Discrete Sampling of the Coulomb Field

The procedures for mapping atoms from a protein into the low-dielectric sphere (LDS) are based on spatial integrals over electrostatic energy densities. In the geometry of the protein, this integral is approximated by sampling points in a series of shells centered at the atom of interest. We place two requirements on our sampling scheme. First, sampled points must provide uniform coverage within each shell. To satisfy this requirement, we use a point distribution based on an icosahedral tessellation of the unit sphere (see ref. 1). Second, each sample point should contribute equally to the discretized integral. The contributions of individual points are controlled through judicious choice of shell widths and point densities. We note that the integrand of Eq. **2** of the manuscript ($\vec{D}^2$) varies as r^{-4} with distance r from the atom of interest. Because surface area increases quadratically with r, we can achieve equal contributions from sampled points by increasing shell width, and decreasing points per shell, linearly with r. Assuming uniform dielectric constant, the contribution of a shell bounded by r_1 and r_2 to the integral in Eq. **2** is

$$\int_{r_1}^{r_2} 4\pi r^2 \frac{C}{r^4}\, dr = 4\pi C \frac{r_2 - r_1}{r_1 r_2} = 4\pi C \frac{\Delta r}{r_{equiv}^2}. \qquad \textbf{[S1]}$$

Therefore, the radius at which sample points are taken for each shell (r_{equiv}) is simply the geometric mean of the radial bounds. Calculations in the manuscript used a set of 24 shells of linearly increasing width that covered the region from each atom's Born radius to a radius of 3R, where R was chosen such that $4/3\pi R^3$ gave the Connolly volume of the protein (2). Five-hundred and fourteen points were sampled in the first shell, and this value decreased linearly with r_{equiv} for each shell.

3-Body Surface Burial

The fine-mapping procedure for pairs of atoms (see *Materials and Methods*) requires knowledge of the fraction of a probe shell l buried within the union of the Born radius of an atom (shell m) and the boundary of the LDS (shell n). This is an example of the general problem of determining the accessible surface area of a collection of spheres. Richmond has given expressions for the solution for an arbitrary number of spheres (3). We present a summary for three shells, using his notation. We consider only the case in which some portion of shell l is intersected by both shells m and n, but neither region of intersection is completely contained within the other. Without loss of generality, we place shell l at the origin, shell m on the z axis, and shell n in the x–z plane. In the following, y coordinates will be omitted, and coordinate pairs will refer to $\{x, z\}$. We take as given the coordinates for each shell: $\{0, 0\}$ for l, $\{0, c_m\}$ for m, and $\{a_n, c_n\}$ for n. Radii for the shells are denoted by ρ_l, ρ_m, and ρ_n. The (positive definite) distances of the center of shells m and n from the origin are denoted d_m and d_n. This information is sufficient to determine the perpendicular distances from the origin of the circles resulting from the intersections of shells m and n with shell l:

$$g_n = \frac{d_n^2 + \rho_l^2 - \rho_n^2}{2d_n} \qquad \textbf{[S2]}$$

and

$$g_m = \frac{d_m^2 + \rho_l^2 - \rho_m^2}{2d_m}, \qquad \textbf{[S3]}$$

as well as the radii of these circles:

$$r_m = \sqrt{(\rho_l^2 - g_m^2)} \qquad \textbf{[S4]}$$

and

$$r_n = \sqrt{(\rho_l^2 - g_n^2)}. \qquad \textbf{[S5]}$$

By applying the Gauss–Bonnet theorem to a collection of spheres, Richmond has shown that $\mathcal{F}$, the fraction of shell l occluded by the pair of shells m and n, can be written as

$$\mathcal{F} = \frac{2\pi - 2\Omega - k_m \Delta s_m - k_n \Delta s_n}{4\pi}. \qquad \textbf{[S6]}$$

Ω is the exterior angle formed by the two circles of intersection where they cross. k_m and k_n are the (constant) geodesic curvatures along the circles of intersection, and Δs is the arc length of each circle that is not enclosed in another shell. Expressions for Ω, k_m, k_n, Δs_m, and Δs_n in terms of the geometric parameters are given by Richmond:

$$\Omega = -\cos^{-1}\left[\left(\frac{c_m}{d_m}\right) \frac{\left(\rho_l^2 c_n - \left(\frac{c_m}{d_m}\right) g_m g_n d_n\right)}{r_m r_n d_n} \right], \qquad \textbf{[S7a]}$$

$$k_m = \frac{g_m}{\rho_l r_m}, \qquad \textbf{[S7b]}$$

$$k_n = \frac{g_n}{\rho_l r_n}, \qquad \textbf{[S7c]}$$

$$\Delta s_m = \begin{cases} 2r_m \cos^{-1}\left(\dfrac{\left(\frac{c_m}{d_m}\right) g_m c_n - g_n d_n}{a_n r_m} \right) & \text{for } a_n \geq 0 \\ 2r_m \left(\pi - \cos^{-1}\left(\dfrac{\left(\frac{c_m}{d_m}\right) g_m c_n - g_n d_n}{a_n r_m} \right) \right) & \text{for } a_n < 0, \end{cases} \qquad \textbf{[S7d]}$$

$$\Delta s_n = \begin{cases} 2r_n \left(\pi - \cos^{-1}\left(\dfrac{\left(\frac{c_m}{d_m}\right) g_m d_n - g_n c_n}{r_n |a_n|} \right) \right) & \text{for } c_m > 0 \\ 2r_n \cos^{-1}\left(\dfrac{\left(\frac{c_m}{d_m}\right) g_m d_n - g_n c_n}{r_n |a_n|} \right) & \text{for } c_m \leq 0. \end{cases} \qquad \textbf{[S7e]}$$

Evaluation of Eqs. **S7a–S7e** from the geometric primitives and substitution into Eq. **S6** yields the fraction of shell l buried in either shell m or n.

Application of Image Transformations to Shell Charges

Definitions. The symbols R, ε_s, and ε_p denote, respectively, the radius of the LDS used in the Tanford–Kirkwood model, the dielectric constant assigned to the solvent, and the dielectric constant assigned to the protein interior. Let $(q)|_{\vec{x}_q}$ denote a point charge of magnitude q located at position $\vec{x}_q$. Let $(\sigma, b_q)|_{\vec{x}_q}$ denote a shell of uniform surface charge σ (total charge $q = 4\pi b_q^2 \sigma$) and radius b_q centered at position $\vec{x}_q$. Let $(\sigma, b_q, \vec{x}_\theta, \theta_a \rightarrow \theta_b)|_{\vec{x}_q}$ denote a polar section of a shell of uniform surface charge σ, radius b_q, polar axis $\vec{x}_\theta$, bounded by the polar angles θ_a and θ_b, and centered at position $\vec{x}_q$. Finally, let $\Phi[\rho(\vec{x})](\vec{x}_{obs})$ denote the electrostatic potential of charge distribution $\rho(\vec{x})$ at position $\vec{x}_{obs}$.

Shells Crossing the LDS Boundary. Consider a shell charge $(\sigma, b_q)|_{\vec{x}_q}$ that crosses the LDS boundary $[|x_q - b_q| < R < (x_q + b_q)]$. Knowledge of several quantities is required for application of image transformations Eqs. **6a–6d** of the manuscript in this case. First, the polar angle θ^* at which $(\sigma, b_q)|_{\vec{x}_q}$ intersects the LDS is evaluated by using the law of cosines (unless otherwise specified, polar angles are defined relative to the polar axis $\vec{x}_q$):

$$\theta^* = \cos^{-1}\left(\frac{R^2 - x_q^2 - b_q^2}{2 x_q b_q} \right). \qquad \textbf{[S8]}$$

Next, we define F_{in}^{LDS} and F_{out}^{LDS} to be the surface fractions of $(\sigma, b_q)|_{\vec{x}_q}$ that lie, respectively, inside and outside of the LDS. By integration of surface area in polar coordinates, these quantities can be evaluated:

$$F_{in}^{LDS} = \frac{1}{2}\left(1 + \frac{(R^2 - b_q^2)}{2b_q x_q} - \frac{x_q}{2b_q}\right);$$

$$F_{out}^{LDS} = \frac{1}{2}\left(1 + \frac{(b_q^2 - R^2)}{2b_q x_q} + \frac{x_q}{2b_q}\right). \quad \textbf{[S9]}$$

Finally, we define $\langle r^{-1}\rangle$ to be the average inverse radius of the portion of $(\sigma, b_q)|_{\vec{x}_q}$ that lies exterior to the LDS. Again by surface integration, $\langle r^{-1}\rangle$ is evaluated as

$$\left\langle \frac{1}{r} \right\rangle = \frac{2}{(R + x_q + b_q)}. \quad \textbf{[S10]}$$

electrostatic potential energy of their interaction *in vacuo*. The property of reciprocity of the image-charge solution for a conducting sphere (4) provides the following relationship:

$$W^{Elec}[\rho_i(\vec{x}),\, I_{R,C}\cdot\rho_j(\vec{x})] = W^{Elec}[I_{R,C}\cdot\rho_i(\vec{x}),\, \rho_j(\vec{x})]. \quad \textbf{[S11]}$$

It follows directly from Eq. **S11** that

$$\Phi[I_{R,C}\cdot(\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs}) = \Phi[(C\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs,inv}), \quad \textbf{[S12]}$$

where $\vec{x}_{obs,inv}$ denotes the inverse point, colinear with $\vec{x}_{obs}$ and the LDS center, at radius R^2/x_{obs}.

Image Shell Solutions. Using the identities described above and Eqs. **6a–6d** of the manuscript, the potential of a shell charge $(\sigma, b_q)|_{\vec{x}_q}$ at position $\vec{x}_{obs}$ inside or outside the LDS can be approximated by summing *in vacuo* the potentials of a set of point charges, shell charges, and shell charge sections as

$$(x_q + b_q) \le R;\, x_{obs} < R\text{: } \left(\frac{\sigma}{\varepsilon_p}, b_q\right)\Big|_{\vec{x}_q}, \left(-\frac{q(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}\frac{R}{x_q}\right)\Big|_{\vec{x}_{inv}}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, R\right)\Big|_{Origin}, \quad \textbf{[S13a]}$$

$$(x_q + b_q) \le R;\, x_{obs} \ge R\text{: } \left(\frac{2q}{(\varepsilon_s + \varepsilon_p)}\right)\Big|_{\vec{x}_q}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, R\right)\Big|_{Origin}, \quad \textbf{[S13b]}$$

$$(x_q - b_q) \ge R;\, x_{obs} > R\text{: } \left(\frac{\sigma}{\varepsilon_s}, b_q\right)\Big|_{\vec{x}_q}, \left(\frac{q(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}\right)\Big|_{\vec{x}_{inv}}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}, R\right)\Big|_{Origin}, \quad \textbf{[S13c]}$$

$$(x_q - b_q) \ge R;\, x_{obs} \le R\text{: } \left(\frac{2q}{(\varepsilon_s + \varepsilon_p)}\right)\Big|_{\vec{x}_q}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}, R\right)\Big|_{Origin}, \quad \textbf{[S13d]}$$

$$(x_q - b_q) < R < (x_q + b_q);\, x_{obs} \le R\text{: } \left(\frac{2\sigma}{(\varepsilon_s + \varepsilon_p)}, b_q\right)\Big|_{\vec{x}_q}, \left(\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}, b_q, \vec{x}_q,\, \theta^* \to \pi\right)\Big|_{\vec{x}_q},$$
$$\left(-\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}\frac{R}{x_{obs}}, b_q, \vec{x}_q,\, \theta^* \to \pi\right)\Big|_{\left(\substack{\vec{x}_q + \vec{x}_{obs} \\ -\vec{x}_{obs,inv}}\right)}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\left[F_{in}^{LDS} + F_{out}^{LDS}\frac{2R}{(R + x_q + b_q)}\right], R\right)\Big|_{Origin}, \quad \textbf{[S13e]}$$

$$(x_q - b_q) < R < (x_q + b_q);\, x_{obs} > R\text{: } \left(\frac{2\sigma}{(\varepsilon_s + \varepsilon_p)}, b_q\right)\Big|_{\vec{x}_q}, \left(-\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, b_q, \vec{x}_q,\, 0 \to \theta^*\right)\Big|_{\vec{x}_q},$$
$$\left(\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_{obs}}, b_q, \vec{x}_q,\, 0 \to \theta^*\right)\Big|_{\left(\substack{\vec{x}_q + \vec{x}_{obs} \\ -\vec{x}_{obx,\,inv}}\right)}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s} \times \left[F_{in}^{LDS} + F_{out}^{LDS}\frac{2R}{(R + x_q + b_q)}\right], R\right)\Big|_{Origin}, \quad \textbf{[S13f]}$$

Inversion Operator Identity. Define $I_{R,C}$ to be the inversion operator that maps a charge of magnitude q at polar coordinates $(r,\, \theta,\, \phi)$ into a charge of magnitude $C(R/r)q$ at polar coordinates $(R^2/r,\, \theta,\, \phi)$. Let $\rho_i(\vec{x})$ and $\rho_j(\vec{x})$ represent two charge distributions that lie entirely inside (or outside) a sphere of radius R, and define $W^{Elec}[\rho_i(\vec{x}),\, \rho_j(\vec{x})]$ to be the electrostatic potential energy of their interaction *in vacuo*. The property of reciprocity of the image-charge solution for a conducting sphere (4) provides the following relationship:

$$W^{Elec}[\rho_i(\vec{x}),\, I_{R,C}\cdot\rho_j(\vec{x})] = W^{Elec}[I_{R,C}\cdot\rho_i(\vec{x}),\, \rho_j(\vec{x})]. \quad \textbf{[S11]}$$

It follows directly from Eq. **S11** that

$$\Phi[I_{R,C}\cdot(\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs}) = \Phi[(C\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs,inv}), \quad \textbf{[S12]}$$

where $\vec{x}_{obs,inv}$ denotes the inverse point, colinear with $\vec{x}_{obs}$ and the LDS center, at radius R^2/x_{obs}.

Image Shell Solutions. Using the identities described above and Eqs. **6a–6d** of the manuscript, the potential of a shell charge $(\sigma, b_q)|_{\vec{x}_q}$ at position $\vec{x}_{obs}$ inside or outside the LDS can be approximated by summing *in vacuo* the potentials of a set of point charges, shell charges, and shell charge sections as

$$(x_q + b_q) \le R;\, x_{obs} < R\text{: } \left(\frac{\sigma}{\varepsilon_p}, b_q\right)\Big|_{\vec{x}_q}, \left(-\frac{q(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}\frac{R}{x_q}\right)\Big|_{\vec{x}_{inv}}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, R\right)\Big|_{Origin}, \quad \textbf{[S13a]}$$

$$(x_q + b_q) \le R;\, x_{obs} \ge R\text{: } \left(\frac{2q}{(\varepsilon_s + \varepsilon_p)}\right)\Big|_{\vec{x}_q}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, R\right)\Big|_{Origin}, \quad \textbf{[S13b]}$$

$$(x_q - b_q) \ge R;\, x_{obs} > R\text{: } \left(\frac{\sigma}{\varepsilon_s}, b_q\right)\Big|_{\vec{x}_q}, \left(\frac{q(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}\right)\Big|_{\vec{x}_{inv}}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}, R\right)\Big|_{Origin}, \quad \textbf{[S13c]}$$

$$(x_q - b_q) \ge R;\, x_{obs} \le R\text{: } \left(\frac{2q}{(\varepsilon_s + \varepsilon_p)}\right)\Big|_{\vec{x}_q}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_q}, R\right)\Big|_{Origin}, \quad \textbf{[S13d]}$$

$$(x_q - b_q) < R < (x_q + b_q);\, x_{obs} \le R\text{: } \left(\frac{2\sigma}{(\varepsilon_s + \varepsilon_p)}, b_q\right)\Big|_{\vec{x}_q}, \left(\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}, b_q, \vec{x}_q,\, \theta^* \to \pi\right)\Big|_{\vec{x}_q},$$
$$\left(-\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_p}\frac{R}{x_{obs}}, b_q, \vec{x}_q,\, \theta^* \to \pi\right)\Big|_{\left(\substack{\vec{x}_q + \vec{x}_{obs} \\ -\vec{x}_{obs,inv}}\right)}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\left[F_{in}^{LDS} + F_{out}^{LDS}\frac{2R}{(R + x_q + b_q)}\right], R\right)\Big|_{Origin}, \quad \textbf{[S13e]}$$

$$(x_q - b_q) < R < (x_q + b_q);\, x_{obs} > R\text{: } \left(\frac{2\sigma}{(\varepsilon_s + \varepsilon_p)}, b_q\right)\Big|_{\vec{x}_q}, \left(-\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}, b_q, \vec{x}_q,\, 0 \to \theta^*\right)\Big|_{\vec{x}_q},$$
$$\left(\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s}\frac{R}{x_{obs}}, b_q, \vec{x}_q,\, 0 \to \theta^*\right)\Big|_{\left(\substack{\vec{x}_q + \vec{x}_{obs} \\ -\vec{x}_{obx,\,inv}}\right)}, \left(-\frac{b_q^2}{R^2}\frac{\sigma(\varepsilon_s - \varepsilon_p)}{(\varepsilon_s + \varepsilon_p)}\frac{1}{\varepsilon_s} \times \left[F_{in}^{LDS} + F_{out}^{LDS}\frac{2R}{(R + x_q + b_q)}\right], R\right)\Big|_{Origin}, \quad \textbf{[S13f]}$$

Inversion Operator Identity. Define $I_{R,C}$ to be the inversion operator that maps a charge of magnitude q at polar coordinates $(r,\, \theta,\, \phi)$ into a charge of magnitude $C(R/r)q$ at polar coordinates $(R^2/r,\, \theta,\, \phi)$. Let $\rho_i(\vec{x})$ and $\rho_j(\vec{x})$ represent two charge distributions that lie entirely inside (or outside) a sphere of radius R, and define $W^{Elec}[\rho_i(\vec{x}),\, \rho_j(\vec{x})]$ to be the electrostatic potential energy of their interaction *in vacuo*. The property of reciprocity of the image-charge solution for a conducting sphere (4) provides the following relationship:

$$W^{Elec}[\rho_i(\vec{x}),\, I_{R,C}\cdot\rho_j(\vec{x})] = W^{Elec}[I_{R,C}\cdot\rho_i(\vec{x}),\, \rho_j(\vec{x})]. \quad \textbf{[S11]}$$

It follows directly from Eq. **S11** that

$$\Phi[I_{R,C}\cdot(\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs}) = \Phi[(C\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs,inv}), \quad \textbf{[S12]}$$

where $\vec{x}_{obs,inv}$ denotes the inverse point, colinear with $\vec{x}_{obs}$ and the LDS center, at radius R^2/x_{obs}.

Image Shell Solutions. Using the identities described above and Eqs. **6a–6d** of the manuscript, the potential of a shell charge $(\sigma, b_q)|_{\vec{x}_q}$ at position $\vec{x}_{obs}$ inside or outside the LDS can be approximated by summing *in vacuo* the potentials of a set of point charges, shell charges, and shell charge sections as

where $\vec{x}_{inv}$ denotes the inverse point, colinear with $\vec{x}_q$ and the LDS center, at radius R^2/x_q. Following ref. 4, the electrostatic potential *in vacuo* of a partial charge shell $(\sigma, b_q, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}$ is numerically evaluated:

$$\Phi[(\sigma, b_p, \vec{x}_\theta,\, \theta_a \to \theta_b)|_{\vec{x}_q}](\vec{x}_{obs}) = \begin{cases} 2\pi\sigma \sum_{l=0}^{\infty} \frac{1}{2l+1}\left[P_{l+1}(\cos\theta_a) - P_{l-1}(\cos\theta_a) + P_{l-1}(\cos\theta_b) - P_{l+1}(\cos\theta_b)\right]\frac{r_{obs}^l}{b_k^{l-1}}P_l(\cos\theta_{obs}) \text{ for } r_{obs} < b_k \\ 2\pi\sigma \sum_{l=0}^{\infty} \frac{1}{2l+1}\left[P_{l+1}(\cos\theta_a) - P_{l-1}(\cos\theta_a) + P_{l-1}(\cos\theta_b) - P_{l+1}(\cos\theta_b)\right]\frac{b_k^{l+2}}{r_{obs}^{l+1}}P_l(\cos\theta_{obs}) \text{ for } r_{obs} \ge b_k, \end{cases} \quad \textbf{[S14]}$$

where r_{obs} and θ_{obs} are the polar coordinates of $\vec{x}_{obs}$ in a system with its origin at $\vec{x}_q$ and polar axis $\vec{x}_\theta$. $P_l(\,)$ denotes the Legendre

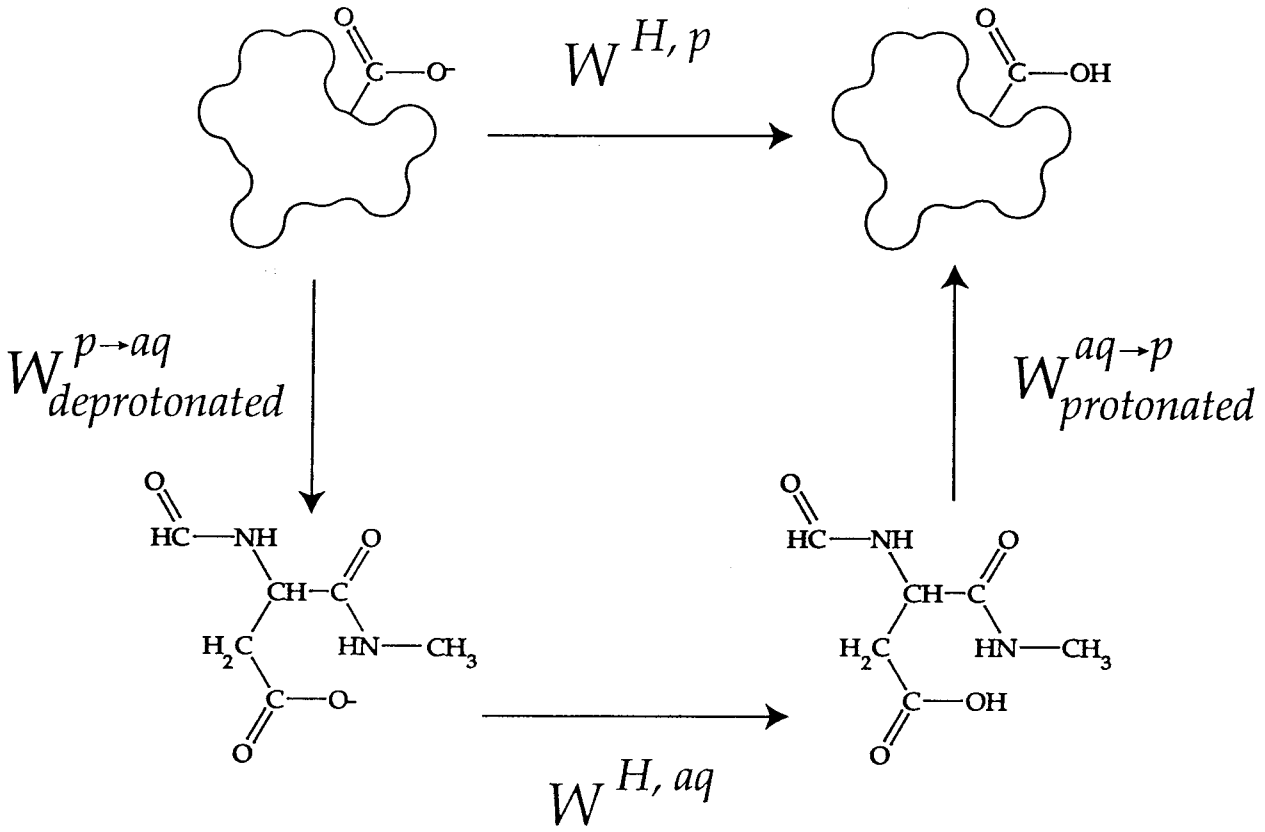

FIG. 4. Thermodynamic cycle for calculation of $W^{H,p}$. The energy of protonating a site in the protein is derived from a thermodynamic cycle connecting a titrating site in a protein (top) to an equivalent site in an isolated *N*-formyl, *N*-methyl amino amide (FMAA) of known pKa (bottom). The protonation potential consists of the transfer free energy of the deprotonated FMAA from the protein into aqueous solution, $W^{p\rightarrow aq}_{deprotonated}$, the free energy of protonation of the FMAA in aqueous solution, $W^{H,aq}$, and the transfer free energy of the protonated FMAA back into the protein, $W^{aq\rightarrow p}_{protonated}$. $W^{H,aq}$ is evaluated as 2.3·RT(pH-IpK_a), where IpK_a is the intrinsic pK_a of the appropriate FMAA. The deprotonated FMAA fragment in the protein (upper left) and both FMAA fragments in solution (bottom left and right) are placed in the most commonly occurring side-chain rotamer conformation of the amino acid (i* in the text). For the purpose of calculating transfer energies, the protein is represented by its polyglycine backbone stripped of all side-chain atoms. The dielectric boundary of the protein is defined by the molecular surface of the appropriate experimentally determined structure (including side chains).

polynomial of order *l*. The series in Eq. **S14** is truncated at the first term whose unsigned ratio to the sum of previous terms falls below 10^{-6}, with a maximum of 64 terms.

Interactions of Shell Charges. The interaction of two shell charges i and j can fall into one of three categories: (*i*) neither shell charge crosses the dielectric boundary; (*ii*) one shell charge crosses the LDS boundary and does not intersect the other shell; or (*iii*) both shell charges cross the LDS boundary, or one shell crosses the LDS boundary and intersects the other shell. If a single shell crosses the LDS boundary, it is taken to be j without loss of generality. For category *i*, Eqs. **S13a–S13d** and Eq. **1** of the manuscript fully describe the interaction. For category *ii*, the interaction of shell i with the image shells of j (Eqs. **S13e** and **S13f**) is evaluated by using Eq. **1** of the text, and the interaction of shell i with the image caps of shell j is determined by evaluating Eq. **S14** at the center of shell i. For category *iii*, the interaction of shell i with the image shells of j is evaluated by using Eq. **1** of the text, and the interaction of shell i with the image caps of shell j is evaluated by averaging Eq. **S14** over points on the surface of shell i. To perform this averaging, a set of points that uniformly covers shell i is generated by using the spherical tessellation of Eisenhaber (1). As noted in the text, all interactions between shells with surfaces separated by >2 Å are treated as though between point charges.

Energy Evaluation for Multicopy Sampling

Multicopy sampling algorithms require knowledge of two types of energy terms: the self energies of individual side-chain rotamers and the interaction energies of rotamer pairs. The self energy of rotamer i at residue position j, W_{ij}, is defined to be the summed self energies of its constituent atoms and their interaction with the fixed protein backbone. The interaction of rotamer i at position j and rotamer i′ at position j′, $I_{ij,i'j'}$, is evaluated as the summed interaction energies of the constituent atoms of i and i′.

Three terms constitute the energy function for all calculations: the modified Tanford–Kirkwood electrostatic potential, W^{Elec}, the CHARMM19 Lennard-Jones potential, W^{LJ}, and a protonation potential, $W^{H,p}$. The protonation potential is non-zero only at titratable sites and contributes only to the self energies of protonated rotamers. For deprotonated rotamers, self energies are evaluated: $W_{ij,deprotonated} = W^{Elec}_{ij,deprotonated} + W^{LJ}_{ij,deprotonated}$.

The self energies of protonated rotamers are defined relative to the self energy of the most commonly occurring deprotonated rotamer (denoted i*). Thus, the self energies of protonated rotamers are evaluated: $W_{ij,protonated} = W_{i^*j,deprotonated} + W^{H,p}_{ij}$. $W^{H,p}$ is evaluated by using the thermodynamic cycle illustrated in Fig. 4.

Interaction energies for protonated and deprotonated rotamers are calculated by applying the W^{Elec} and W^{LJ} potentials to the appropriate protonated and deprotonated atom sets.

1. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. (1995) *J. Comput. Chem.* **16,** 273–284.
2. Connolly, M. L. (1985) *J. Am. Chem. Soc.* **107,** 1118–1124.
3. Richmond, T. J. (1984) *J. Mol. Biol.* **178,** 63–89.
4. Jackson, J. D. (1975) *Classical Electrodynamics* (Wiley, New York).