Supporting Information for:

Accurate $\mathbf{pK_a}$ Calculations in Proteins with Reactive Molecular Dynamics Provide Physical Insight Into the Electrostatic Origins of Their Values

Joshua Zuchniarz, Yu Liu, Chenghan Li, and Gregory A. Voth*

Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA

*Corresponding Author: gavoth@uchicago.edu

$$U_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{S1}$$

Equation S1. The 12-6 Lennard-Jones potential, a function of interatomic distance r. Parameters σ and ϵ were fit in the present model for four pairs of atoms, described in the main text.

$$\mathbf{n}_{\mathrm{prot}} = \frac{\mathbf{r}_{C2} - \mathbf{r}_{C1}}{|\mathbf{r}_{C2} - \mathbf{r}_{C1}|} \tag{S2}$$

$$r_{\perp} = |\mathbf{v}_{\mathrm{CEC}} - (\mathbf{v}_{\mathrm{CEC}} \cdot \mathbf{n}_{\mathrm{prot}})\mathbf{n}_{\mathrm{prot}}| \tag{S3}$$

$$d_{\mathrm{SC}} = \mathbf{v}_{\mathrm{SC}} \cdot \mathbf{n}_{\mathrm{prot}} \tag{S4}$$

Equations S2-S4. Several relationships first defined in ref[9] and used for the enhanced sampling performed in this work. $\mathbf{r}_{C1}$ and $\mathbf{r}_{C2}$ are the centroids of the backbone atoms of residues 15-19 and residues 61-65 of SNase, respectively. $\mathbf{v}_{\mathrm{CEC}}$ is the vector pointing from the D/E66 sidechain carboxyl centroid to the excess proton CEC. $\mathbf{v}_{\mathrm{SC}}$ is the vector pointing from the D/E66 α-carbon to the sidechain carboxyl centroid.

Table S1. fitRMD Parameters Used to Compute Potentials of Mean Force.

	Asp	Glu		Asp	Glu
B	0.000928477	3.94487	V_{ii}	-139.912	-153.282
b	1.41581	1.41583	ϵ^{LJ}_{OE-HH}	0.231526	0.227986
b'	1.08883	1.09180	σ^{LJ}_{OE-HH}	1.36801	1.37334
A	2.72026	3.85746	ϵ^{LJ}_{Ow-HEP}	0.717000	0.730093
a	1.15572	1.15358	σ^{LJ}_{Ow-HEP}	1.22018	1.24711
g_1	-20.2207	-25.0434	ϵ^{LJ}_{OE-OH}	0.141951	0.112701
g_2	3.03394	2.99967	σ^{LJ}_{OE-OH}	3.00880	3.00179
g_3	1.43771	1.40739	ϵ^{LJ}_{OEP-Ow}	0.150728	0.195512
			σ^{LJ}_{OEP-Ow}	3.08218	3.11138
C	5.0917317 [a, b]		γ	1.783170 [a]	
c	8.9920023 [a, b]		P	0.1559053 [a]	
B	9.9178410 [a, b]		k	5.0664471 [a]	
b	1.1021518 [a, b]		D_{OO}	2.8621690 [a]	
b'	2.0066249 [a, b]			5.2394128 [a]	
q_O^{ex}	-0.0895456 [a]		R_{OO}^0	2.9425969 [a]	
q_H^{ex}	0.0252683 [a]		P'	7.6147672 [a]	
$q_{H^*}^{ex}$	0.0780180 [a]		α	7.4062624 [a]	
$q_O^{H_3O^+}$	-0.32 [a]		r_{OO}^0	1.8 [a]	
$q_H^{H_3O^+}$	0.44 [a]		V_{const}^{ij}	-21.064268 [a]	
d_{OO}^0	2.4 [a]		d_{OH}^0	1.0 [a]	
$\epsilon_{O^*O^*}$	0.098609686 [a]		$\sigma_{O^*O^*}$	3.118508 [a]	
$\epsilon_{H^*H^*}$	0.000040458 [a]		$\sigma_{H^*H^*}$	0.0 [a]	
$\epsilon_{O^*H_w}$	3.0 [a]		$\sigma_{O^*H_w}$	1.6 [a]	
D_{OH}	136.2026 [a]		α_{OH}	2.0834 [a]	
r_{OH}^0	0.98 [a]		k_α	77.4868 [a]	
			α_0	111.7269 [a]	

[a]Parameters of MS-EVB 3.2, published in ref[28] of the main text. Definitions of all these parameters can be found in ref[15]. [b]Parameters for water-hydronium interactions only. Similarly-named parameters in the main text and the top half of this table are the equivalent for amino acid-water interactions.