase−barstar, growth hormone−receptor, neuraminidase−antibody, and RafRBD−Rap1A complexes were 343 × 343 × 343, 403 × 403 × 403, 479 × 479 × 479, and 361 × 361 × 361, respectively. Each calculation required 1.5−4 CPU hours on a single core of a dual-core 2.6 GHz Opteron node.

Explicit Solvent Calculations. Explicit solvent calculations were performed using the thermodynamic integration approach[27] with explicit solvent MD simulations, as implemented in the GROMACS 4.0.4 software package.[18] In particular, we first computed differences in the overall free energy of each salt bridge relative to its hydrophobic isostere in its unbound and bound states $\Delta G^{(un)bound}$, which is the sum of contributions from both nonbonded protein−protein and protein−solvent interactions, $\Delta G_{protein}^{(un)bound}$ and $\Delta G_{solv}^{(un)bound}$, respectively. Next, to obtain differences in solely the solvation free energies, all nonbonded protein−protein interactions were subtracted from differences in the overall free energies. Finally, the desolvation penalty of each salt bridge upon protein binding relative to its hydrophobic isostere was computed using $\Delta\Delta G_{solv} = \Delta G_{solv}^{unbound} - \Delta G_{solv}^{bound}$.

Differences in the overall free energies of each salt bridge relative to its hydrophobic isostere in its unbound and bound states were computed using the following

$$\Delta G^{(un)bound} = \int_0^1 d\lambda \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_\lambda \qquad (1)$$

where $H(\lambda)$ is the system Hamiltonian as a function of the coupling parameter λ and the brackets represent ensemble averaging at a given λ value; the λ values of 0 and 1 represent the wild-type and hydrophobic isostere versions of the salt bridge, respectively. Separate MD simulations of the proteins (unbound and bound states) were performed at each of the following eight λ values, linearly discharging the side chains of the salt bridge: 0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9, and 1. The trapezoidal method was then used to numerically solve the thermodynamic integral to obtain $\Delta G^{(un)bound}$. Uncertainties in the free energies were derived from sampling errors in $\langle \partial H(\lambda)/\partial\lambda \rangle_\lambda$; errors at each λ value were estimated using block averaging,[28] as implemented in the g_analyze utility of GROMACS.[18]

MD simulations were performed with explicit solvent (TIP3P,[19] TIP4P,[19] or SPC/E)[20] in the *NVT* ensemble, with the number of atoms in the unbound and bound states of each system enforced to be exactly the same (see Supporting Information). Proteins were kept rigid throughout the simulations using the GROMACS "frozen" option, setting velocities of all protein atoms to 0. Real space electrostatic interactions were truncated at 10 Å, while the long-range components of these interactions were calculated using the PME method[16] with periodic boundary conditions, a spline order of 6, Fourier spacing of 1.0 Å, and relative tolerance of 10^{-6} between long- and short-range energies. van der Waals interactions were switched off smoothly between 8 and 9 Å. Each λ simulation was performed for 1 ns at constant temperature (25 °C) and volume. Prior to each λ simulation, the solvent was equilibrated in two stages, (1) 10 ps at constant temperature (25 °C) and

DOI: 10.1021/jz1010863 | J. Phys. Chem. Lett. 2010, 1, 2844–2848

volume and (2) 100 ps at constant temperature (25 °C) and pressure (1 atm). The Langevin thermostat (frictional constant of 1 ps^{-1}) and a weak Berendsen barostat[29] (coupling time constant of 5 ps) were used to maintain constant temperature and pressure, respectively. A 2 fs time step was used for all simulations. Each λ simulation required 1−6 CPU days on a dual-quad core 2.66 GHz Xeon node.

SUPPORTING INFORMATION AVAILABLE Full details of protein models; Figures S1−S6; and Tables S1−S3. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: ltchong@pitt.edu.

ACKNOWLEDGMENT This work was supported by NSF CAREER Award MCB-0845216 to L.T.C. and the University of Pittsburgh's A&S Fellowship to R.S. We thank the University of Pittsburgh's CMMS for use of its Linux cluster.

REFERENCES

(1) Novotny, J.; Sharp, K. A. Electrostatic Fields in Antibodies and Antibody/Antigen Complex. *Prog. Biophys. Mol. Biol.* **1992,** *58*, 203–224.

(2) Hendsch, Z. S.; Tidor, B. Do Salt Bridges Stabilize Proteins? A Continuum Electrostatic Analysis. *Protein Sci.* **1994,** *3*, 211–226.

(3) Elcock, A. H. The Stability of Salt Bridges at High Temperatures: Implications for Hyperthermophilic Proteins. *J. Mol. Biol.* **1998,** *284*, 489–502.

(4) Hendsch, Z. S.; Tidor, B. Electrostatic Interactions in the GCN4 Leucine Zipper: Substantial Contributions Arise from Intramolecular Interactions Enhanced on Binding. *Protein Sci.* **1999,** *8*, 1381–1392.

(5) Sheinerman, F. B.; Honig, B. On the Role of Electrostatic Interactions in the Design of Protein−Protein Interfaces. *J. Mol. Biol.* **2002,** *318*, 161–177.

(6) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994,** *98*, 1978–1988.

(7) Nina, M.; Beglov, D.; Roux, B. Atomic Radii for Continuum Electrostatics Calculations Based on Molecular Dynamics Free Energy Simulations. *J. Phys. Chem.* **1997,** *101*, 5239–5248.

(8) Swanson, J. M. J.; Adcock, S. A.; McCammon, J. A. Optimized Radii for Poisson−Boltzmann Calculations with the AMBER Force Field. *J. Chem. Theory Comput.* **2005,** *1*, 484–493.

(9) Wagoner, J. A.; Baker, N. A. Assessing Implicit Models for Nonpolar Mean Solvation Forces: The Importance of Dispersion and Volume Terms. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103*, 8331–8336.

(10) Zhang, L. Y.; Gallicchio, E.; Friesner, R. A.; Levy, R. M. Solvent Models for Protein−Ligand Binding: Comparison of Implicit Solvent Poisson and Surface Generalized Born Models with Explicit Solvent Simulations. *J. Comput. Chem.* **2001,** *22*, 591–607.

(11) Yu, Z.; Jacobson, M. P.; Josovitz, J.; Rapp, C. S.; Friesner, R. A. First-Shell Solvation of Ion Pairs: Correction of Systematic Errors in Implicit Solvent Models. *J. Phys. Chem. B* **2004,** *108*, 6643–6654.

(12) Tan, C.; Yang, L.; Luo, R. How Well Does Poisson−Boltzmann Implicit Solvent Agree with Explicit Solvent? A Quantitative Analysis. *J. Phys. Chem. B* **2006,** *110*, 18680–18687.

(13) Thomas, A. S.; Elcock, A. H. Direct Observation of Salt Effects on Molecular Interactions through Explicit-Solvent Molecular Dynamics Simulations: Differential Effects on Electrostatic and Hydrophobic Interactions and Comparisons to Poisson−Boltzmann Theory. *J. Am. Chem. Soc.* **2006,** *128*, 7796–7806.

(14) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. Removal of Pressure and Free Energy Artifacts in Charged Periodic Systems via Net Charge Corrections to the Ewald Potential. *J. Chem. Phys.* **1998,** *108*, 3017–3020.

(15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001,** *105*, 6474–6487.

(16) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995,** *103*, 8577–8593.

(17) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995,** *268*, 1144–1149.

(18) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008,** *4*, 435–447.

(19) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79*, 926–935.

(20) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987,** *91*, 6269–6271.

(21) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. Rapid Grid-Based Construction of the Molecular Surface and the Use of Induced Surface Charge to Calculate Reaction Field Energies: Applications to the Molecular Systems and Geometric Objects. *J. Comput. Chem.* **2002,** *23*, 128–137.

(22) Dong, F.; Vijayakumar, M.; Zhou, H.-X. Comparison of Calculation and Experiment Implicates Significant Electrostatic Contributions to the Binding Stability of Barnase and Barstar. *Biophys. J.* **2003,** *85*, 49–60.

(23) Dong, F.; Zhou, H.-X. Electrostatic Contribution to the Binding Stability of Protein−Protein Complexes. *Proteins* **2006,** *65*, 87–102.

(24) Kumar, S.; Nussinov, R. Salt Bridge Stability in Monomeric Proteins. *J. Mol. Biol.* **1999,** *293*, 1241–1255.

(25) Hess, B.; van der Vegt, N. F. A. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B* **2006,** *110* (35), 17616–17626.

(26) Woo, H.-J.; Dinner, A. R.; Roux, B. Grand Canonical Monte Carlo Simulations of Water in Protein Environments. *J. Chem. Phys.* **2004,** 6392–6400.

(27) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935,** *3*, 300–313.

(28) Bishop, M.; Frinks, S. Error Analysis in Computer Simulations. *J. Chem. Phys.* **1987,** *87*, 3675–3676.

(29) Berendsen, H.; Postma, J.; van Gunsteren, W.; DiNola, A.; Haak, J. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984,** *81*, 3684–3690.

Desolvation Costs of Salt Bridges Across Protein Binding Interfaces: Similarities and Differences between Implicit and Explicit Solvent Models

*Reza Salari and Lillian T. Chong**

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260

* To Whom Correspondence should be addressed. Email: ltchong@pitt.edu

RECEIVED DATE (to be automatically inserted after your manuscript is accepted if required according to the journal that you are submitting your paper to)

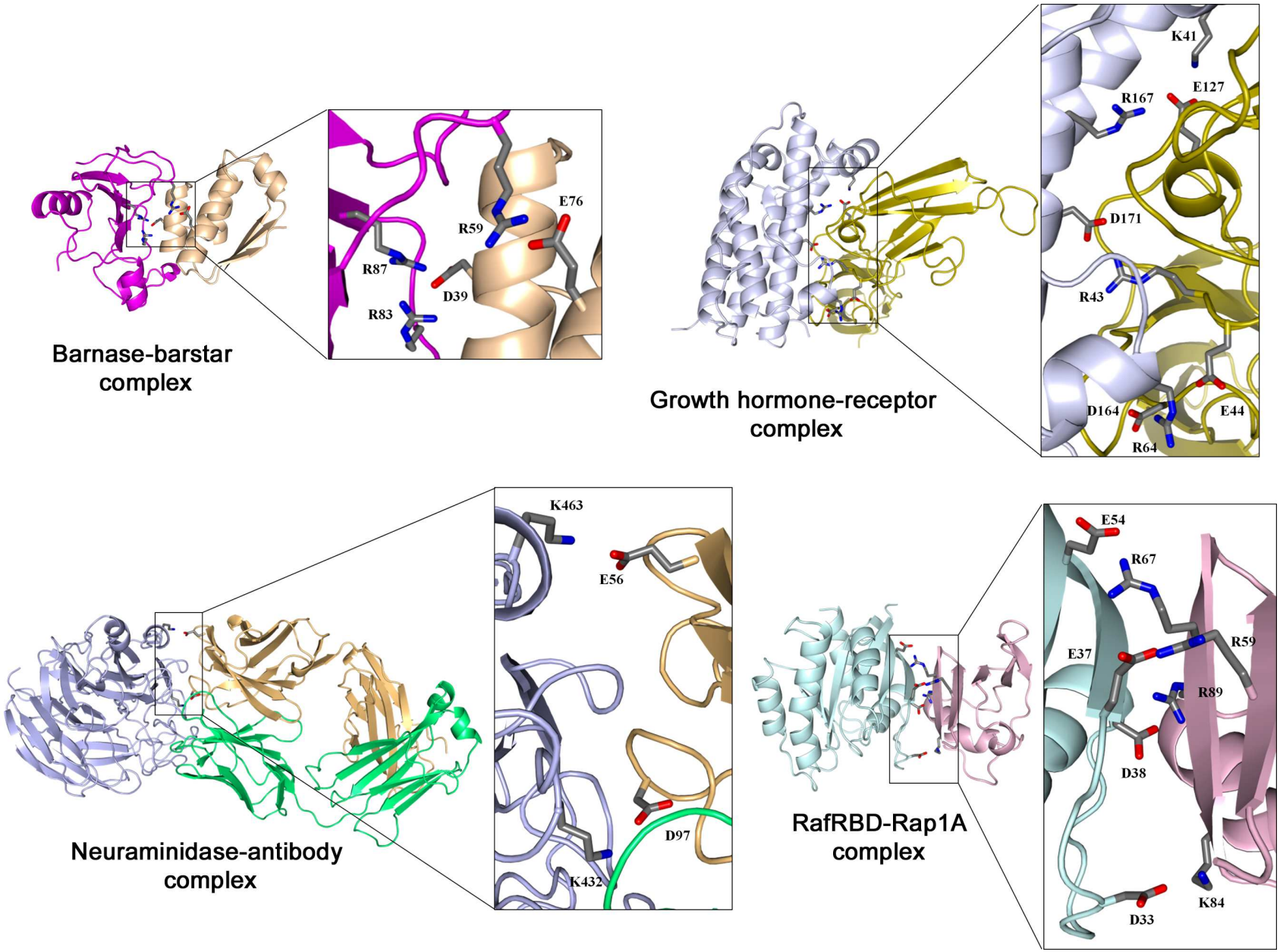


Figure S1. Locations of salt bridges across the binding interfaces of each protein-protein complex in this study.

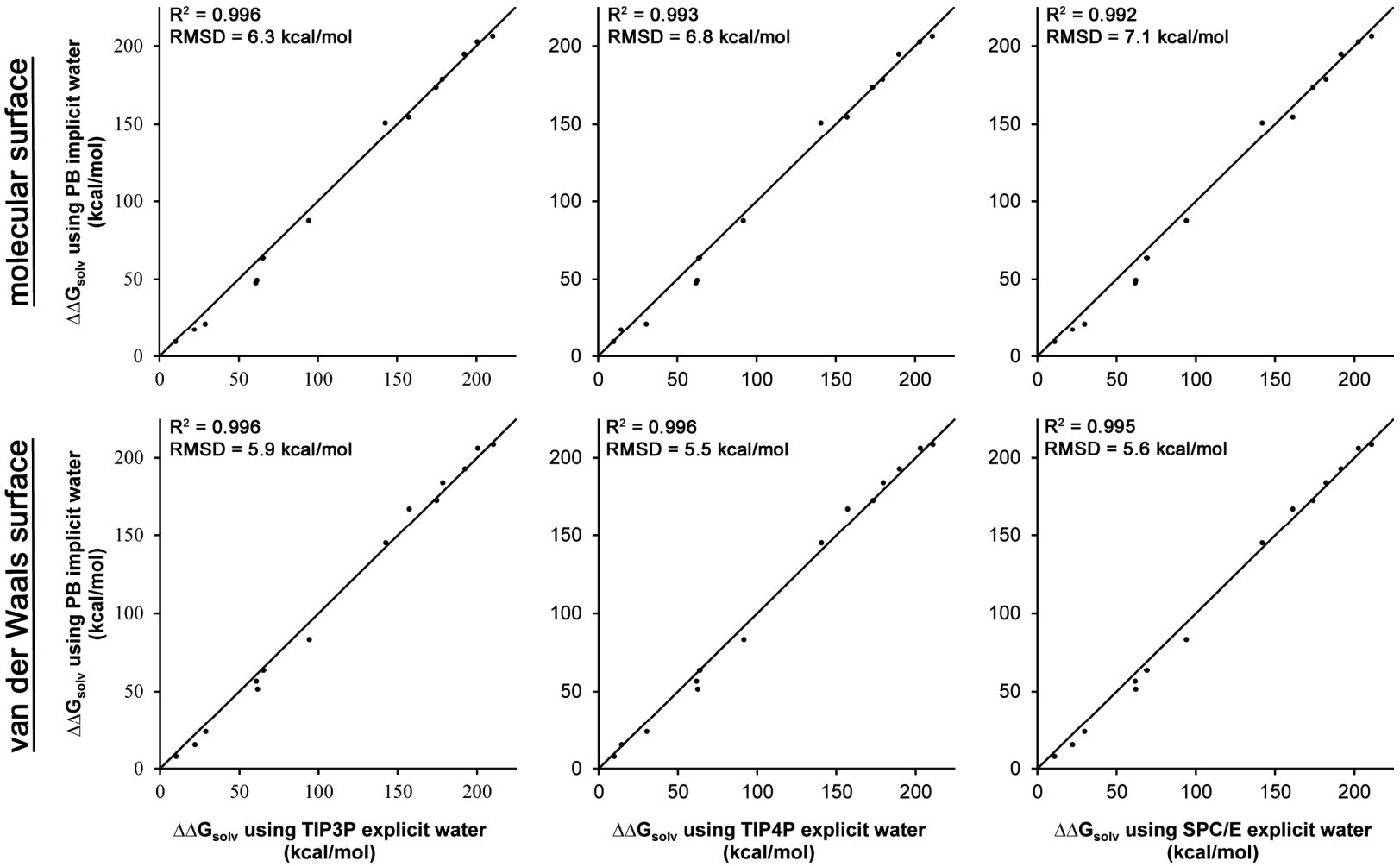


Figure S2. Comparison of implicit vs. explicit solvent models for computing desolvation penalties of salt bridges upon protein binding ($\Delta\Delta G_{solv}$). In the implicit solvent calculations, two different representations of the dielectric boundary were tested: the molecular surface of the protein (top row) and the van der Waals surface of the protein (bottom row). Diagonal lines represent perfect agreement. Error bars are included, but difficult to see since they are small.

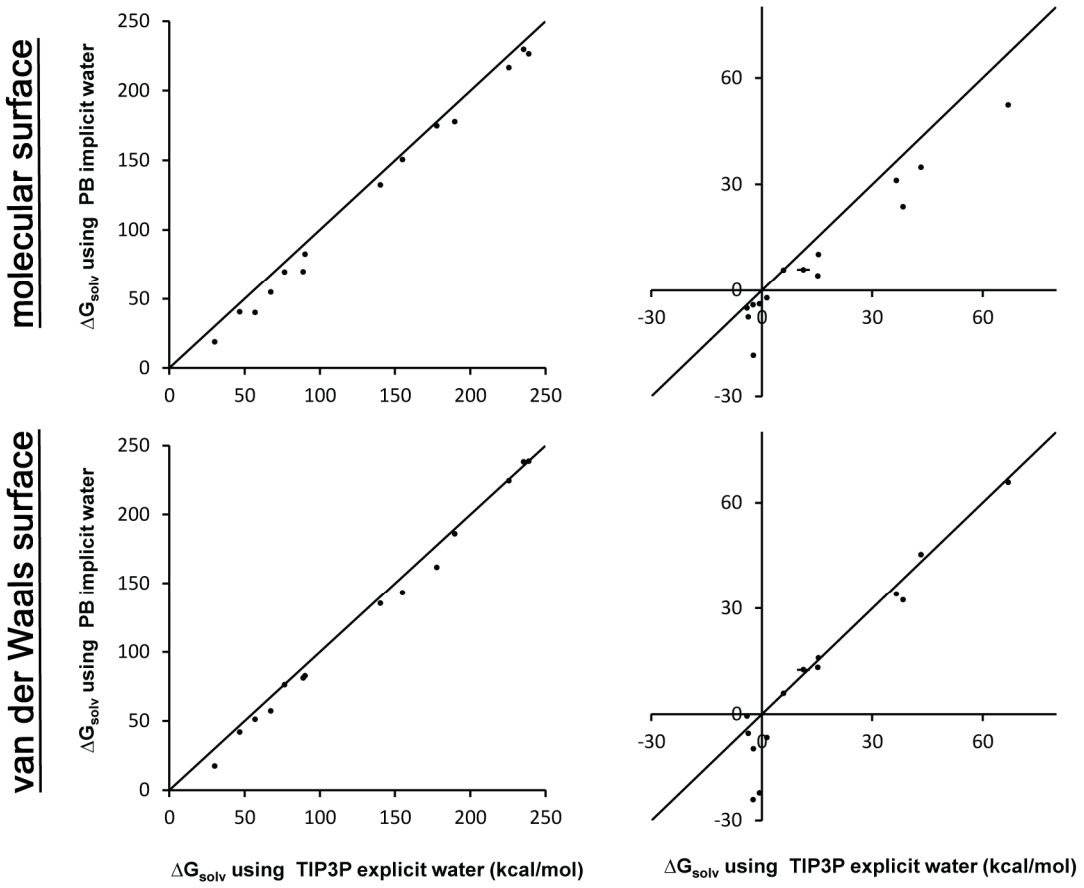


Figure S3. Comparison of implicit vs. explicit solvent models for computing the solvation free energies of salt bridges (ΔG_{solv}) in their unbound (left) and bound states (right). In the implicit solvent calculations, both the molecular (top row) and van der Waals surfaces (bottom row) were tested as representations of the dielectric boundary. The diagonal lines represent perfect agreement. Error bars are included, but difficult to see since they are very small.

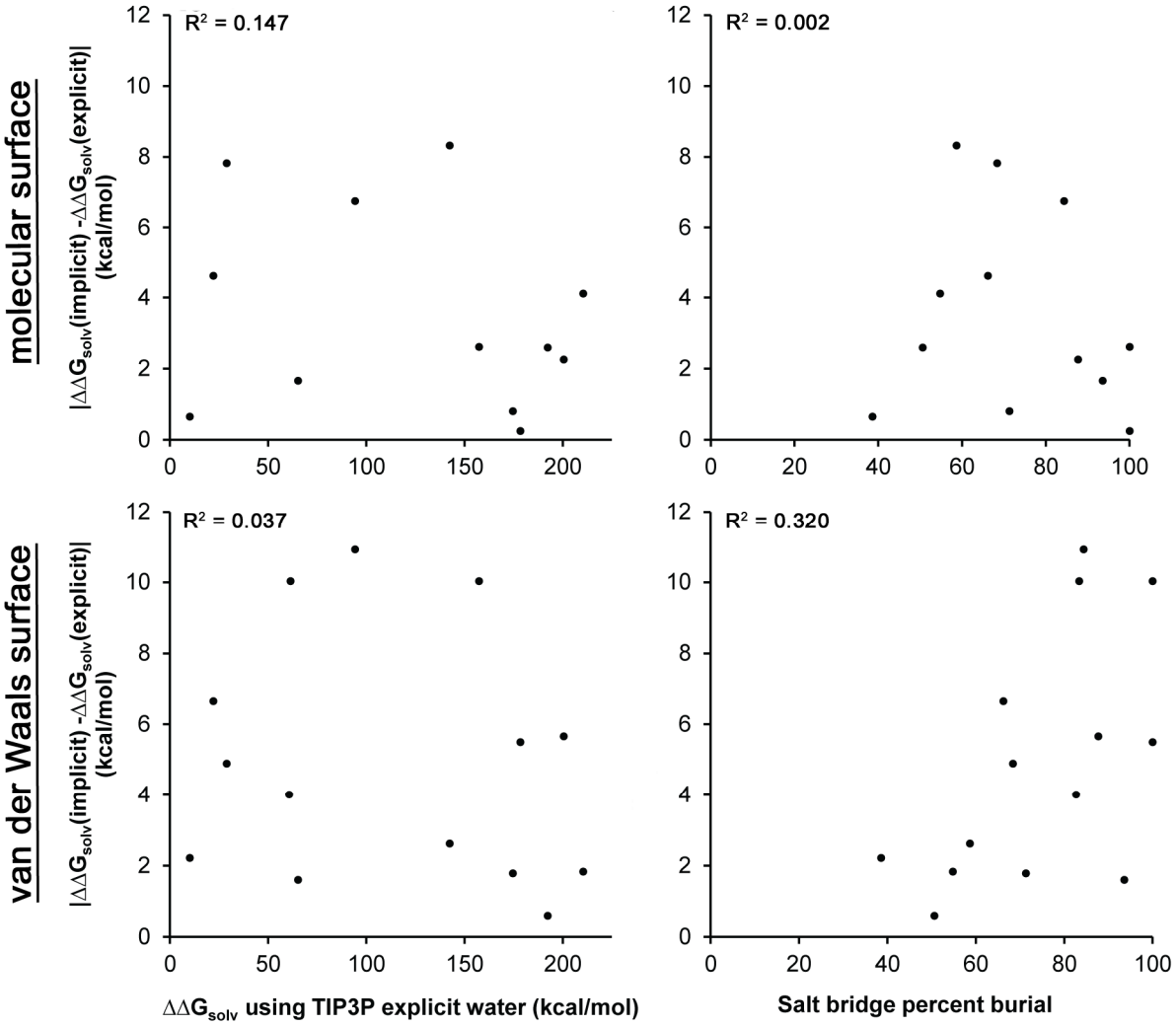


Figure S4. Correlation of the magnitudes of implicit-explicit differences in $\Delta\Delta G_{solv}$ vs. $\Delta\Delta G_{solv}$ (left) and percent burial of salt bridges (right). Implicit solvent calculations were performed separately using the molecular (top row) and van der Waals surfaces (bottom row). The TIP3P water model was used for the explicit solvent calculations.

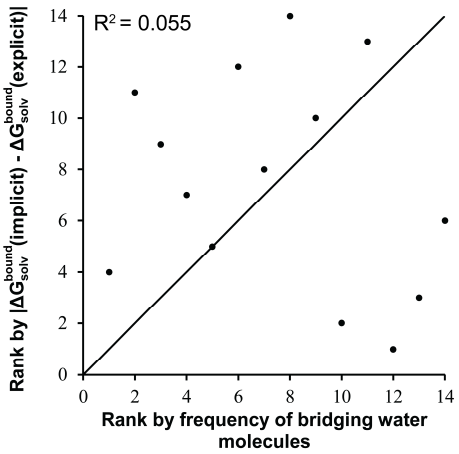


Figure S5. Correlation of the magnitude of implicit-explicit differences in ΔG_{solv}^{bound} vs. probability of observing bridging water molecules during simulations in TIP3P explicit water when the dielectric boundary in the implicit solvent calculations is represented by the van der Waals surface.

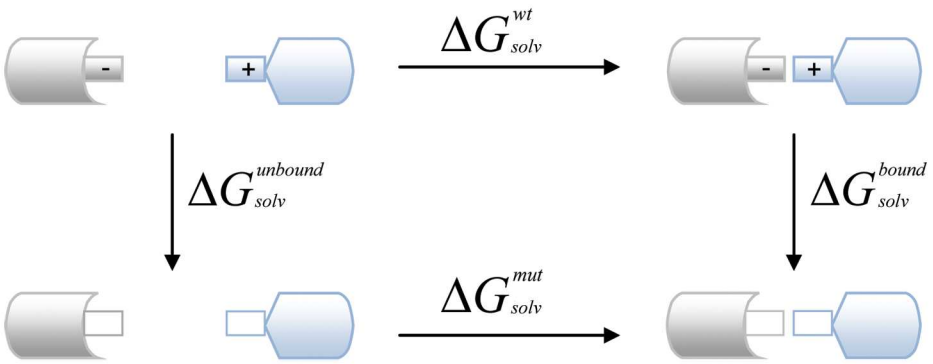


Figure S6. Thermodynamic cycle used for computing the desolvation penalty of a salt bridge upon protein binding relative to its hydrophobic isostere ($\Delta\Delta G_{solv}$). The wild-type salt bridge and its "mutant" hydrophobic isostere are represented by the filled and empty rectangles, respectively. The $\Delta\Delta G_{solv}$ of each salt bridge was computed using the vertical arrows of the thermodynamic cycle in Equation (2), which is much more straighforward to evaluate than Equation (1) as it circumvents the need to simulate the diffusional association of the proteins:

$$\Delta\Delta G_{solv} = \Delta G_{solv}^{wt} - \Delta G_{solv}^{mut} \quad (1)$$

$$\Delta\Delta G_{solv} = \Delta G_{solv}^{unbound} - \Delta G_{solv}^{bound} \quad (2)$$

where ΔG_{solv}^{wt} and ΔG_{solv}^{mut} are the desolvation penalties of the salt bridge upon binding to form the wild-type and mutant complexes, respectively; $\Delta G_{solv}^{unbound}$ and ΔG_{solv}^{bound} are the solvation free energies of the wild-type unbound and bound states, respectively, relative to the corresponding mutant states.

Table S1. Computed desolvation penalties of salt bridges upon protein binding ($\Delta\Delta G_{solv}$) using implicit and explicit solvent models. In the implicit solvent calculations, two different representations of the dielectric boundary were tested: the molecular surface and van der Waals surface of the protein. Errors for desolvation penalties are reported in parentheses.

		implicit solvent		*explicit solvent*		
complex	*salt Bridge*	*molecular surface*	*van der Waals surface*	*TIP3P*	*TIP4P*	*SPC/E*
Barnase-barstar	R59-E76	173.7 (0.2)	172.8 (0.2)	174.6 (0.2)	173.3 (0.9)	174.1 (0.3)
	R83-D39	178.7 (0.2)	184.0 (0.3)	178.5 (0.3)	179.5 (0.2)	182.1 (0.5)
	R87-D39	154.7 (0.1)	167.3 (0.3)	157.3 (0.3)	157.2 (0.2)	161.2 (0.3)
Growth hormone-receptor	K41-E127	49.2 (0.2)	51.5 (0.2)	61.5 (0.2)	62.4 (0.2)	62.0 (0.4)
	R64-E44	17.4 (0.2)	15.4 (0.3)	22.1 (0.2)	14.4 (0.3)	22.2 (0.4)
	R64-D164	21.0 (0.2)	24.0 (0.3)	28.9 (0.2)	30.4 (0.2)	29.7 (0.3)
	R167-E127	47.5 (0.2)	56.7 (0.3)	60.7 (0.3)	61.7 (0.2)	61.7 (0.3)
	D171-R43	63.6 (0.1)	63.7 (0.2)	65.3 (1.7)	63.8 (1.8)	69.1 (1.9)
Neuraminidase-antibody	K432-D97	87.5 (0.2)	83.3 (0.3)	94.2 (0.3)	91.5 (0.2)	94.0 (0.5)
	K463-E56	9.5 (0.2)	7.9 (0.2)	10.1 (0.2)	9.7 (0.2)	10.9 (0.3)
Raf1-Rap1A	D33-K84	194.9 (0.2)	192.9 (0.2)	192.3 (0.3)	189.7 (0.5)	191.7 (0.4)
	E37-R59	206.3 (0.3)	208.6 (0.2)	210.4 (0.2)	210.9 (0.2)	210.9 (0.3)
	D38-R89	202.8 (0.1)	206.2 (0.2)	200.5 (0.3)	203.0 (0.7)	202.5 (0.3)
	E54-R67	150.8 (0.2)	145.1 (0.2)	142.5 (0.2)	140.6 (0.3)	141.9 (0.4)

Table S2. Implicit-explicit differences in $\Delta\Delta G_{solv}$ for each salt bridge and whether or not the salt bridge is networked. Salt bridges are considered "networked" if at least one of the charged partners forms another salt bridge. In the implicit solvent calculations, both the molecular and van der Waals surfaces were tested as representations of the dielectric boundary. Implicit-explicit differences (with errors in parentheses) are presented in order of increasing magnitude according to implicit solvent results with the molecular surface. Explicit solvent calculations are associated with the TIP3P water model.

salt bridge	$\lvert\Delta\Delta G_{solv}$*(implicit)-* $\Delta\Delta G_{solv}$*(explicit)*$\rvert$ *(kcal/mol)*		*networked?*
	molecular surface	*van der Waals surface*	
R83-D39	0.2 (0.3)	5.5 (0.4)	X
K463-E56	0.7 (0.3)	2.2 (0.3)	
R59-E76	0.8 (0.2)	1.8 (0.2)	
D171-R43	1.7 (1.7)	1.6 (1.7)	
D38-R89	2.3 (0.3)	5.7 (0.3)	
D33-K84	2.6 (0.3)	0.6 (0.3)	
R87-D39	2.6 (0.3)	10.0 (0.4)	X
E37-R59	4.1 (0.3)	1.8 (0.2)	
R64-E44	4.6 (0.2)	6.7 (0.4)	X
K432-D97	6.7 (0.3)	10.9 (0.4)	
R64-D164	7.8 (0.2)	4.9 (0.3)	X
E54-R67	8.3 (0.3)	2.6 (0.3)	
K41-E127	12.3 (0.2)	10.0 (0.2)	X
R167-E127	13.2 (0.3)	4.0 (0.4)	X

Table S3. Effect of the solvent-region dielectric constant on RMS deviations between implicit and explicit predictions of $\Delta\Delta G_{solv}$ for the TIP3P, TIP4P, and SPC/E explicit water models.

	implicit solvent		*RMS deviations from explicit solvent results (kcal/mol)*		
	dielectric boundary	*dielectric constant*	*TIP3P*	*TIP4P*	*SPC/E*
PB	*molecular surface*	*78.4*	6.3	6.8	7.1
		52.0	6.4	6.8	7.3
	van der Waals surface	*78.4*	5.9	5.5	5.6
		52.0	5.5	4.7	5.4

Protein Models

All model building was performed using the GROMACS 4.0.4 software package.[1] Coordinates of heavy atoms for each protein-protein complex were taken from the highest-resolution crystal structure of the complex from the Protein Data Bank: 1BRS (chains A and D) for the barnase-barstar complex,[2] 1C1Y for the complex between the Ras binding domain of c-Raf1 kinase (RafRBD) and Ras homologue, Rap1A,[3] 3HHR (chains A and B) for the complex between human growth hormone and its receptor,[4] and 1NCA for the complex between influenza virus N9 neuraminidase and the NC41 antibody.[5] Hydrogen atoms were added using ionization states present at neutral pH. The N- and C-termini were modeled as charged except for the N-terminus of the human growth hormone receptor, which was capped with an acetyl group. All crystallographic water molecules were removed. In addition, the following ions and molecules that are distant from the binding sites were removed: Mg^{2+} ions, Ca^{2+} ions, and the GTP-analog from the RafRBD-Rap1A complex; the Ca^{2+} ion and sugar molecules from the neuraminidase-antibody complex. The following missing residues, which are all distant from the binding site, were reconstructed using the MODELLER 9v6 software package:[6] residues 1 and 2 of barnase (chain A), residues 64 and 65 of barstar (chain D), residues 149-153 of the human growth hormone (chain A), and residues 57-61, 74, 235 and 236 of the human growth hormone receptor (chain B). As necessary for the PME treatment of long-range electrostatics[7] in the explicit solvent calculations, the net charge of each protein-protein complex was neutralized by introducing the following mutations at the most distant locations from the binding site: E8K and E57K of barstar; E129K of Rap1A; E1K, E60K, D121K, E152K, D174K, and E178K of the human growth hormone receptor; E185K of the NC41 antibody (chain L).

To relieve unfavorable interactions, each protein-protein complex was energy minimized in two stages with the OPLS-AA/L force field[8] in a cubic box of TIP3P water molecules,[9] with each stage involving 1000 steps of steepest descent minimization. During the first stage, the positions of all hydrogens, reconstructed missing residues, and water molecules were minimized while applying harmonic position restraints to the heavy atoms of residues that are present in the crystal structure. During the second stage, minimization was conducted with no position restraints. The resulting, minimized protein structures (after removing all water molecules) were used as the bound states for both implicit and explicit solvent calculations. To generate the unbound states for these calculations, the two proteins in each protein-protein complex were separated by 30 Å between their centers of mass, resulting in a minimum distance of 15 Å between the proteins. For the explicit solvent calculations, the unbound and bound states of each protein-protein complex were solvated in cubic boxes of explicit water (TIP3P,[9] TIP4P,[9] or SPC/E[10]) with identical volumes that allowed for a minimum solute-wall distance of 12 Å. The number of atoms in the unbound and bound states of each system was enforced to be exactly the same by removing extra water molecules in the bound state, which contained more water molecules than the unbound state in all cases. When the TIP3P or SPC/E water model was used, the total number of water molecules in each protein system was the following: 47,753 for the barnase-barstar complex; 56,116 for the RafRBD-Rap1A complex; 77,780 for the growth hormone-receptor complex; and 178,663 for the neuraminidase-antibody complex; when the TIP4P water model was used, the numbers are the following (in the same order): 45,644; 55,685; 78,203; and 176,538.

References

(1) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008,** *4*, 435-447.
(2) Buckle, A. M.; Schreiber, G.; Fersht, A. R. Protein-Protein Recognition: Crystal Structural Analysis of a Barnase-Barstar Complex at 2.0 Ang. Resolution. *Biochemistry* **1994,** *33*, 8878-8889.

(3) Nassar, N.; Horn, G.; Herrmann, C.; Scherer, A.; McCormick, F.; Wittinghofer, A. The 2.2 Å Crystal Structure of the Ras-Binding Domain of the Serine/Threonine Kinase C-Raf1 in Complex with Rapla And a GTP Analogue. *Nature* **1995,** *375*, 554-560.
(4) de Vos, A. M.; Ultsch, M.; Kossiakoff, A. A. Human Growth Hormone and Extracellular Domain of its Receptor: Crystal Structure of the Complex. *Science* **1992,** *255*, 306-312.
(5) Tulip, W. R.; Varghese, J. N.; Laver, W. G.; Webster, R. G.; Colman, P. M. Refined Crystal Structure of the Influenza Virus N9 Neuraminidase-NC41 Fab Complex. *J. Mol. Biol.* **1992,** *227*, 122-148.
(6) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Protein Restraints. *J. Mol. Biol.* **1993,** *234*, 779-815.
(7) Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995,** *103*, 8577-8593.
(8) Kaminski, G.A.; Friesner, R.A.; Tirado-Rives, J.; Jorgensen, W.L. Evaluation and Reparametrization of The OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001,** *105*, 6474-6487.
(9) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79*, 926-935.
(10) Berendsen, H.J.C.; Grigera, J.R.; Straatsma, T.P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987,** *91*, 6269-6271.