Supporting Information

Why the Electron Chooses One of Two Symmetry-Related Paths in the Type II Bacterial Photosynthetic Reaction Centers

Rongmei Judy Wei[1,2], Christine Kirmaier[3], Dewey Holten[3], Jan Kern[4], Philip D. Laible[5], Deborah K. Hanson[5], Gehan Ranepura[2,6], Md Raihan Uddin[2,7], Jose Ortiz-Soto[1,2] M.R. Gunner[1,2,6,7]*

[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]Department of Chemistry, Washington University, St. Louis, MO 63130, USA
[4]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[5]Biosciences Division, Argonne National Laboratory, Lemont, IL 60439, USA
[6]Ph.D. Program in Physics, The Graduate Center of the City University of New York, New York, NY 10016, USA
[7]Department of Biochemistry, The Graduate Center, City University of New York, New York, NY 10016, USA

* **Correspondence:** Department of Physics, City College of New York, 160 Convent Avenue, New York, NY 10031, USA.
E-mail address: mgunner@ccny.cuny.edu (M.R. Gunner)

Table of Contents

Part I

1a) MCCE calculated microstate energy.

The free energy of each MCCE microstate is:

$$\Delta H^x = \sum_{i=1}^{M} \delta_{x,i}\{[2.3 m_i k_b T(\mathrm{pH} - \mathrm{p}K_{\mathrm{sol},i}) + n_i F(E_h - E_{\mathrm{msol},i})]$$

$$+ (\Delta\Delta G_{\mathrm{rxn},i} + \Delta\Delta G^{\mathrm{CE}}_{\mathrm{bkbn},i} + \Delta\Delta G^{\mathrm{LJ}}_{\mathrm{bkbn},i} + \Delta\Delta G_{\mathrm{torsion},i} + \Delta\Delta G_{\mathrm{SAS},i}) \qquad \text{Eqn 1}$$

$$+ \sum_{j=i+1}^{M} \delta_{x,j}[\Delta G^{\mathrm{CE}}_{ij} + \Delta G^{\mathrm{LJ}}_{ij}]\}$$

In a microstate each residue has a defined protonation state and conformation and each cofactor a defined redox state. The first line of equation 1 gives the energy each group would have in solution at the imposed pH and E_h, in its assigned microstate protonation or redox state. The neutral form is the reference. *M* denotes the total number of conformers, while $\delta_{x,i}$ equals 1 if conformer *i* is present in microstate *x* and 0 otherwise. The parameter m_i is assigned a value of 1 for protonated basic residues, -1 for deprotonated acidic residues, and 0 for neutral conformers. The reference pK_a of a residue type in solution is denoted as $\mathrm{p}K_{\mathrm{sol,i}}$, and E_m is the reference electrochemical midpoint potential of the redox-active group. *F* stands for the Faraday constant, and n_i indicates the number of electrons exchanged during a redox titration involving the conformer. The solution parameters *pH* and E_h define the chemical potential of protons and electrons, respectively, which establish equilibrium with the protein.

The second line of equation 1 describes the self-energies of the conformers that make up the microstates, which remain independent of conformer selection for other residues. These self-energies include the loss of solvation energy as the conformer transitions from the reference solvent to its position within the protein, continuum electrostatics (CE), Lennard-Jones (LJ) van der Waals interactions with the fixed backbone amides and torsion energy, as well as favorable van der Waals interactions between exposed side-chain surfaces and the implicit solvent ($\Delta\Delta G_{SAS}$).

The third line of the equation accounts for the pairwise continuum electrostatic and Lennard-Jones interactions between conformers of different residues. The MCCE force field, previously benchmarked, is used for these calculations. All energy calculations are performed before Monte Carlo (MC) sampling.

MCCE microstates are subjected to Metropolis Hastings sampling at pH 8. The E_m calculations find the probability of reduced and ground state of each individual BChl and BPh as a function of E_h. The E_m is the E_h where the probability of the two redox forms are equal in the simulated titration.

1b) Mean Field Energy (MFE) derived following MC sampling.

The MCCE MC analysis calculates the probability of each conformer choice for each residue and ligand in the full ensemble of microstate. A mean field analysis of the results allows an estimate of the interaction of individual residues with the cofactor that shift its midpoint potential from

$E_{m,sol}$. Here the mean field interaction of a conformer with the cofactor = (the average occupancy of the conformer) times (its interaction with the cofactor of interest). The mean field interaction of a residue with the cofactor sums the conformer interactions.

Thus, the average residue interactions in the tables are calculated from:

$$\Delta G_{ij}^{MFE} = \sum_{j=1} P_j \left(\Delta G_{ij}^{CE} + \Delta G_{ij}^{LJ} \right) \qquad \text{Eqn 2}$$

Where the i is the cofactor of interest and j runs over the conformers in the residue of interest. P_j is the probability of conformer j in the MC sampling.

ΔG_{ij}^{mfe} is a good measure of the true interaction if the change in the cofactor redox state does not change the conformer choice of the residue.[1] This method breaks down when a cofactor and residue conformer choices are coupled together. A test is whether the sum of all ΔG_{ij}^{mfe} equals the ΔG_{ij}^{MC}. The two differ by only a few meV for any cofactor here (S.I. Table 3).

Part II

S.I. Table 1. The averaged charge of residues that do not adopt their canonical protonation state at pH 8 in the MCCE calculations.

pH	8	8
dielectric constant(ε)	4	2
AspL210	-0.65	0
GluM236	-0.23	0
HisH68	0.29	0
LysH146	0	0
LysH197	0	0
GluH43	-0.4	0
GluL104	0	0
LysH197	0	0
His H126	0.73	0.21
His H128	0.44	0.27
LysH232	1	0
LysM144	1	0.92
CysH234	0	-1

In solution at pH 8 Asp and Glu will be deprotonated with a negative charge, Arg and Lys protonated with a positive charge, Tyr, Cys and His will be neutral. The identity and ionization of L and M branch Asp, Glu, Arg and Lys that are less than 95% ionized or Cys, Tyr or His which are more than 5% ionized are reported. A non-integer probability does not mean the residue has a fractional charge. Rather, this percentage of the accepted microstates have this protonation state. The calculations here are carried out at ε=2. The protonation states are provided at ε=4 for comparison with earlier studies.[2, 3]

S.I. Table 2. The contributions of individual free energy components to the E_ms of the four cofactors in the three structures.

S.I. Table 2a: Calculated E_m and contributions to the shift of the free energy of ionization in the protein of the A and B branch BChl.

$BChl_B$	1AIG	2GNU	2J8C	av	std		$BChl_A$	1AIG	2GNU	2J8C	av	std
E_m (mV)	-920	-951	-1050	-984	68		E_m (mV)	-811	-803	-847	-820	23
ΔG° (meV)				av	std		**ΔG° (meV)**				av	std
Backbone	-167	-143	-207	-172	32		Backbone	-140	-115	-145	-133	16
Desolvation	158	191	321	223	86		Desolvation	156	239	116	170	63
Cof+Lig	225	198	303	242	55		Cof+Lig	224	130	302	219	86
Side chains + Fe^{3+}	-228	-189	-223	-213	21		Side chains + Fe^{3+}	-363	-380	-375	-373	9
Background side chains	74	38	3	38	36		Background side chains	66	72	90	76	12
All side chains	-154	-151	-220	-175	39		All side chains	-297	-308	-285	-297	12
Breakdown Cof_Lig Interactions (meV)				av	std		**Breakdown Cof_Lig Interactions (meV)**				av	std
HisM182	134	114	214	154	53		HisL153	124	139	207	157	44
P_B	7	-7	-29	-10	18		P_A	0	-43	-44	-29	25
P_A	82	70	48	67	17		P_B	51	22	57	43	19
$BChl_A$	25	33	33	30	5		$BChl_B$	36	44	43	41	4
$BChl_B$	NA	NA	NA	NA	NA		$BChl_A$	NA	NA	NA	NA	NA
BPh_A	-5	4	-4	-2	5		BPh_B	2	3	-5	0	4
BPh_B	-17	-13	50	7	38		BPh_A	11	-32	48	9	40
Q_A	-1	-1	-1	-1	0		Q_B	1	-1	-1	0	1
Q_B	0	-2	-8	-3	4		Q_A	-1	-2	-3	-2	1
Sum_Cof+Lig	225	198	303	242	55		Sum_Cof+Lig	224	130	302	219	86
Side chains + Fe^{3+}(meV)				av	std		**Side chains + Fe^{3+}(meV)**				av	std
Fe^{3+}	-130	-126	-158	-138	17		Fe^{3+}	-151	-151	-152	-151	1
GluM234	35	31	40	35	5		GluM234	39	33	38	37	3
ArgM164	-59	-61	-63	-61	2		ArgL135	-60	-67	-68	-65	4
ArgL217	-24	-17	-26	-22	5		ArgM253	-31	-33	-33	-32	1
ArgL231	-35	-26	-29	-30	5		ArgM267	-29	-21	-28	-26	4
ArgM132	-67	-49	-59	-58	9		ArgL103	-89	-86	-75	-83	7
PheL181	0	0	0	0	0		TyrM210	-75	-80	-75	-77	3
LeuM135	0	0	0	0	0		GluL106	41	33	27	34	7
ThrM186	-54	-35	-46	-45	10		ValL157	0	0	0	0	0
GluL212	32	26	37	32	6		ArgM241	-8	-8	-9	-8	1
AspM88	32	32	40	35	5		none	na	na	na	na	na
AspM184	42	36	41	40	3		none	na	na	na	na	na
Sum_Side chain/Fe^{3+}	-228	-189	-223	-213	21		Sum_Side chain/Fe^{3+}	-363	-380	-375	-373	9

Free energy (meV): Negative values favor reduction. Desolvation: loss of solvation energy of the charge moving from the reference solvent, water, to the position within the protein; Backbone: interaction with the amide backbone dipoles; Residue interaction: Side chain + Fe^{3+}: Residues that shift the ΔG of reduction by more than ±29 meV;; Background side chains: Summed interactions with residues that interact with the cofactor by less than ±29 meV; Cofactors+Ligand: interaction with the neutral neighboring cofactors and the axial ligands to $BChl_B$ and $BChl_A$ (HisM182 and HisL153). Interactions with side chains determined with mean field energy approximation (SI Eqn 2).

Free energies are the difference in interaction between ionized and neutral conformation, i.e. $(cofactor)^-$ - $(cofactor)^0$. F ; negative ΔG_{A-B} favor A-branch reduction. Fe^{3+} and GluM234 are along the midpoint of the protein and their interaction with A-and B-branch cofactors are given on a single row. All other rows provide interactions with residues aligned by BLAST considering the interactions of A-branch residues with nearby A-branch cofactors or B-branch residue with B-branch cofactors. In each case at least one of the residues on each row has an interaction greater than ±29 meV with its cofactor. Individual residue interactions calculated with the mean field energy approximation (SI Eqn 2).

S.I.Table 2b: Calculated E_m and contributions to the shift of the free energy of ionization in the protein of the A and B branch BPh.

BPh_B	1AIG	2GNU	2J8C	av	std
E_m (mV)	-336	-361	-367	-355	16
ΔG° (meV)				ave	std
Backbone	-183	-158	-187	-176	16
Desolvation	475	466	519	487	28
Cof+Lig	-188	-222	-187	-199	20
Side chains + Fe^{3+}	-356	-301	-329	-329	28
Background side chains	11	-5	-29	-8	20
Sum side chains	-345	-306	-358	-336	27
Cof_Lig Interactions (meV)				av	std
P_B	-146	-124	-139	-136	11
P_A	-57	-54	-65	-59	6
$BChl_A$	7	16	20	14	7
$BChl_B$	16	-60	20	-8	45
BPh_A	-6	5	-5	-2	6
BPh_B	NA	NA	NA	NA	NA
Q_A	-1	-1	-1	-1	0
Q_B	-1	-4	-17	-7	9
Sum_Cof+Lig	-188	-222	-187	-199	20
Side chain+ Fe^{3+}(meV)				av	std
Fe^{3+}	-293	-272	-336	-300	33
GluL212	84	71	92	82	11
GluM234	78	74	88	80	7
GluM232	34	30	36	33	3
ArgM132	-175	-133	-155	-154	21
ArgM164	-45	-48	-48	-47	2
ArgL217	-52	-41	-54	-49	7
ArgL231	-52	-40	-48	-47	6
ArgM267	-48	-34	-62	-48	14
HisM182	26	18	30	25	6
ThrM133	-6	-11	7	-3	9
LeuM135	0	0	0	0	0
AlaM139	0	0	0	0	0
LysM144	-24	-26	-31	-27	4
AspL218	55	49	59	54	5
AspL213	51	46	55	51	5
GluH173	39	37	41	39	2
none	na	na	na	na	na
none	na	na	na	na	na
none	na	na	na	na	na
GluM263	25	24	40	30	9
ArgM136	-53	-45	-43	-47	5
Sum	-356	-301	-329	-329	28

BPh_A	1AIG	2GNU	2J8C	av	std
E_m (mV)	-250	-293	-321	-288	36
ΔG° (meV)				ave	std
Backbone	-192	-145	-141	-159	28
Desolvation	459	520	505	495	32
Cof+Lig	-184	-179	-130	-164	30
Side chains + Fe^{3+}	-530	-587	-566	-561	29
Background side chains	126	104	73	101	27
Sum side chains	-404	-483	-493	-460	49
Cof_Lig Interactions (meV)				av	std
P_A	-151	-151	-147	-150	2
P_B	-71	-55	-45	-57	13
$BChl_B$	16	25	28	23	6
$BChl_A$	22	1	46	23	23
BPh_B	0	4	-8	-1	6
BPh_A	NA	NA	NA	NA	NA
Q_B	1	-1	-1	0	1
Q_A	-1	-2	-3	-2	1
Sum_Cof+Lig	-184	-179	-130	-164	30
Side chain+ Fe^{3+}(meV)				av	std
Fe^{3+}	-310	-324	-323	-319	8
GluL212	40	39	42	40	2
GluM234	81	74	83	79	5
GluM232	37	40	40	39	2
ArgL103	-165	-219	-183	-189	27
ArgL135	-52	-55	-58	-55	3
ArgM253	-58	-68	-74	-67	8
ArgM267	-36	-30	-48	-38	9
ArgL231	-49	-48	-41	-46	4
HisL153	19	30	20	23	6
GluL104	-65	-62	-64	-64	2
GluL106	65	81	66	71	9
LysL110	-37	-40	-37	-38	2
TyrL115	-1	-4	5	0	5
TrpM255	14	18	16	16	2
none	na	na	na	na	na
none	na	na	na	na	na
ArgM247	-35	-38	-36	-36	2
GluL6	46	54	51	50	4
ArgM228	-27	-29	-32	-29	3
GluM263	28	31	40	33	6
ArgL10	-25	-37	-33	-32	6
Sum	-530	-587	-566	-561	29

S.I. Table 3: Validation of use of mean field energy to decompose contributions to the E_ms derived by MC sampling.

	Backbone	Desolvation	Residues + cofactors	$E_{m,}$ref	MFE E_m, mV	MC E_m, mV	MC E_m - MFE E_m
$BChl_B$							
1AIG	-167	158	71	-860	-922	-920	2
2GNU	-143	191	47	-860	-955	-951	4
2J8C	-207	321	83	-860	-1057	-1050	7
$BChl_A$							
1AIG	-140	156	-73	-860	-803	-811	-8
2GNU	-115	239	-178	-860	-806	-803	3
2J8C	-145	116	17	-860	-848	-847	1
BPh_B							
1AIG	-183	475	-533	-580	-339	-336	3
2GNU	-158	466	-528	-580	-360	-361	-1
2J8C	-187	519	-545	-580	-367	-367	0
BPh_A							
1AIG	-192	459	-588	-580	-259	-250	9
2GNU	-145	520	-662	-580	-293	-293	0
2J8C	-141	505	-623	-580	-321	-321	0

The MC E_ms are obtained by MC sampling using the microstate energies from SI Eqn 1.

$$\Delta\Delta G_{protein}^{MC} = -nF(E_{m,MCCE} - E_{m,ref}) \qquad \text{Eqn 3a}$$

The MFE E_ms sum the individual energy terms using the conformer probabilities obtained with MC sampling using SI Eqn 2. Residues and Cofactors here are the (cofactors + ligands) + (side chains + Fe^3) + (background side chains).

$$\Delta\Delta G_{protein}^{MFE} = \Delta G_{desolvation} + \Delta G_{backbone} + \Delta G_{residues+cofactors} \qquad \text{Eqn 3b}$$

S.I. Table 4. Conservation of residues with significant interaction with cofactors.

Conservation A-branch residues		Residues influencing BChl A and B					Conservation B-branch residues	
aa	%cons	$BChl_B$	meV	$BChl_A$	meV	diff	%cons	aa
		Fe^{3+}	-130	Fe^{3+}	-151	-21		
E:100	1	GluM234	35	GluM234	39	4	1	E:100
R:99;(Q)	1	ArgM164	-58	ArgL135	-60	-2	1	R:100
R:99;(Q)	0.99	ArgL217	-24	ArgM253	-31	-7	0.99	R:99;(-)
R:100	1	ArgL231	-35	ArgM267	-29	6		
R:99;(H)	1	ArgM132	-69	ArgL103	-89	-20	1	R:100
F:98; (W)	0.98	PheL181	0	TyrM210	-75	-75	0.99	Y:99;(F)
T:99;(V)	1	ThrM186	-54	ValL157	0	54	1	V:100
E:100	1	GluL212	32	ArgM241	-8	-40	1	R:100
D:59;Q:23:E:14; (K:N,T,K,M)	0.59	AspM88	32	none	na	-32		
L:36;V:15;R:14; (T,N,M,I,,F,Q,D, E,K,Y,D,H,W))	0.36	LeuM135	0	GluL106	41	41	1	E:100
		Sum	-271	Sum	-363	-92		

Interactions energies with cofactors from Table 3 in main text. Free energy (meV): Negative values favor reduction. Free energies are the difference in interaction between ionized and neutral states; i.e. $(cofactor)^-$ - $(cofactor)^0$. Fe^{3+} and GluM234 are along the axis of twofold symmetry of the protein and their interactions with A- and B-branch cofactors are given on a single row. All other rows provide interactions with the aligned residues considering the interactions of A-branch residues with nearby A-branch cofactors or B-branch residue with B-branch cofactors. In each case at least one of the residues in the row has an interaction > ±29 meV with its cofactor. Individual residue interactions are calculated with the mean field energy approximation (S.I. Eqn 2).

Conservation is the percent identity of the residue found at this position (with *R. sphaeroides* residue numbering) from within the 250 BLASTp sequences from the multi-sequence alignment. The conservation is for the symmetry-related residues on the same row, with the left column for the residue interacting with the B-side cofactor and the right column for the residue interacting with the A-side cofactor.

aa: The alternative residue found when a residue is not fully conserved are given.The numerical value gives the probability of that residue being found. Residue type in parenthesis and no number indicates the probability is <5%.

Conservation A-branch residues		Residues influencing BPhs A and BPhs B					Conservation B-branch residues	
aa	aa	BPh_B	meV	BPh_A	meV	diff	%cons	aa
		Fe^{3+}	-293	Fe^{3+}	-310	-17		
E:100	1	GluM234	78	GluM234	81	3	1	E:100
R:99;(H)	99	ArgM132	-175	ArgL103	-165	10	1	R:100
R:99;(Q)	99	ArgM164	-45	ArgL135	-52	-7	1	R:100
E:100	1	ArgL217	-52	ArgM253	-58	-6	0.99	R:99;(-)
R:100	1	ArgL231	-52	ArgM267	-36	16	0.98	R:99;(-,H)
R:99;(-,H)	99	ArgM267	-48	ArgL231	-49	-1	1	R:100
T:53;S:34;(V,A,C,M)	0.53	ThrM133	-6	GluL104	-65	-59	0.91	E:91;(1:8;(M)
D:100	1	AspL218	55	TrpM255	0	-55	0.98	T:98;(-,V,C0
L:36;V:15;R:14;(T,N,M,I,,F,Q,D,E,K,Y,D,H,W))	0.36	LeuM135	0	GluL106	65	65	1	E:100
A:38;Q:34;E17;(D,N,M,R,S,K)	0.38	AlaM139	0	LysL110	-37	-37	1	R:100
E:81;N:18	0.82	AspL213	51	GlyM242	0	-51	1	G:100
E:100	1	GluL212	84	ArgM241	-16	-100	1	R:100
R:100	1	ArgM241	-7	GluL212	40	47	1	E:100
R:99;(H)	0.99	GluM232	34	na	na	-34		
		gap	na	GluM232	37	37	1	R:99;(H)
		gap	na	ArgM247	-35	-34	0.99	R:99;(G)
		gap	na	GluL6	46	45	0.99	E:98;(0)
		GluH173	39	none	na	-38		
R:59;L:41		ArgM136	-53	ArgL10	-25	28	0.99	R:99;(-)
		Sum	-390	Sum	-579	-189		

S.I. Table 5: Change in the E_ms of three parent structures by the YF swap imposed within MCCE.

(mV)	1AIG	1AIG YF	*1AIG* $\Delta E_{m,yf-}$	2GNU	2GNU YF	*2GNU* $\Delta E_{m,yf-}$	2J8C	2J8C YF	*2J8C* $\Delta E_{m,YF}$
$BChl_B$	-920	-839	*81*	-951	-916	*35*	-1050	-881	*169*
$BChl_A$	-811	-911	*-100*	-803	-813	*-10*	-847	-1114	*-267*
$\Delta E_{m,A-B}$	*109*	*-72*		*180*	*103*		*203*	*-233*	
BPh_B	-336	-369	*-33*	-361	-416	*-55*	-367	-350	*17*
BPh_A	-250	-247	*3*	-293	-322	*-29*	-321	-261	*60*
$\Delta E_{m,A-B}$	*86*	*122*		*68*	*94*		*46*	*89*	

In the YF swap TryL210 is replaced by PheL210 and PheM181 is replaced by TyrM181. No additional relaxation of the structure beyond MCCE side chain optimization is performed.

S.I.Table 6: E_m shift in YF swap bRCs relaxed within a Molecular Dynamics trajectory.

	$BChl_B$	$BChl_A$	$\Delta E_{m,A-B}$	BPh_B	BPh_A	$\Delta E_{m,A-B}$
1AIG	-911	-831	80	-366	-250	116
snapshot80	-940	-1040	-100	-436	-387	49
snapshot128	-700	-991	-291	-433	-403	30
snapshot186	-773	-817	-44	-411	-362	49
snapshot180	-838	-1102	-264	-449	-429	20
snapshot86	-727	-871	-144	-411	-362	49
Ave(snapshots)	-796±96	-964±118	*-168*	-428±17	-389±29	*39*

S.I. Table 7: The energy terms that contribute to the E_ms difference of different cofactors from the five YF MD snapshots.

(meV)	$BChl_B$	$BChl_A$	BPh_B	BPh_A
Backbone	-93±45	-101±23	-92±28	-101
Desolvation	162±91	179±72	417±33	447
Residues/Cof	-128±41	27±63	-487±49	-522

Reference:

1. J. Mao, K. Hauser and M. R. Gunner, *Biochemistry*, 2003, **42**, 9829-9840.
2. R. J. Wei, U. Khaniya, J. Mao, J. Liu, V. S. Batista and M. R. Gunner, *Photosynth Res*, 2023, **156**, 101-112.
3. R. J. Wei, Y. Zhang, J. Mao, D. Kaur, U. Khaniya and M. R. Gunner, *Photosynth Res*, 2022.