Supporting information
Tools for tracing proton transfer pathways with examples from the *Rb. sphaeroides* photosynthetic reaction centers

Rongmei Judy Wei[1,2], Umesh Khaniya[2,3], Junjun Mao[2], Jinchan Liu[4], Victor S. Batista[5], M.R. Gunner*[1,2,3]

[1]*Ph.D. Program in Chemistry, The Graduate Center, City University of New York, New York, NY 10016, USA*
[2]*Department of Physics, City College of New York, New York 10031, USA*
[3]*Ph.D. Program in Physics, The Graduate Center of the City University of New York, New York, NY 10016, USA*
[4]*Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA*
[5]*Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA*

**Correspondent author: M.R. Gunner email: mgunner@ccny.cuny.edu or marilyn.gunner@gmail.com*

SI. Table of content:

Part I: Procedures for generating the ubisemiquinone parameter for MD simulation.

SI.1.1: Molefacture: preparing initial structure: prepare head group:
Parameters were generated for the anionic semiubiquinone with a single isoprene unit for the tail as described for generation of the neutral quinone parameter (Ceccarelli et al. 2003). The semiquinone structure was drawn using VMD molefacture plugin (tutorial: http://www.ks.uiuc.edu/Research/vmd/plugins/molefacture/). And the structure was optimized with VMD and the coordinates were saved as pdb and psf files.

SI.1.2: Build parameter:
The psf and pdb files identify the missing parameters. The van der Waals parameters are assigned by analogy with the same atom type in the VMD force field toolkit, as described in: https://www.ks.uiuc.edu/Training/Tutorials/science/ffTK/fftk-tutorial.pdf (Mayne et al. 2015).
DFT calculations (https://github.com/judywei2333/semi_brc_parameters) were applied using Gaussian to optimize the structure geometry. The $Q^{\bullet-}$ has a charge of -1 and multiplicity of 2.

SI.1.3: Fitting partial charge with water hydrogen bonds
Water molecules are introduced to make hydrogen bonds to the $Q^{\bullet-}$ carbonyl oxygens. DFT is applied to calculate the atomic partial charges in the presence of these waters.

SI.1.3: Fitting bonds, angles, and dihedral-angles
Bond and angle optimization is further conducted in with DFT, with the parameters updated in the parameter file.
Lastly, the dihedral bond information is calculated as the torsional energy mapped using DFT. The Molecular mechanics energy is minimized using the new force field parameters. The periodicity (n) and phase shifts are taken from the neutral ubiquinone.

SI.1.4: Validating force filed parameters
After refitting the plot and refinement, the RMSD value of the QM and MM energy is 0.123 kcal/mol.

SI.1.5: Patching
The $Q^{\bullet-}$ head group was patched to the tail used for the neutral ubiquinone using the CHARM-GUI (Jo et al. 2008). The ubisemiquinone is called SQU, and the neutral ubiquinone is U10. The parameters can be found at https://github.com/judywei2333/semibrcparameters.

SI.1.6. Preparing the system for MD. The CHARMM36 force field is used for the protein, cofactors are patched to their protein amino acid ligands and ≈30,000 TIP3P water molecules and ≈300 POPC lipid molecules and 86 Na^+ and 84 Cl^- are added to the protein membrane system as described in previous work (Wei et al. 2022).

Part II

Table SI.1: The 15 most populated protonation microstates in the MCCE simulation of one snapshot from the MD trajectory SQ2.

There were 16.2 million μstates sampled in MCCE. There are 15 residues in the Q_B network.

Residues that are not shown here: AspM17, HisL190, LysH130, and GluH224 maintain their charge in all protonation microstates. HisH68 and LysH132 are always protonated and AspH124, GluH173 and GluM326 are always deprotonated in these 15 protonation microstates, although they change their charge in less populated protonation microstates.

H^+ μstate: Protonation microstate; Count μstates: how many times this protonation microstate is accepted in MC sampling; Probability of accepted microstates in this protonation microstates; Accumulated probability: the running (aggregated) sum of the probability of this protonation microstates added to that of all more probable states.

	Residue Charge								
H^+μstate	Asp L210	Asp L213	His H128	His H126	Asp M240	Glu L212	count μstates	probability	accumulated probability
1	-1	0	1	1	-1	-1	4189608	0.259	0.259
2	-1	0	1	1	-1	0	2146165	0.132	0.391
3	-1	0	0	1	-1	-1	2055004	0.127	0.518
4	-1	0	0	1	-1	0	1490145	0.092	0.610
5	-1	-1	1	1	-1	0	1049171	0.065	0.675
6	0	-1	0	1	-1	0	645169	0.040	0.715
7	0	-1	1	1	-1	0	581425	0.036	0.750
8	-1	-1	0	1	-1	0	497109	0.031	0.781
9	0	0	0	1	-1	-1	472555	0.029	0.810
10	0	0	1	1	-1	-1	361262	0.022	0.833
11	-1	0	1	1	0	-1	334960	0.021	0.853
12	-1	0	1	0	-1	-1	259261	0.016	0.869
13	-1	0	0	1	0	-1	197670	0.012	0.881
14	0	-1	1	0	-1	0	197023	0.012	0.894
15	0	0	1	0	-1	-1	171095	0.011	0.904

SI.2: MCCE calculated average protonation states for residues in the Q_B proton transfer network.

SI.2A: MCCE calculated average protonation states for residues with different protonation states in MD.

Neutral	NeutralQB				SemiQB•-								average	stdev	N
	Q1a	Q1b	Q1c	Q1d	SQ1a	SQ1b	SQ2a	SQ2b	SQ3a	SQ3b	SQ4a	SQ4b			
AspL210_	0	0	0	-0.03									-0.01	0.02	4
AspL213_					-0.75	-0.5	-0.16	-0.22			0	0	-0.27	0.30	6
GluL212_					-0.01	0.0	-0.28	-0.53	-0.06	0	-0.17	-0.24	-0.16	0.18	8
GluM236_	0	0	-0.92	-0.84	-0.39	-0.6			0	-0.85	-0.06	-0.75	-0.44	0.40	10
HisH68_	0.64	0.67	0.02	0.41							0.42	0.45	0.44	0.23	6
LysH197_					0.95	0.91	0.97	0.99	0.96	0	1	0.41	0.77	0.37	8
LysH146_					0.98	0.99	1	1	1	1	1	1	1.00	0.01	8

SI.2B: MCCE calculated average charge of residues in snapshots from trajectories where the residue is ionized. N is the number of snapshots averaged.

Ionized	NeutralQB				SemiQB•-								average	stdev	N
	Q1a	Q1b	Q1c	Q1d	SQ1a	SQ1b	SQ2a	SQ2b	SQ3a	SQ3b	SQ4a	SQ4b			
AspL210_					-0.69	-0.95	-0.99	-0.82	-0.99	-0.99	-0.99	-1	-0.93	0.11	8
AspL213_	-0.99	-0.95	-1	-0.99					-0.95	-0.98			-0.98	0.02	6
GluL212_	-0.98	-0.88	-0.11	-0.11									-0.52	0.48	4
GluM236_							-1	-0.99					-1.00	0.01	2
HisH68_					0.51	0.59	0.6	0.99	0.85	0.91			0.74	0.20	6
LysH197_	1	1	1	1									1	0	4
LysH146_	1	1	1	1									1	0	4

SI.2C: The MCCE calculated charge of residues in the Q_B network whose charge is the same in all MD trajectories.

	SQ1a	SQ1b	SQ2a	SQ2b	SQ3a	SQ3b	SQ4a	SQ4b	average	stdev
Asp_M17	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	0.00
Asp_M23	-1	-1	-1	-1	-1	-1	-1	-1	-1.00	0.00
Asp_M240	-1	-1	-1	-0.94	-1	-1	-0.99	-0.99	-0.99	0.02
Glu_H173	-1	-1	-0.89	-0.99	-0.98	-1	-1	-1	-0.98	0.04
Glu_H224	-1	-1	-0.99	-1	-1	-1	-1	-1	-1.00	0.00
His_M301	0.83	0.9	0.84	0.91	0.87	0.89	0.72	0.92	0.86	0.07
His_H126	0.7	0.77	0.85	0.93	0.84	0.79	0.8	0.88	0.82	0.07
His_H128	0.33	0.81	0.78	0.63	0.64	0.43	0.29	0.22	0.52	0.23
Lys_H130	1	1	1	1	1	1	1	1	1.00	0.00
Lys_H132	1	1	1	0.99	1	1	1	1	1.00	0.00

Table SI.3: Number of connections made to each residue in the Q_B or Q_A network found the MD trajectories. Connections can be made with a maximum of four intervening waters. ^: Surface residue. *: Charged residue. This can be compared to Table 3 in the main text where connections are made with a maximum of two intervening waters.

Table SI.3A. Connections to residues that have different protonation states in different trajectories. Av crg: Average number of other residues to which the residue is connected to in trajectory where the residue is charged; Av neu: This count in trajectory where residue is neutral.

	Q1	SQ1	SQ2	SQ3	SQ4	av crg	av neu
Q_B	4	5	6	7	6	6.00	4.00
Asp_L213	10*	10	11	13*	8	11.50	9.67
Glu_M236	8	4	10*	7	6	10.00	6.25
^His_H68	7	9*	12*	8*	4	9.67	5.50
Asp_L210	9	8*	10*	11*	8*	9.25	9.00
Glu_L212	8*	4	5	6	2	8.00	4.25

Table SI.3B: Connectivity in Q_B network for residues with the same charge in all MD trajectories.

	Q1	SQ1	SQ2	SQ3	SQ4	average		stdev
^His_H126*	20	20	25	24	13	20.4	±	4.7
Ser L223	21	14	19	8	16	15.6	±	5
Lys H130*	12	11	13	13	11	12.0	±	1
Thr L226	13	10	11	11	10	11.0	±	1.2
Thr L214	10	12	11	11	9	10.6	±	1.1
Asp M17*	9	11	12	10	9	10.2	±	1.3
Glu H173*	11	9	12	11	6	9.8	±	2.4
Asp M240*	6	4	10	12	8	8.0	±	3.2
Asp H124*	9	7	8	7	7	7.6	±	0.9
^Tyr_M3	1	9	9	11	5	7.0	±	4
Thr L208	9	2	10	11	3	7.0	±	4.2
^His H128*	1	8	14	8	3	6.8	±	5.1
Ser_M8	4	6	7	6	8	6.2	±	1.5
^Lys_H132*	3	6	7	7	5	5.6	±	1.7
^Glu H224*	6	5	5	5	5	5.2	±	0.4
Glu M234*	7	0	0	2	0	4.5	±	3.5
^Thr M21	4	3	4	6	5	4.4	±	1.1
Ser M227	3	4	5	4	2	3.6	±	1.1
Ser H143	1	2	4	2	2	2.2	±	1.1
Tyr L222	0	1	1	5	2	1.8	±	1.9
Asp H170	2	0	2	1	0	1.7	±	0.6
Glu H230*	3	0	1	1	0	1.7	±	1.2
His L116	0	2	0	2	0	1.3	±	1.2
Thr M 261	0	0	1	2	2	1.3	±	1
Glu_H122*	1	1	2	1	0	1.3	±	0.5
His L211	1	0	1	2	1	1.3	±	0.5
Glu M232*	1	1	2	1	1	1.2	±	0.4
His L190	1	1	1	1	2	1.2	±	0.4
Ser L4	1	1	1	2	1	1.2	±	0.4

Table SI.3C: Number of connections to residues in the network connected to Q_A.

	Q1	SQ1	SQ2	SQ3	SQ4	ave		std
Q_A	2	2	2	4	4	2.8	±	1.1
Thr_M222	1	1	1	1	1	1.0	±	0.0
His_M219	1	1	1	1	1	1.0	±	0.0
Thr_M261	0	0	1	2	2	1.3	±	1.0
Tyr_H40	0	0	1	0	1	0.7	±	0.6
Ser_L4	1	1	1	2	1	1.2	±	0.4

Table SI.4A: **Number of the 2500 frames (snapshots) from 51-100 ns making the residue-residue connection via a maximum of two intervening waters in the four SQ MD trajectories.** Residues whose protonation states change are shown with highlight. Orange: ionized acid as proton acceptors; blue: protonated bases as proton donor; pink: both proton donor and acceptor are ionized; no color residue is neutral in the MD trajectory.

donor	acceptor	SQ1	SQ2	SQ3	SQ4
His_H126	Asp_H124	1563	62	1539	2500
His_L190	Q_B	2500	2500	2500	2500
Lys_H130	Glu_H173	2500	2500	2500	2500
Lys_H130	Asp_L210	2435	1432	2500	2490
Thr_L226	Glu_H173	436	35	267	2073
Thr_L226	Asp_L213	398	31	270	2069
Ser_H221	Glu_H224	1800	2194	1755	1809
Thr_H223	Glu_H224	507	1579	1462	1484
Ser_L223	Asp_L213	144	816	0	1046
Ser_L223	Q_B	2384	2257	2500	947
Ser_L223	Asp_M17	16	113	0	546
Thr_L214	Asp_M17	1075	714	125	448
Thr_L214	Asp_L210	1074	706	125	444
Lys_H132	Asp_H124	172	469	456	403
Ser_M8	Tyr_M3	141	309	23	358
His_H68	Glu_M236	1610	1514	804	342
Lys_H130	Glu_M236	2097	1788	331	292
His_H128	His_H126	31	16	117	178
Lys_H132	Glu_H224	66	538	150	175
Ser_L223	Asp_L210	2	6	0	117
Tyr_M3	Glu_M232	0	16	1	108
Thr_L214	Asp_L213	346	369	115	66
Tyr_M3	Ser_M8	44	121	518	63
His_H126	His_H128	18	24	237	38
His_H126	Asp_L210	1157	195	1408	28
Lys_H130	Asp_L213	746	1035	2493	27
His_H126	Asp_M17	1135	220	2034	22
Lys_H220	Glu_H224	31	7	41	14
Lys_H130	Glu_L212	14	19	1271	11
Ser_L223	Glu_H173	1	129	0	8
Thr_L226	Ser_L223	4	3	62	8
Thr_L226	Asp_L210	42	0	42	5
Thr_M21	His_H126	0	12	4	5

Continued

donor	acceptor	SQ1	SQ2	SQ3	SQ4
Ser_M8	GLU_H173	2	3	2	1
Thr_L226	Q_B	2	6	16	1
Thr_M21	Asp_M17	0	1	7	1
Tyr_L222	Asp_L213	0	0	1	1
His_H68	Asp_L210	442	1624	413	
Lys_H132	His_H128	35	495	341	0
Lys_H130	Asp_M17	0	201	159	0
His_H126	Glu_M22	82	101	0	0
Ser_L4	Asp_M240	0	86	0	0
His_H68	Glu_H173	0	32	0	0
Thr_L214	Glu_H173	0	28	0	0
His_H126	Thr_M21	0	23	18	0
His_H126	Asp_L213	33	14	120	0
His_H68	Asp_M240	280	14	137	0
Thr_L208	Asp_L210	0	14	1664	0
Tyr_M3	Ser_M227	0	13	0	0
His_H68	Asp_M17	0	8	0	0
Ser_L223	Thr_L226	0	7	0	0
His_H128	Asp_L210	5	5	6	0
His_H128	Asp_M17	2	5	25	0
Lys_H62	Asp_M240	0	5	0	0
His_H68	His_H126	0	4	0	0
Thr_L214	His_H126	82	3	14	0
His_H68	Glu_L212	0	2	0	0
Lys_H130	Asp_H124	0	2	0	0
Ser_M227	Glu_H173	7	2	0	0
Thr_L208	Asp_M240	0	2	4	0
Tbr_L214	Asp_H124	6	2	934	0
His_H126	Asp_M23	0	1	0	0
His_H126	Thr_L214	0	1	0	0
His_H68	Asp_L213	0	1	0	0
Thr_L208	Asp_H124	0	1	0	0

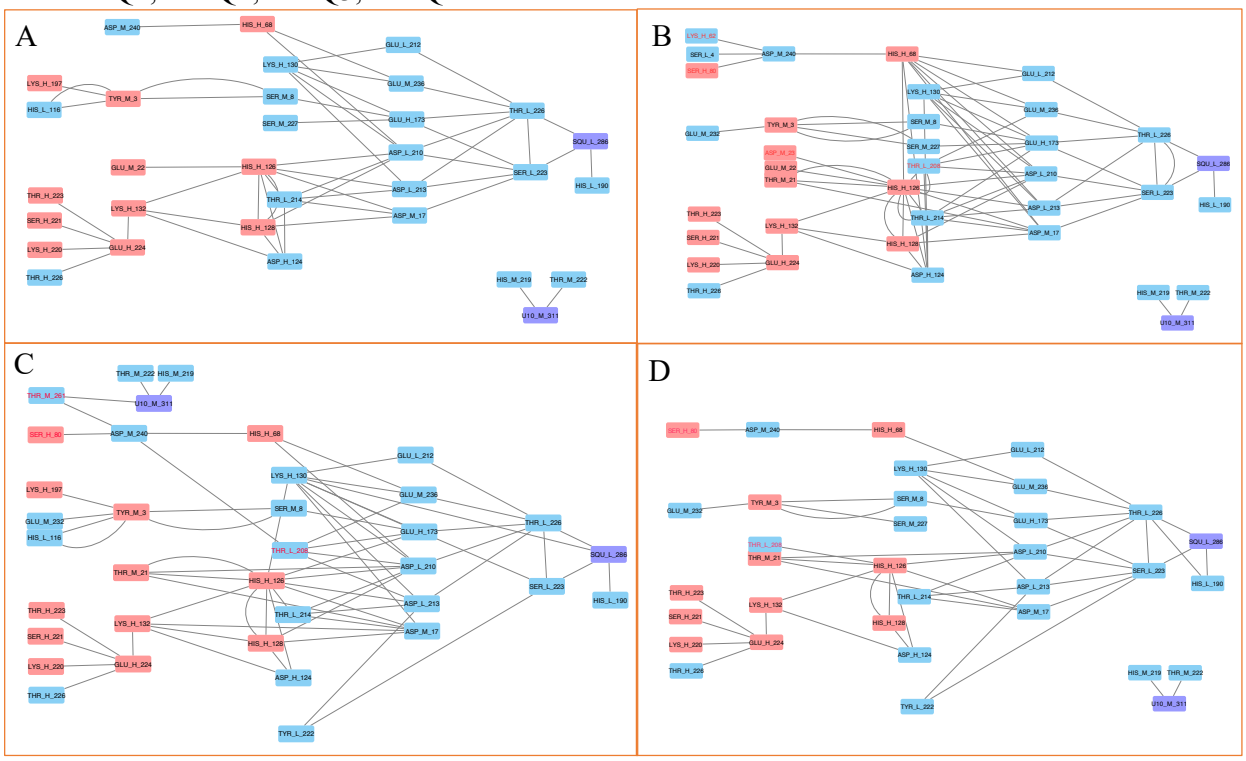
Thr_M21	Asp_L210	0	0	3	5
His_H128	Asp_H124	1231	2500	1287	4
Ser_H80	Asp_M240	0	7	2	4
Ser_M227	Tyr_M3	0	26	0	4
Thr_H226	Glu_H224	19	3	5	4
Thr_L226	Glu_L212	16	7	142	4
Lys_H132	His_H126	1	2	17	3
Thr_L226	His_L190	0	0	0	3
Thr_L226	Glu_M236	30	2	1	2
Tyr_L222	Ser_L223	0	0	1	2
His_H126	Thr_L208	0	12	0	1
His_H68	Asp_H124	0	6	0	1

Thr_L208	Asp_L213	0	1	29	0
Thr_L208	Glu_H173	0	1	78	0
His_H126	Glu_H173	0	0	2	0
His_L116	Tyr_M3	405	0	142	0
Lys_H130	His_H126	0	0	15	0
Lys_H130	Q_B	0	0	3	0
Lys_H130	Ser_L223	0	0	12	0
Lys_H132	Asp_M17	0	0	15	0
Lys_H197	Tyr_M3	25	0	17	0
Thr_M133	Thr_M146	0	0	1	0
Thr_M261	Asp_M240	0	0	8	0
Tyr_M3	His_L116	3	0	5	0

Table SI.4B: Counting the time each residue type serves as a donor or acceptor in Table SI.4A.

	SQ2	SQ2	SQ4	SQ4
	donor	acceptor	donor	acceptor
Ser	13	3	10	3
Thr	22	4	17	1
Tyr	3	2	4	2
Asp	0	32	0	20
Glu	0	22	0	15
Lys	13	8	9	0
His	25	0	11	5

SI.5A: Networks of hydrogen bonded connection with a maximum two intervening water molecules. Quinones are in purple, and surface residues are in orange-pink and buried residues on the path are in blue. A: SQ1; B: SQ2; C: SQ3; D: SQ4.



SI.5B: Examples of proton transfer paths in each trajectory:

SQ1 (&SQ2, &SQ4) entry 1: $HisH126^{+}$→$AspM17^{-}$→1w→ $SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ1 (&SQ2) entry 2: $HisH128^{+}$→ 1w →$AspM17^{-}$→1w→ $SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ1 (&SQ2, &SQ3) entry 3: $HisH68^{+}$ →1W →$AspL210^{-}$→2W→$SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
or: $HisH68^{-}$→1W → $GluM236^{0}$→1W→ $ThrL226^{0}$→2w→ $\mathbf{Q_B}^{\bullet-}$
SQ1 entry4: $TyrM3^{0}$→1W→$SerM8^{0}$→2w→$GluH173^{-}$→1w→ $THRL226^{0}$→ 2W→ $\mathbf{Q_B}^{\bullet-}$
Or: $TyrM3^{0}$→1W→$SerM8^{0}$→2w→$GluH173^{-}$→2w→$SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$

SQ2: See in main text.

SQ3 entry1: $HisH126^{+}$→2w→$AspL210^{-}$ →2w→$ThrL226^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ3 entry2: $HisH128^{+}$→ 2w → $AspL210^{-}$ →2w→$ThrL226$→ $\mathbf{Q_B}^{\bullet-}$
SQ3 entry 3: $HisH68^{+}$→$GluM236^{0}$ →1W →$ThrL226^{0}$→2w→ $\mathbf{Q_B}^{\bullet-}$
or: SQ3 entry 3(&SQ1, &SQ2): $HisH68^{+}$→1W →$AspL210^{-}$ →2W→$SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ3 entry04: $TyrM3^{0}$→1W→$SerM8^{0}$ →2w→$GluH173^{-}$ →2w→$SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$

SQ4(&SQ1, &SQ2) entry 1: $HisH126^{+}$→$AspM17^{-}$→1w→ $SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ4 entry 2: $LysH132^{0}$→ $AspH124^{-}$→3w→ $SerL223^{0}$→ $\mathbf{Q_B}^{\bullet-}$
SQ4 entry 3: $HisH68^{0}$→2w→ $GluM236^{0}$→2W →$ThrL226^{0}$→2w→ $\mathbf{Q_B}^{\bullet-}$
SQ4 entry 3: $HisH68^{0}$→3w→ $GluH173^{0}$→1W →$ThrL226^{0}$→2w→ $\mathbf{Q_B}^{\bullet-}$

Reference:

Ceccarelli M, Procacci P, Marchi M (2003) An ab initio force field for the cofactors of bacterial photosynthesis. Journal of Computational Chemistry 24 (2):129-142. doi:https://doi.org/10.1002/jcc.10198

Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29 (11):1859-1865. doi:10.1002/jcc.20945

Mayne CG, Muller M, Tajkhorshid E (2015) Parameterizing small molecules using the force field toolkit (fftk). University of Illinois at Urbana-Champaign

Wei RJ, Zhang Y, Mao J, Kaur D, Khaniya U, Gunner MR (2022) Comparison of proton transfer paths to the QA and QB sites of the Rb. sphaeroides photosynthetic reaction centers. Photosynth Res. doi:10.1007/s11120-022-00906-x