Biophysical Journal, Volume *120*

Supplemental information

Role of the CarH photoreceptor protein environment in the modulation of cobalamin photochemistry

Courtney L. Cooper, Naftali Panitz, Travyse A. Edwards, and Puja Goyal

Role of the CarH Photoreceptor Protein Environment in the Modulation of Cobalamin Photochemistry

Supporting Material

Courtney L. Cooper,[†,¶] Naftali Panitz,[†,¶] Travyse A. Edwards,[‡] and Puja Goyal[*,†]

†*Department of Chemistry, State University of New York at Binghamton, PO Box 6000, Binghamton, NY, USA 13902*

‡*Department of Physics, State University of New York at Binghamton, PO Box 6000, Binghamton, NY, USA 13902*

¶*Contributed equally*

E-mail: pgoyal@binghamton.edu

Phone: +1 (607) 777-4308

Details of molecular dynamics (MD) simulations

Table S1: Details of the different independent MD simulations for the wild type (WT) and mutant proteins [a]

Trajectory index (Traj.)	Mutation	N-terminal restraints	Distance restraints	His142 protonation state
1	None	None	None	Neutral
2	None	Residues 5-75 k: 5 kcal/mol/Å^2	None	Neutral
3	None	None	None	Neutral
4	None	None	None	Protonated
5	E175Q	Residues 5-75 k: 1 kcal/mol/Å^2	OE1 (Gln175) - ND1(His177) r_0: 3.0 Å; k: 50 kcal/mol/Å^2	Neutral
6	E175Q	Residues 5-75 k: 1 kcal/mol/Å^2	OE1 (Gln175) - ND1(His177) r_0: 3.0 Å; k: 50 kcal/mol/Å^2	Neutral
7	E175D	Residues 5-75 k: 1 kcal/mol/Å^2	CG (Asp175) - ND1(His177) r_0: 3.9 Å; k: 50 kcal/mol/Å^2	Neutral
8	E175D	None	CG (Asp175) - ND1(His177) r_0: 3.9 Å; k: 50 kcal/mol/Å^2	Protonated

[a]The k values indicate the force constants for the harmonic restraints and the r_0 values indicate the equilibrium distances for the harmonic restraints. The WT or mutated crystal structure was used as the reference structure for the application of restraints. Any restraints were only applied during the initial energy minimization, 1 ns NVT equilibration and an additional 50 ns NPT equilibration. Subsequently, all restraints were released and NPT simulations were carried out. The data presented in the figures and tables correspond to the completely unrestrained NPT simulations.

Methods: Effect of implicit vs. explicit solvent on excited states

To test the effect of explicit solvent vs. implicit solvent on the excited electronic states of 5′-deoxy-5′-adenosylcobalamin (AdoCbl), the following procedure was adopted. The structure at 16 ns from traj. 1 was extracted and all atoms except those belonging to AdoCbl were deleted. The lower axial ligand, His177, was capped with acetyl groups at the N- and C-termini. AdoCbl was then solvated in

a water box using the *psfgen* utility in VMD,[1] with the minimum distance between an AdoCbl atom and the edges of the water box being 15 Å. An energy minimization and a 500 ps MD simulation in the NVT ensemble at a temperature of 298 K were carried out with the AdoCbl atoms frozen in space. Subsequently, the energy minimization and MD simulation were repeated with weak harmonic restraints on the AdoCbl atoms with a force constant of 2.0 kcal/mol/Å^2. Finally, a 20 ns MD simulation with weak harmonic restraints on the AdoCbl atoms with a force constant of 2.0 kcal/mol/Å^2 was carried out in the NPT ensemble, at a temperature of 298 K and a 1 atm pressure. The final structure from this simulation was used to carry out a quantum mechanical/molecular mechanical (QM/MM) calculation, with the same QM region as described in the main text. Besides the point charges from the water molecules, the point charges of the corrin ring substituents were also included in the calculation. This structure was also used to calculate the excited electronic states for the QM region in implicit water using the polarizable continuum model (PCM),[2] with all point charges deleted, including those of the corrin ring substituents.

QM/MM partitioning

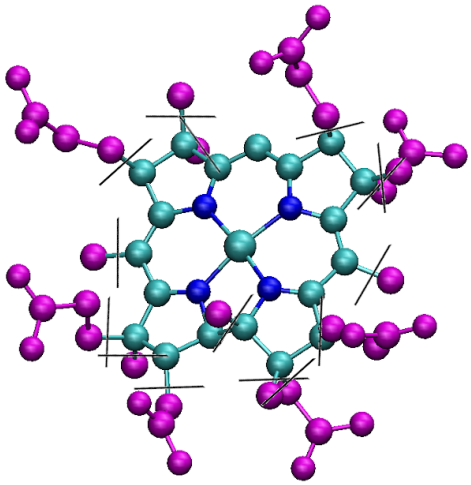

Figure S1: The black lines indicate the location of the QM/MM[3] boundary for the corrin ring. All atoms colored in purple, along with their covalently bonded H atoms were treated as point charges, while all other atoms, along with their covalently bonded H atoms were treated as QM.

Analysis of MD simulation data

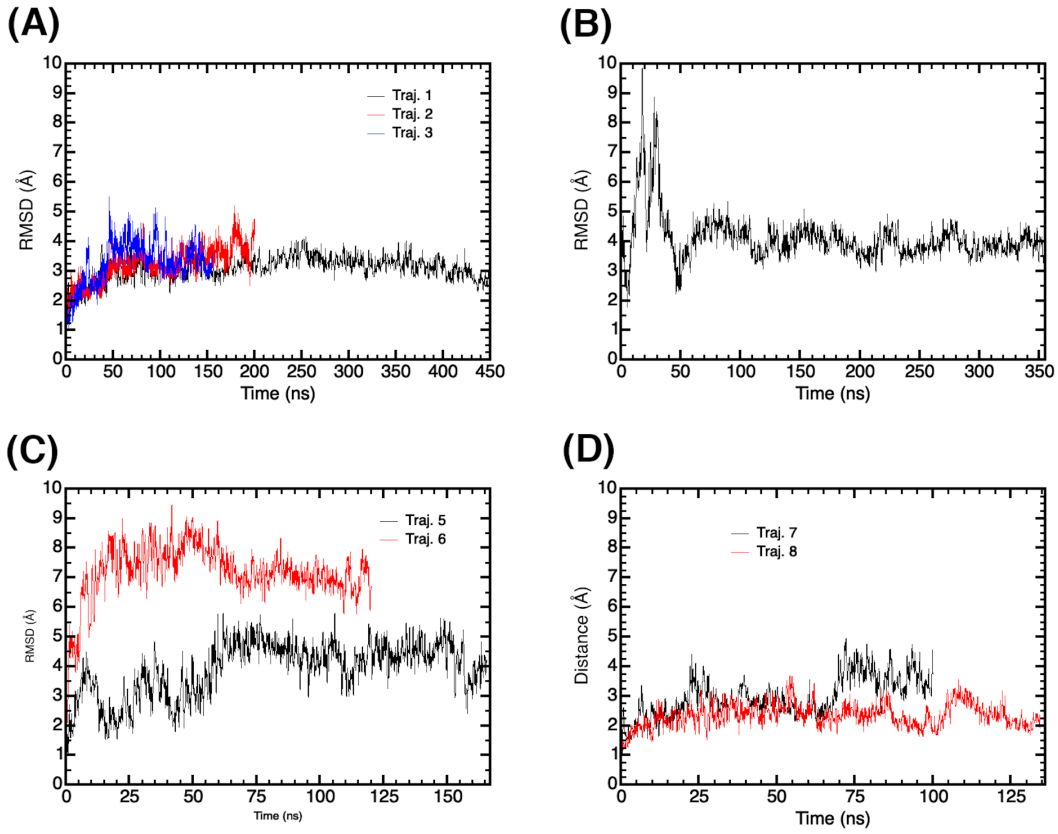

Figure S2: Root mean square deviation (RMSD) of the protein backbone heavy atoms with respect to the first frame in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142. The relatively large RMSD observed in traj. 6 results from the flexibility of the N-terminal region.

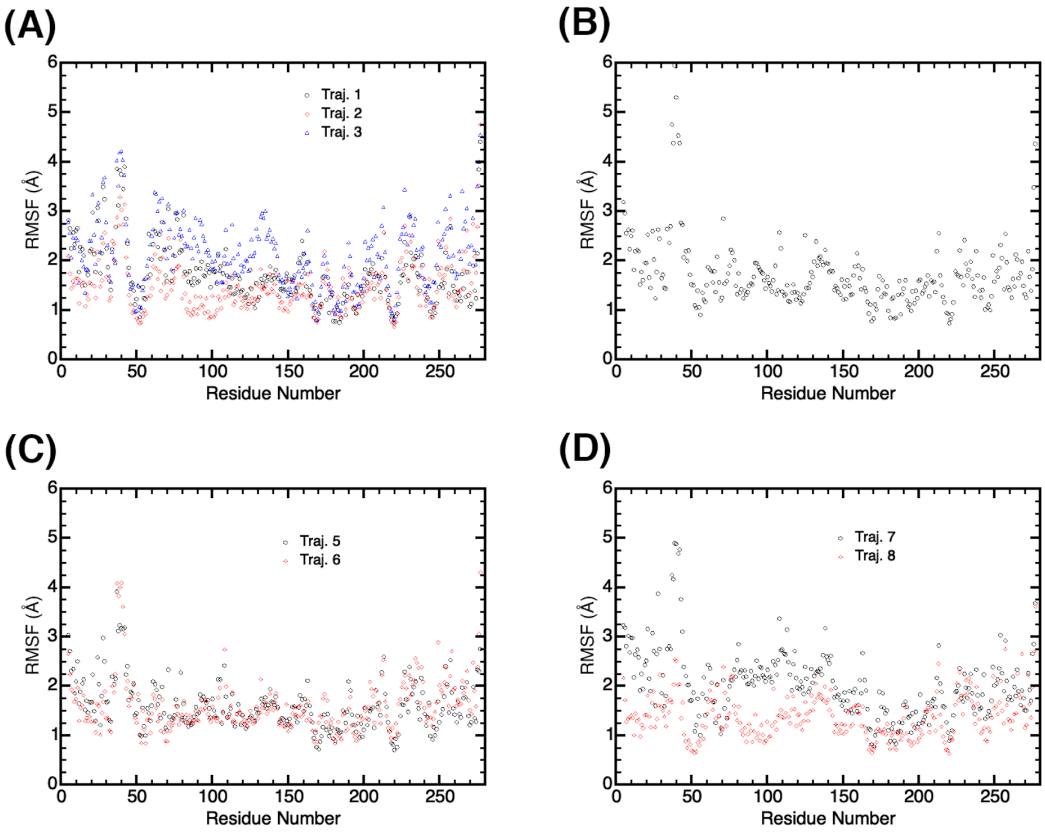

Figure S3: Root mean square fluctuation (RMSF) of the amino acid residues in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142. The RMSF is calculated with respect to the average positions of the atoms in the trajectory. The first 50 ns of each trajectory was excluded from the analysis.

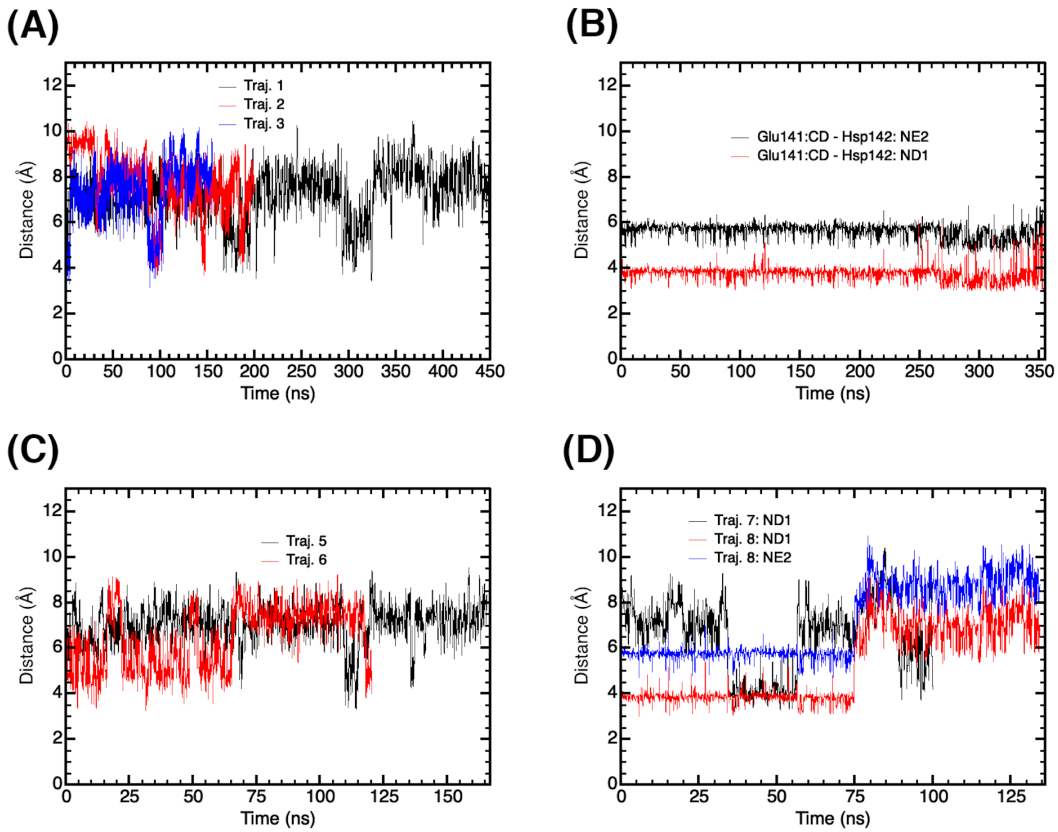

Figure S4: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the sidechain N atom(s) of His142 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142. Traj. 4 (panel (B)) and traj. 8 have His142 protonated at both the N_δ (ND1) and N_ϵ (NE2) sites while the other trajectories have His142 protonated only at the N_δ (ND1) site.

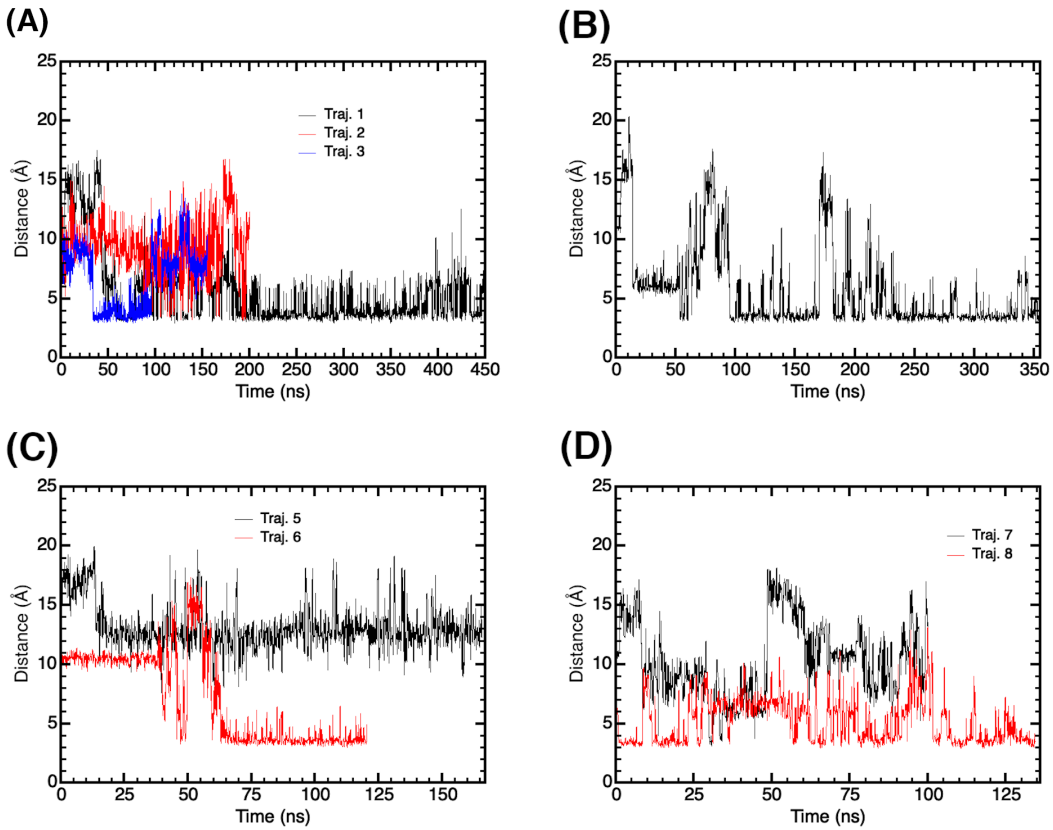

Figure S5: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the N6 atom (N atom of -NH_2 group, also see Fig. S6) of adenine in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

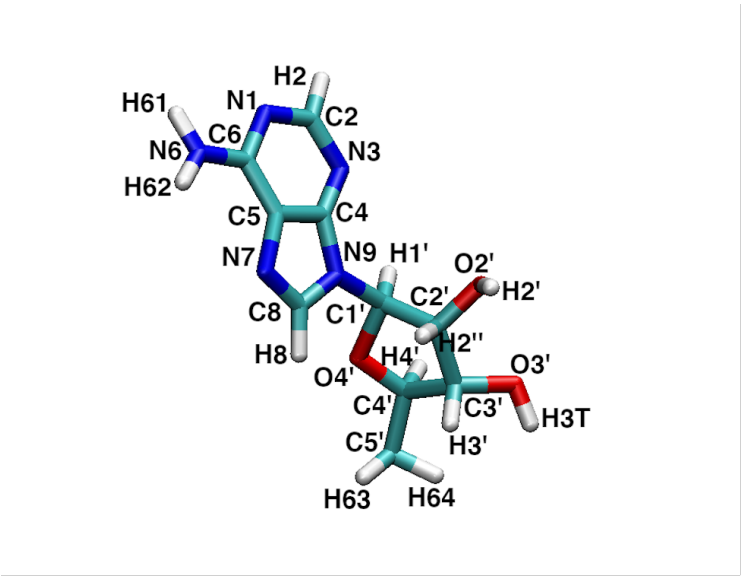

Figure S6: A depiction of the atom names in the adenosyl ligand, also referred to as 5AD.

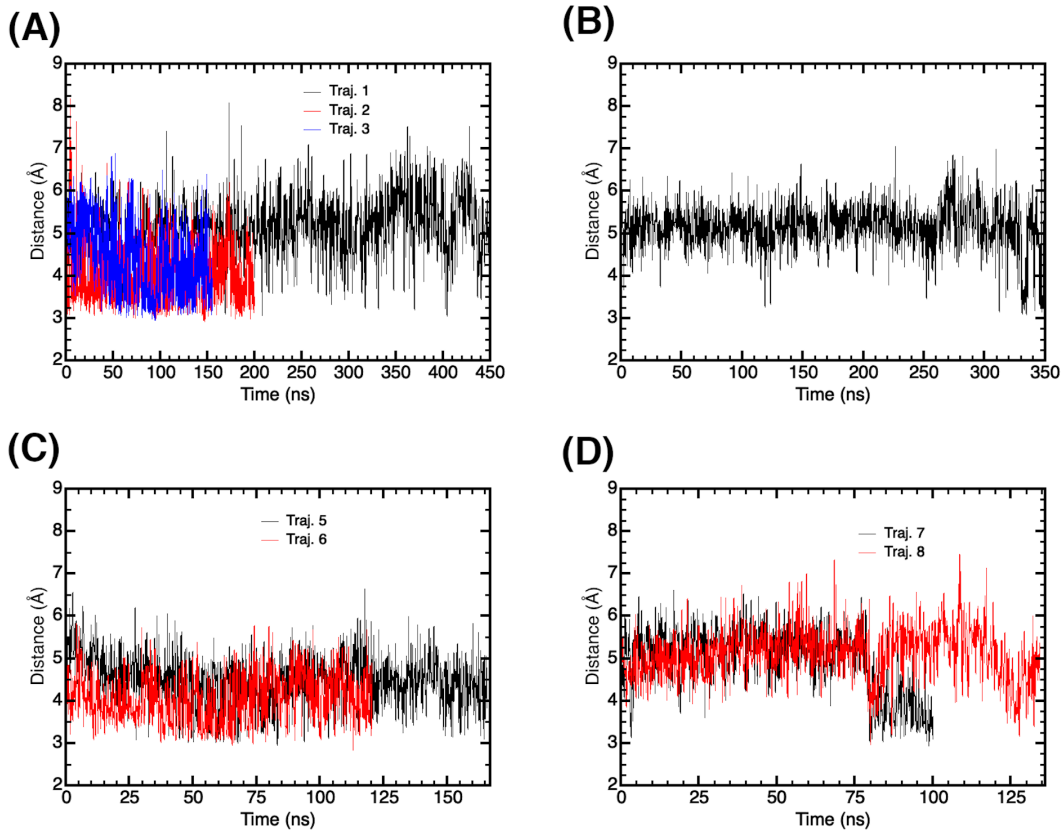

Figure S7: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the sidechain N atom of Trp131 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

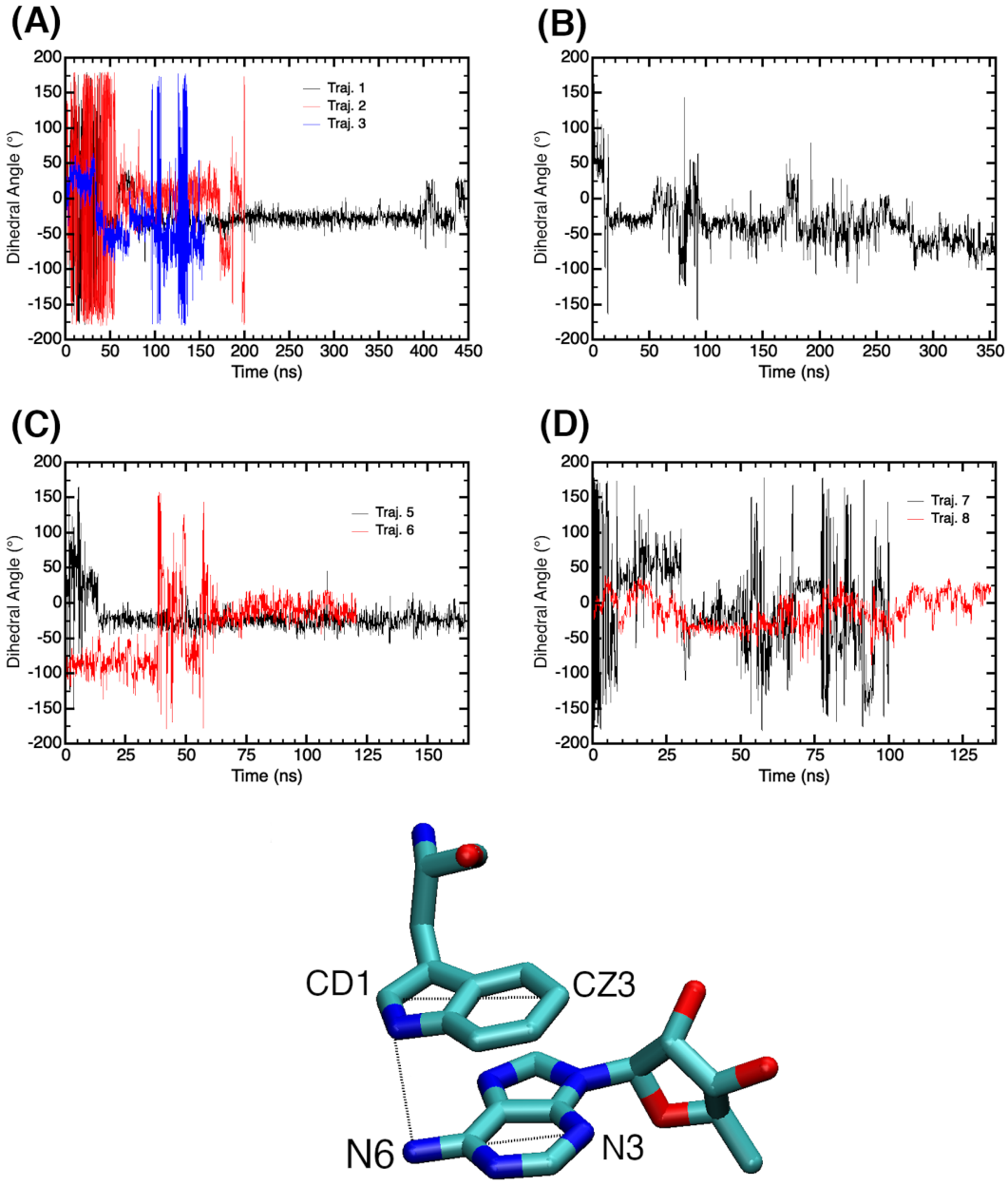

Figure S8: Time evolution of the CZ3-CD1-N6-N3 dihedral angle between Trp131 and adenosyl in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142. The dihedral angle under consideration is illustrated in the bottom image from which hydrogen atoms have been omitted for clarity.

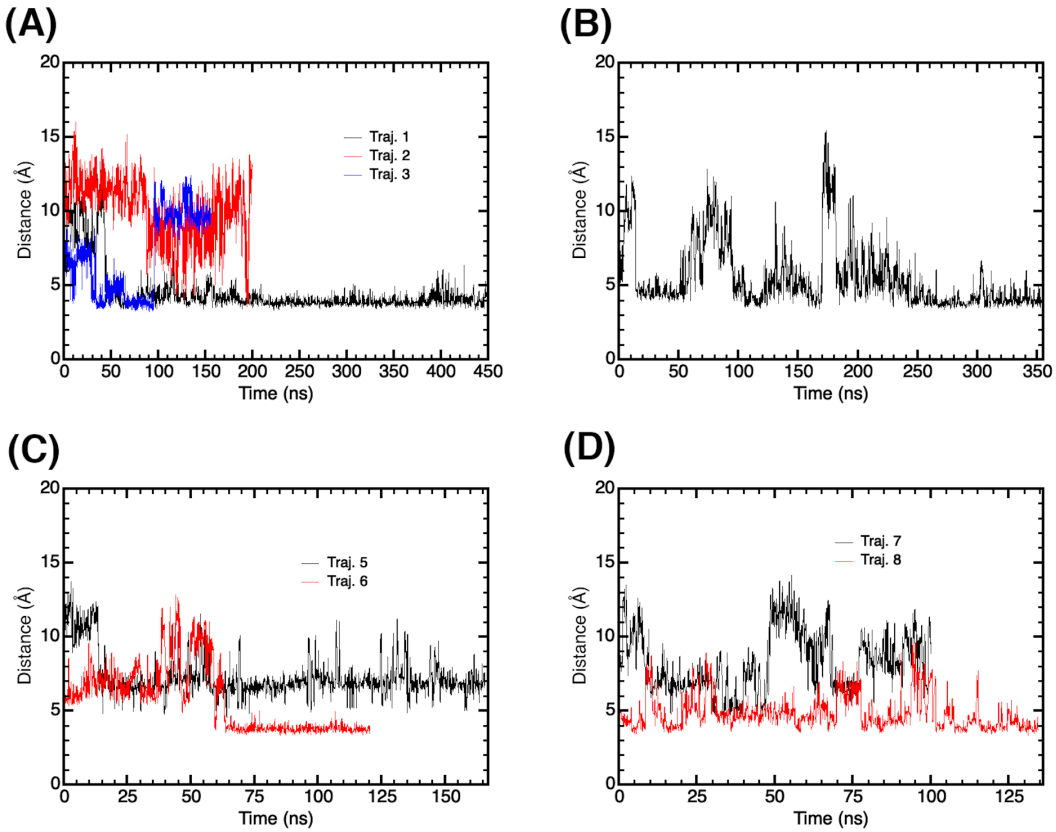

Figure S9: Time evolution of the distance between the center of mass of the heavy atoms of the Trp131 sidechain and the heavy atoms of the five- and six-membered rings of adenine in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

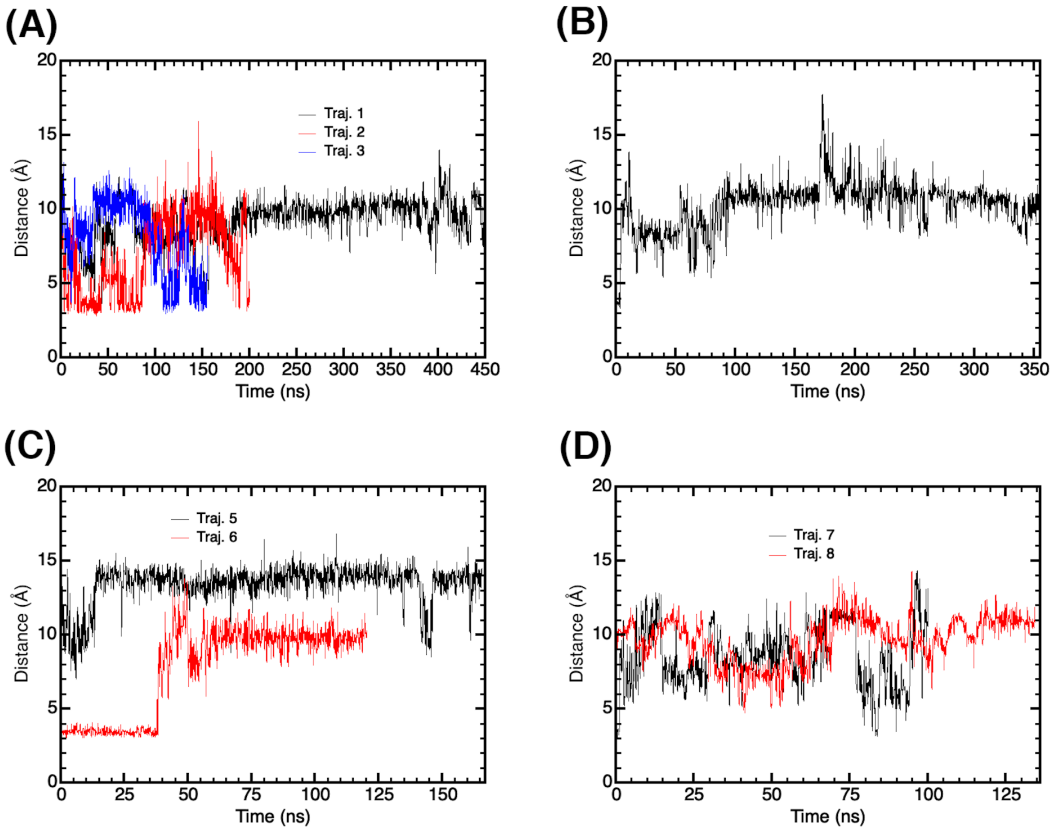

Figure S10: Time evolution of the distance between the ribose O2′ atom of adenosyl (see Fig. S6) and the carboxylate C atom (C_δ) of Glu141 for (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

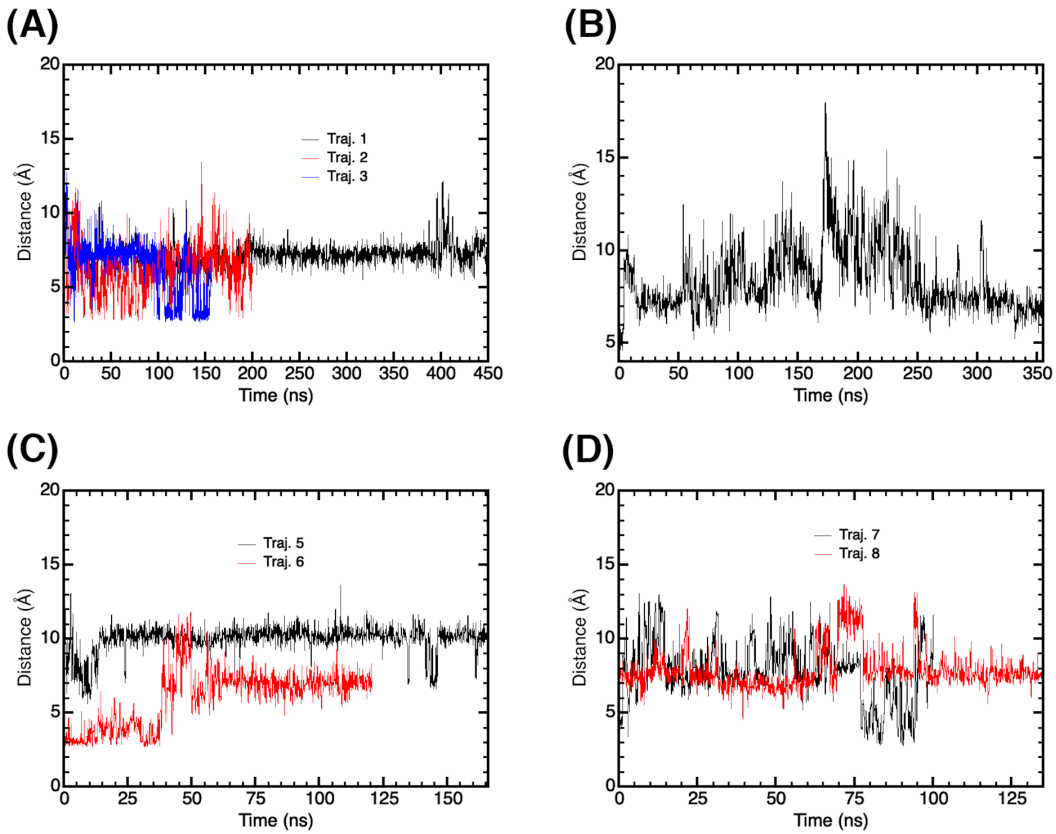

Figure S11: Time evolution of the distance between the ribose O2′ atom of adenosyl (see Fig. S6) and the side chain N atom of Trp131 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

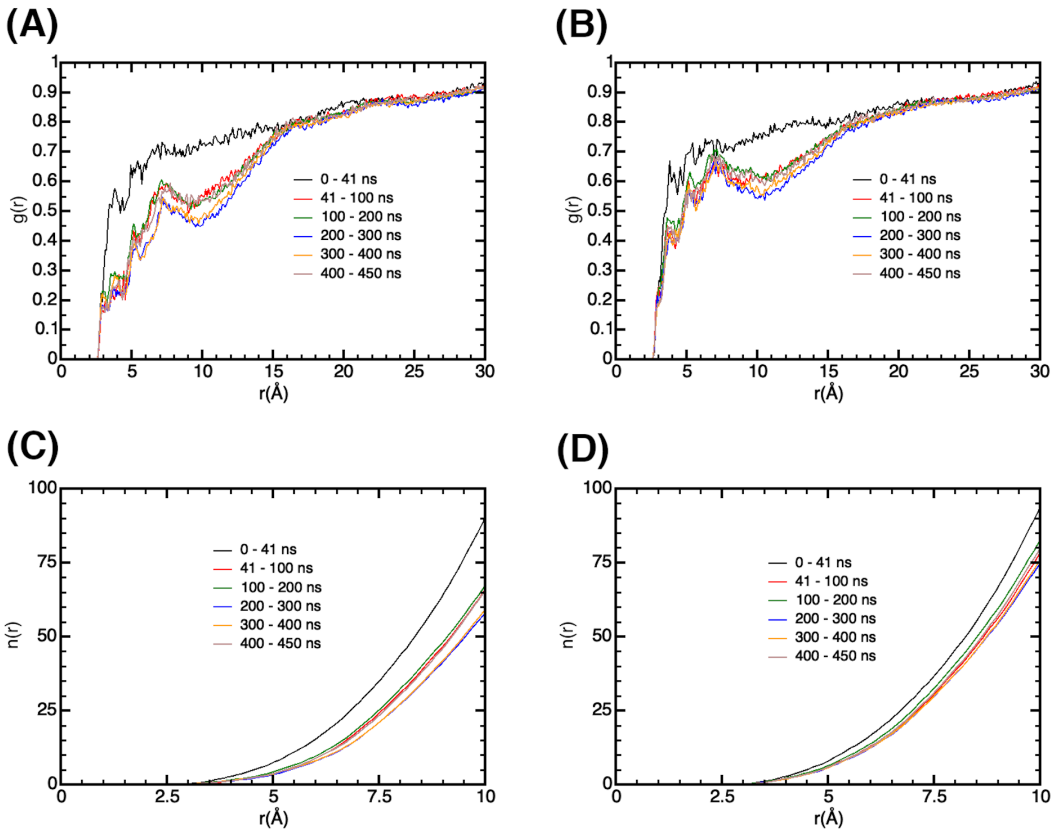

Figure S12: Radial distribution function (RDF) for water O atoms around the heavy atoms (excluding the N atom of the -NH_2 substituent) of the five-membered ring (A) and the six-membered ring (B) of the adenine moiety in traj. 1. The RDFs were examined for separate time intervals based on the time evolution of the Glu141-adenine distance. The RDFs remain similar beyond the 0-41 ns time interval, justifying the use of structures only from the first 100 ns of the trajectory for QM/MM calculations. The integrated RDFs corresponding to panels (A) and (B) are depicted in panels (C) and (D), respectively. The RDFs were calculated using the Radial Pair Distribution Function extension in VMD.[4]

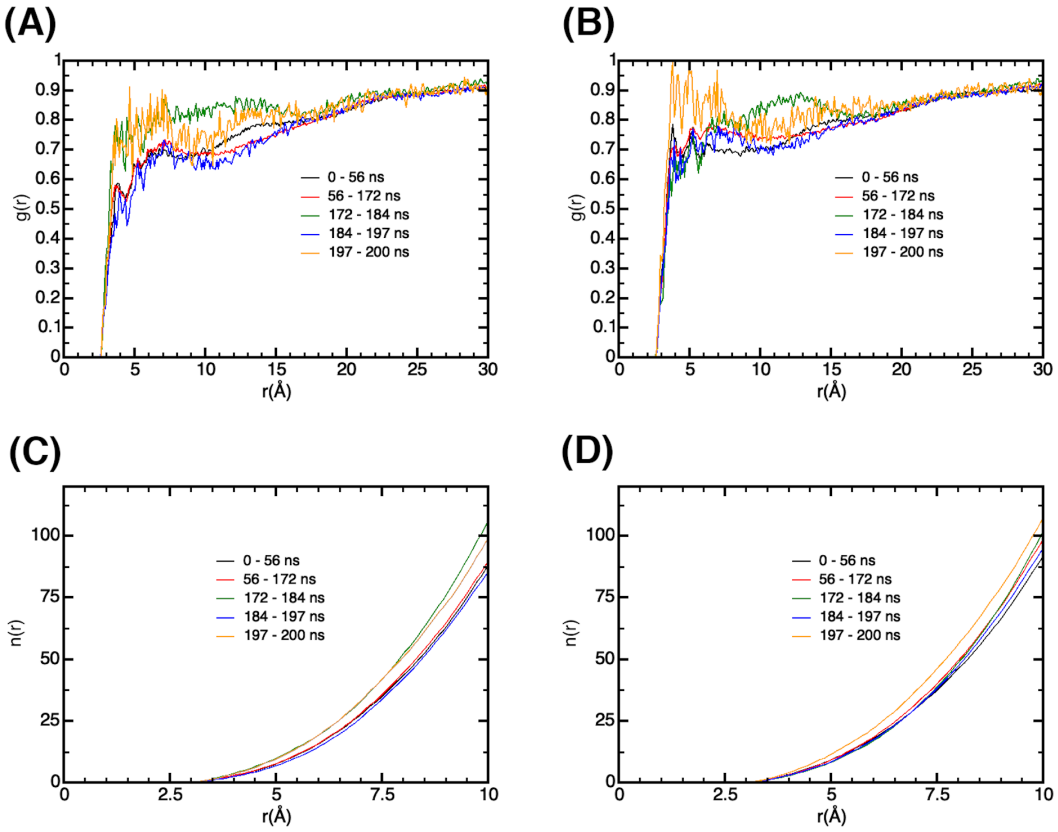

Figure S13: RDF for traj. 2 calculated using the same procedure as for Fig. S12.

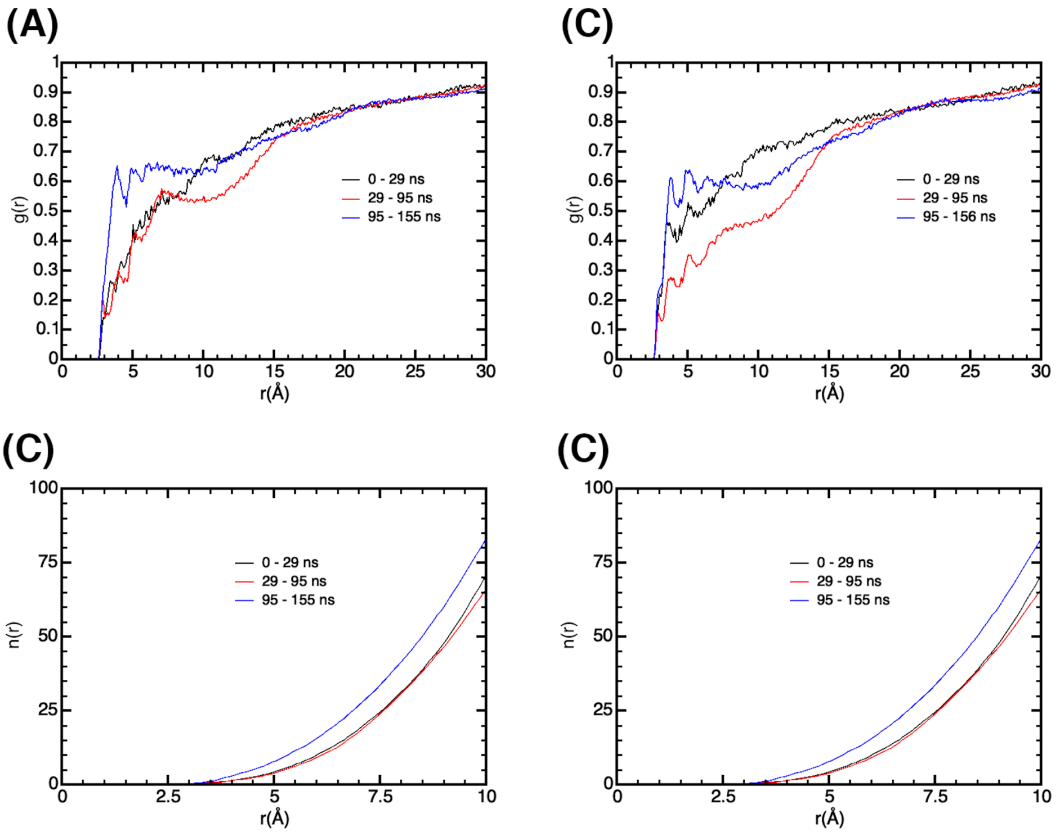

Figure S14: RDF for traj. 3 calculated using the same procedure as for Fig. S12.

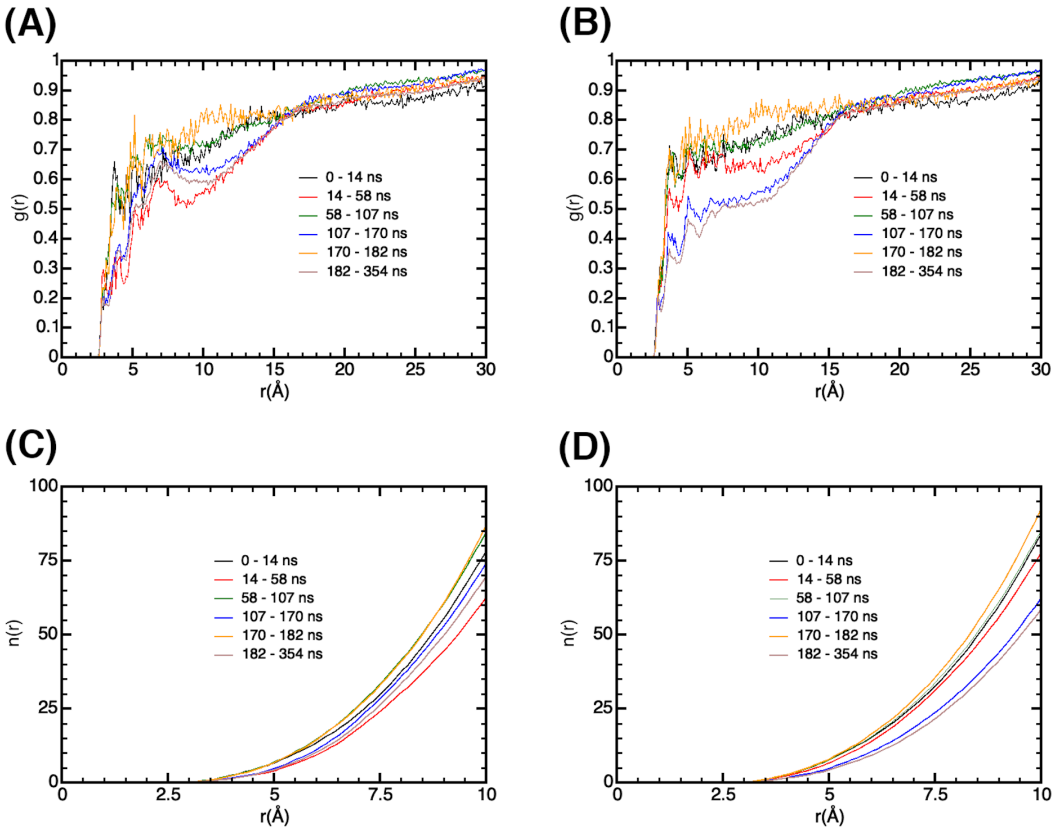

Figure S15: RDF for traj. 4 calculated using the same procedure as for Fig. S12.

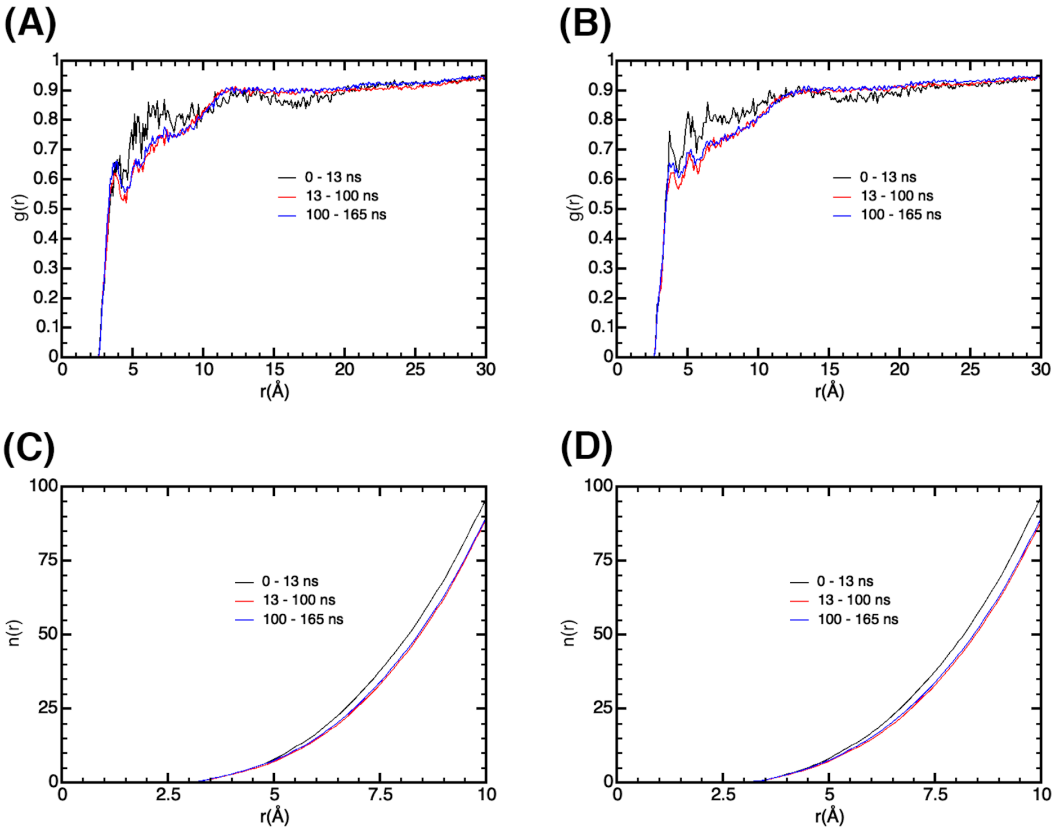

Figure S16: RDF for traj. 5 calculated using the same procedure as for Fig. S12.

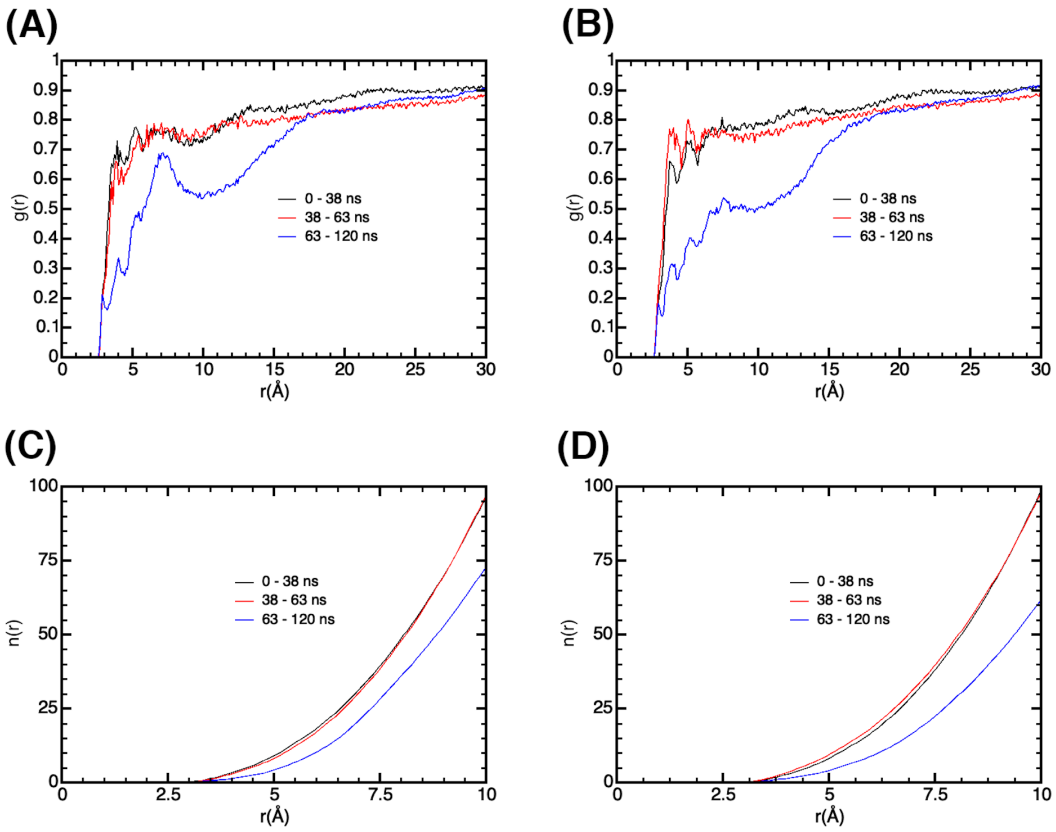

Figure S17: RDF for traj. 6 calculated using the same procedure as for Fig. S12.

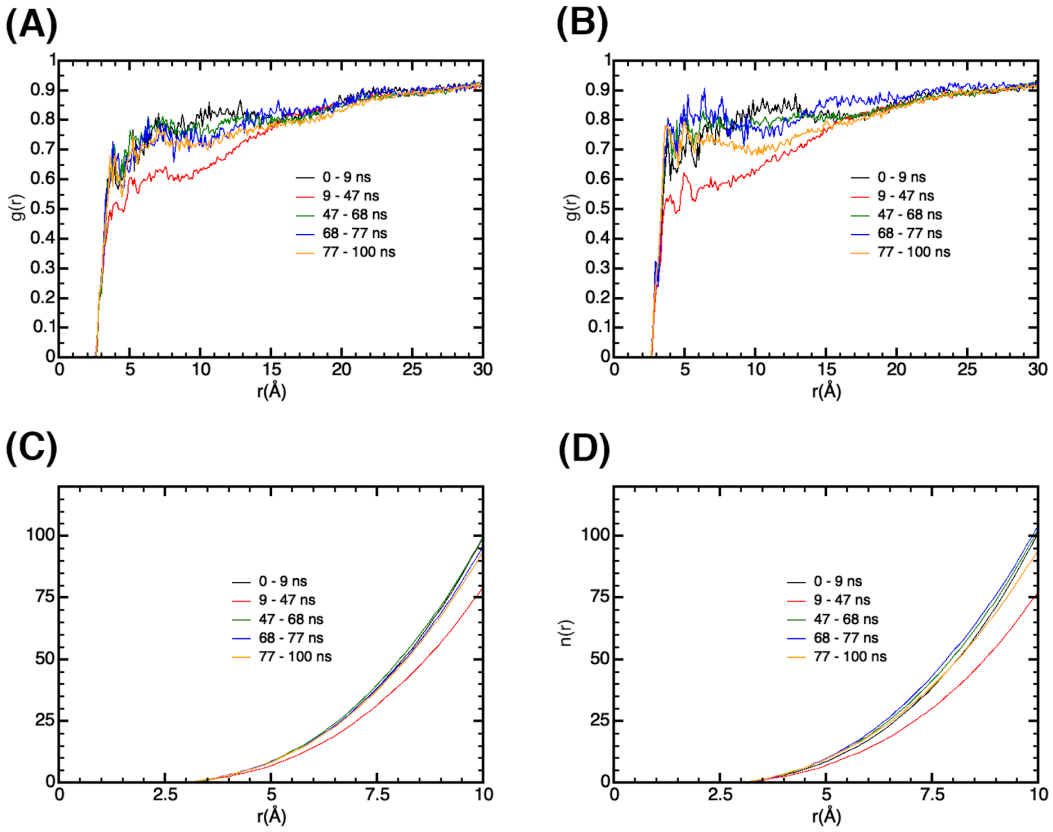

Figure S18: RDF for traj. 7 calculated using the same procedure as for Fig. S12.

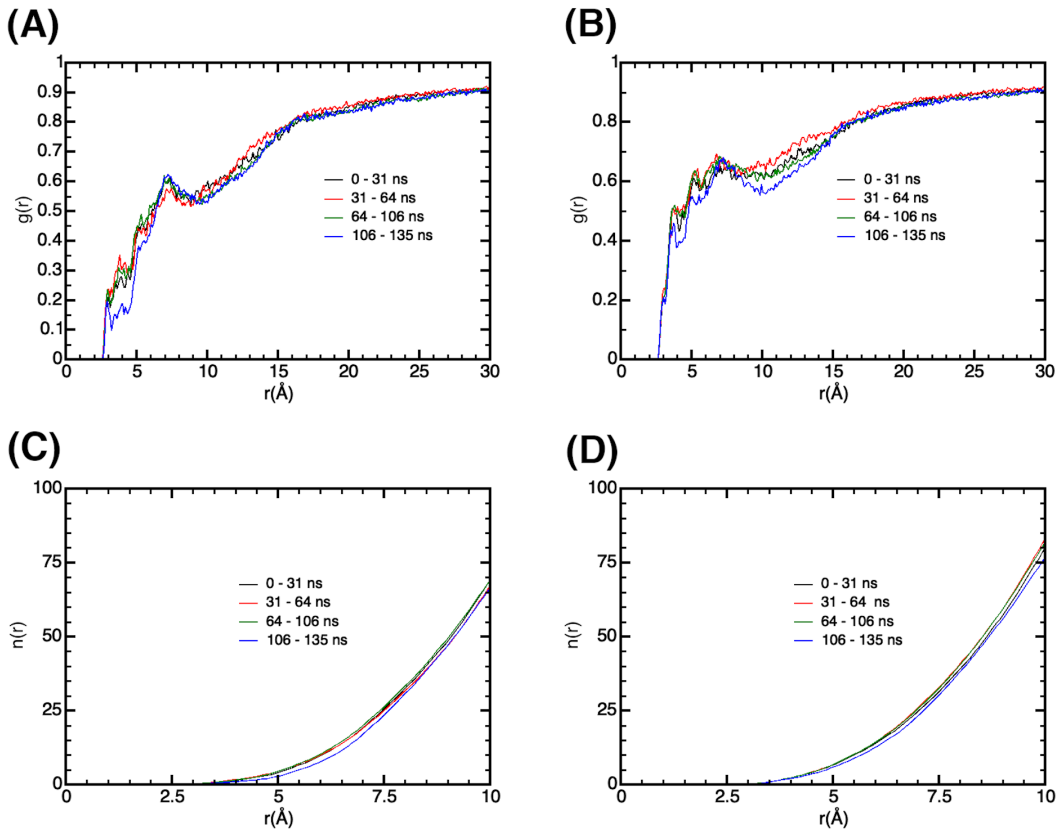

Figure S19: RDF for traj. 8 calculated using the same procedure as for Fig. S12. Adenine and Trp131 remain π-stacked for the majority of the 135 ns simulation.

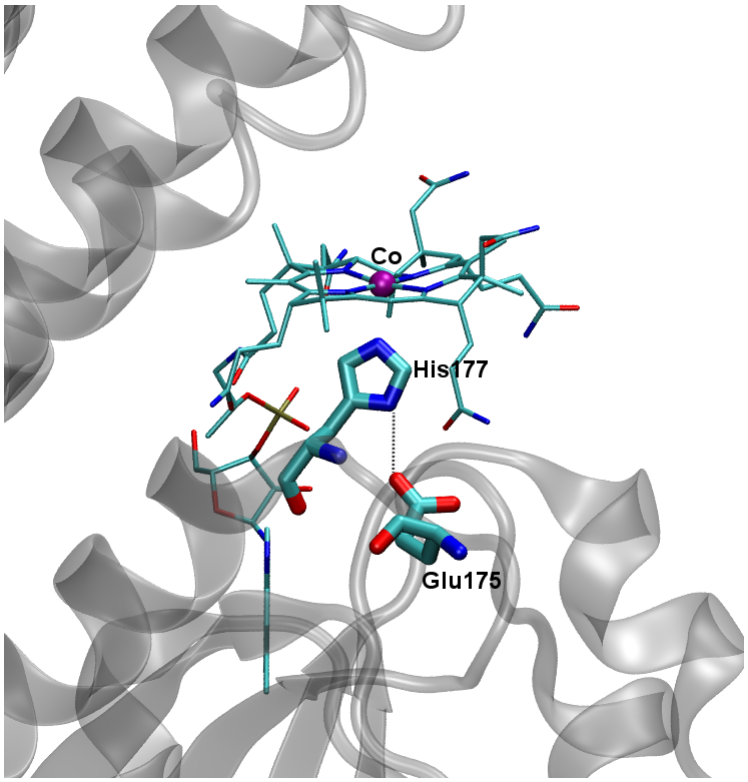

Figure S20: An illustration of the hydrogen bond between His177 (the lower axial ligand of Co) and Glu175 in CarH. In chain A of the crystal structure (PDB ID 5C8E),[5] the distance depicted as the dotted line is 2.92 Å. The hydrogen atoms have been omitted for clarity.

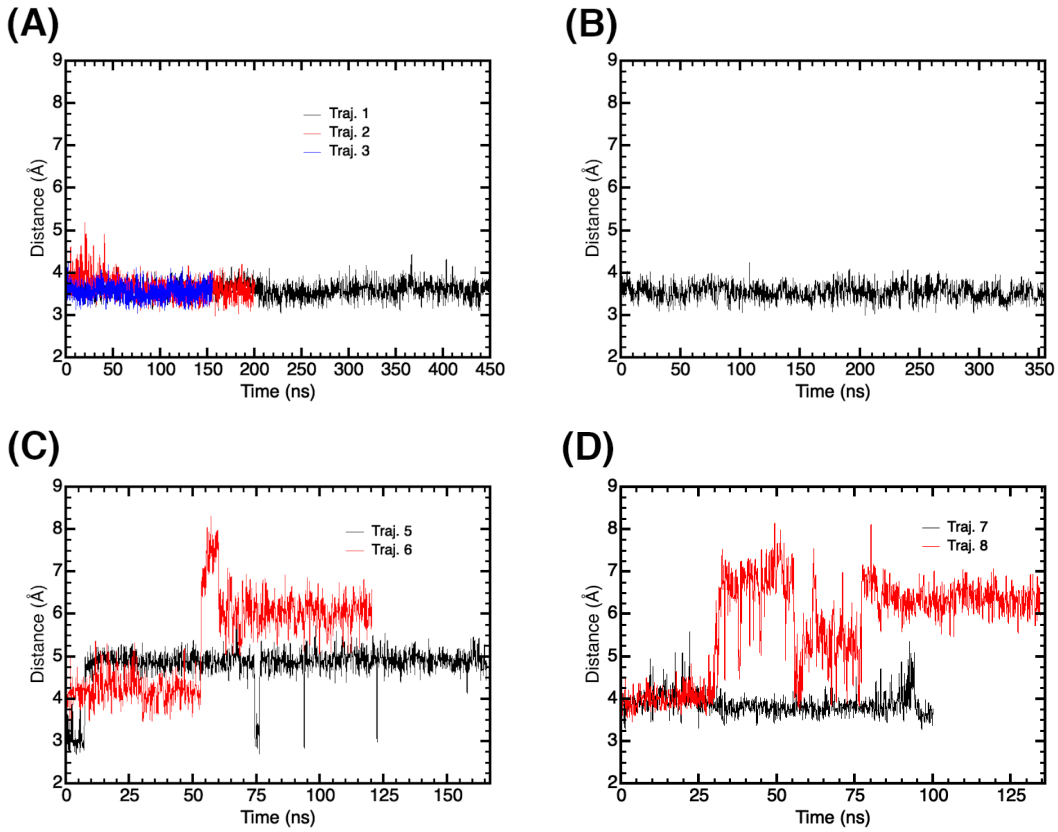

Figure S21: Time evolution of the distance between the N_δ (ND1) atom of His177 and a selected sidechain atom of residue number 175 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142. Residue 175 corresponds to Glu175 in panels (A) and (B), Gln175 in panel (C), and Asp175 in panel (D). For Glu175, Gln175 and Asp175, the C_δ (CD) atom, an O_ϵ (OE1) atom and the C_γ (CG) atom, respectively, were used for the distance measurements.

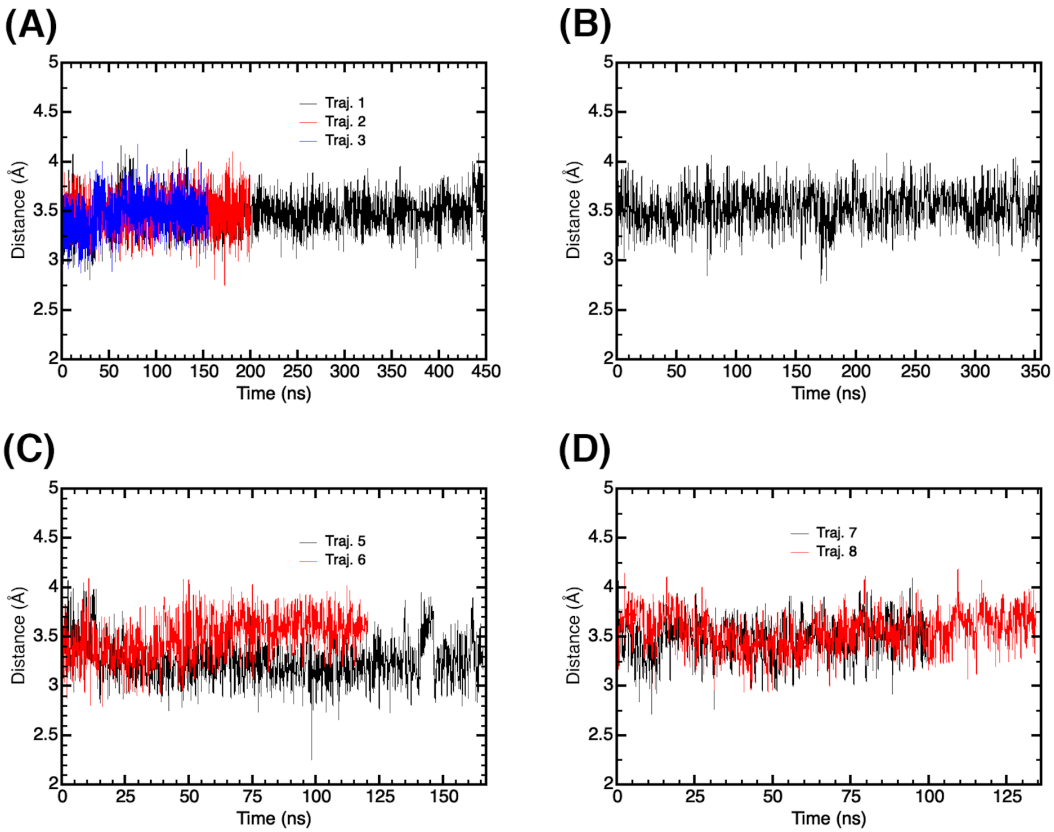

Figure S22: Time evolution of the distance between the H4′ atom of adenosyl (see Fig. S11) and Co in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 8 having a protonated His142.

QM/MM calculations: Effect of environment, solvent treatment, basis set size and mutations

Table S2: Effect of environment, explicitly solvated protein vs. implicit water, on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Time (ns)	Env.	Co/corrin→Ade	Co/corrin→Ade/σ^*	Co→σ^*
4	Protein	3 (S_{34})	0	9 (S_{12})
	Solvent	7 (S_{26})	0	5 (S_{18})
8	Protein	14 (S_4)	2 (S_5)	9 (S_8)
	Solvent	7 (S_{16})	0	11 (S_7)
12	Protein	10 (S_8)	0	10 (S_3)
	Solvent	8 (S_{12})	0	11 (S_4)
16	Protein	12 (S_6)	1 (S_{28})	8 (S_3)
	Solvent	11 (S_{14})	0	10 (S_3)
20	Protein	15 (S_4)	0	12 (S_7)
	Solvent	7 (S_{17})	0	12 (S_6)
24	Protein	13 (S_8)	1 (S_9)	11 (S_4)
	Solvent	8 (S_{20})	0	14 (S_7)
28	Protein	16 (S_4)	0	12 (S_7)
	Solvent	7 (S_{25})	0	10 (S_9)
32	Protein	16 (S_5)	4 (S_{16})	8 (S_7)
	Solvent	7 (S_{24})	0	13 (S_8)
36	Protein	8 (S_{14})	0	11 (S_4)
	Solvent	7 (S_{16})	0	13 (S_{10})
40	Protein	9 (S_{15})	0	13 (S_4)
	Solvent	6 (S_{27})	0	16 (S_9)
44	Protein	13 (S_4)	0	10 (S_7)
	Solvent	6 (S_{20})	0	11 (S_6)
48	Protein	11 (S_{12})	0	12 (S_4)
	Solvent	6 (S_{20})	0	14 (S_6)
52	Protein	17 (S_3)	0	10 (S_8)
	Solvent	10 (S_{14})	0	12 (S_9)
56	Protein	14 (S_5)	0	13 (S_8)
	Solvent	8 (S_{17})	1 (S_{54})	10 (S_9)
60	Protein	7 (S_{13})	0	13 (S_2)
	Solvent	7 (S_{17})	0	16 (S_2)
64	Protein	11 (S_{12})	3 (S_{29})	9 (S_4)
	Solvent	11 (S_{21})	1 (S_{43})	10 (S_6)
68	Protein	17 (S_2)	3 (S_{16})	10 (S_6)
	Solvent	6 (S_{15})	1 (S_{14})	10 (S_8)
72	Protein	20 (S_2)	2 (S_{54})	9 (S_7)
	Solvent	9 (S_{16})	0	15 (S_4)
76	Protein	11 (S_9)	1 (S_{30})	8 (S_3)
	Solvent	7 (S_{22})	0	16 (S_3)
80	Protein	9 (S_{10})	4 (S_8)	13 (S_3)
	Solvent	7 (S_{18})	0	16 (S_3)
84	Protein	11 (S_7)	0	12 (S_4)
	Solvent	7 (S_{15})	0	10 (S_{17})
88	Protein	14 (S_8)	1 (S_{48})	15 (S_4)
	Solvent	9 (S_{20})	0	15 (S_3)
92	Protein	13 (S_5)	1 (S_{41})	10 (S_2)
	Solvent	6 (S_{18})	0	9 (S_2)
96	Protein	20 (S_1)	2 (S_{57})	8 (S_3)
	Solvent	10 (S_{10})	0	14 (S_2)
100	Protein	13 (S_7)	0	6 (S_9)
	Solvent	8 (S_{15})	1 (S_{46})	8 (S_{11})

[a] All structures were taken from traj. 1. The index of the lowest energy excited state of each kind is indicated within parentheses. S_n indicates the n^{th} excited singlet state. Gas phase results are included in the supporting Excel sheet. In the gas phase, most structures show no transitions involving Co/Corrin to Ade charge transfer.

Table S3: Effect of implicit vs. explicit solvent on the number of different kinds of electronic transitions within 60 lowest-lying excited states for the structure at 16 ns of traj. 1.[a]

Transition	Implicit	Explicit
Co/corrin→Ade	9 (S_{12}, 451.8)	7 (S_{12}, 437.4)
Co/corrin→Ade/σ^*	0	0
Co→σ^*	11 (S_5, 520.8)	9 (S_4, 531.3)
σ→σ^*	0	0

[a]The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses.

Table S4: Effect of basis set size on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Time (ns)	Transition	Def2SVP	Def2TZVP
32	Co/corrin→Ade	16 (S_5, 487.8)	19 (S_4, 503.4)
	Co/corrin→Ade/σ^*	4 (S_{16}, 414.5)	3 (S_5, 497.6)
	Co→σ^*	8 (S_7, 478.2)	9 (S_7, 481.0)
	σ→σ^*	0	0
72	Co/corrin→Ade	20 (S_2, 706.8)	23 (S_2, 716.1)
	Co/corrin→Ade/σ^*	2 (S_{54}, 316.7)	1 (S_{25}, 406.3)
	Co→σ^*	9 (S_7, 579.1)	11 (S_7, 582.0)
	σ→σ^*	0	1

[a] Structures were taken from traj. 1. The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses.

Table S5: For the structure at 24 ns of traj. 1, effect of zeroing out the sidechain charge of selected residues on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Mutation	Co/corrin→Ade	Co/corrin→Ade/σ^*	Co→σ^*	$\sigma\rightarrow\sigma^*$
WT	13 (S_8, 445.3)	1 (S_9, 440.4)	11 (S_4, 491.0)	1 (S_{54}, 299.6)
Arg125	15 (S_6, 480.3)	1 (S_{56}, 298.1)	9 (S_7, 460.0)	0
Glu126	10 (S_{10}, 433.4)	2 (S_{37}, 328.9)	12 (S_3, 509.5)	0
Glu129	10 (S_{10}, 424.9)	3 (S_{15}, 397.8)	10 (S_3, 509.1)	0
Arg133	13 (S_8, 441.3)	0	14 (S_7, 460.3)	0
Glu141	13 (S_9, 441.7)	1 (S_{10}, 438.8)	9 (S_7, 510.2)	0
Glu175	3 (S_{41}, 360.5)	0	8 (S_8, 521.0)	0
Arg176	15 (S_6, 472.8)	1 (S_8, 442.9)	15 (S_3, 509.6)	0
Glu178	10 (S_{13}, 410.1)	1 (S_{56}, 301.9)	13 (S_8, 460.7)	0

[a]The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses. WT indicates the wild type protein with no sidechain charges zeroed out.

Table S6: For the structure at 32 ns of traj. 1, effect of zeroing out the sidechain charge of selected residues on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Mutation	Co/corrin→Ade	Co/corrin→Ade/σ^*	Co→σ^*	$\sigma\rightarrow\sigma^*$
WT	16 (S_5, 487.8)	4 (S_{16}, 414.5)	8 (S_7, 478.2)	0
Arg125	19 (S_4, 521.5)	2 (S_{13}, 438.5)	8 (S_7, 481.1)	0
Glu126	16 (S_7, 476.3)	4 (S_6, 479.6)	6 (S_{10}, 441.9)	0
Glu129	16 (S_8, 465.9)	1 (S_7, 473.7)	9 (S_6, 479.0)	0
Arg133	13 (S_7, 474.5)	1 (S_8, 469.2)	9 (S_5, 480.9)	0
Glu141	18 (S_4, 510.8)	0	10 (S_8, 477.8)	0
Glu175	7 (S_{24}, 379.7)	0	9 (S_6, 504.7)	0
Arg176	18 (S_4, 520.4)	2 (S_{11}, 444.0)	12 (S_8, 478.6)	0
Glu178	11 (S_9, 444.4)	0	12 (S_5, 478.3)	0

[a]The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses. WT indicates the wild type protein with no sidechain charges zeroed out.

Table S7: For the structure at 72 ns of traj. 1, effect of zeroing out the sidechain charge of selected residues on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Mutation	Co/corrin→Ade	Co/corrin→Ade/σ^*	Co→σ^*	$\sigma\rightarrow\sigma^*$
WT	20 (S_2, 706.8)	2 (S_{54}, 316.7)	9 (S_7, 579.1)	0
Arg125	22 (S_2, 685.3)	0	7 (S_7, 576.8)	0
Glu126	21 (S_1, 777.0)	3 (S_{12}, 518.8)	9 (S_7, 581.7)	1 (S_{15}, 488.0)
Glu129	23 (S_1, 773.4)	1 (S_{49}, 331.4)	10 (S_7, 582.5)	1 (S_{46}, 336.9)
Arg133	18 (S_2, 630.6)	2 (S_{31}, 371.3)	9 (S_6, 575.7)	1 (S_{12}, 471.1)
Glu141	25 (S_1, 776.5)	0	8 (S_{15}, 583.7)	0
Glu175	12 (S_8, 508.8)	0	11 (S_4, 580.7)	0
Arg176	24 (S_1, 784.5)	2 (S_8, 536.3)	9 (S_7, 584.1)	2 (S_{36}, 369.5)
Glu178	19 (S_2, 685.3)	0	8 (S_7, 578.0)	1 (S_{36}, 362.5)
Glu227	19 (S_2, 630.5)	1 (S_{16}, 451.6)	9 (S_6, 573.8)	1 (S_{12}, 472.5)

[a]The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses.

Table S8: For the structure at 96 ns of traj. 1, effect of zeroing out the sidechain charge of selected residues on the number of different kinds of electronic transitions within 60 lowest-lying excited states.[a]

Mutation	Co/corrin→Ade	Co/corrin→Ade/σ^*	Co→σ^*	$\sigma\rightarrow\sigma^*$
WT	20 (S_4, 596.9)	0	8 (S_3, 601.9)	0
Arg125	19 (S_1, 679.2)	2 (S_{47}, 335.1)	7 (S_6, 600.7)	0
Glu126	16 (S_2, 572.5)	1 (S_{17}, 440.5)	7 (S_3, 604.2)	1 (S_{59}, 303.2)
Glu129	17 (S_2, 563.5)	1 (S_{23}, 407.3)	7 (S_3, 604.6)	0
Arg133	17 (S_2, 584.0)	0	8 (S_3, 599.2)	0
Glu141	24 (S_1, 688.4)	0	8 (S_{10}, 610.0)	0
Glu175	11 (S_4, 555.7)	0	10 (S_2, 604.2)	0
Arg176	18 (S_1, 628.8)	0	10 (S_5, 602.3)	0
Glu178	20 (S_2, 553.9)	0	10 (S_3, 602.7)	0
Glu227	17 (S_2, 574.9)	0	11 (S_3, 598.4)	0

[a]The index and excitation wavelength (in nm) of the lowest energy excited state of each kind is indicated within parentheses.

Natural transition orbital images

Images of natural transition orbitals corresponding to all the calculated excited states in all the time-dependent density functional theory[6] calculations performed are available in a Google Drive repository.

`https://drive.google.com/drive/folders/1cnldtIQksAtH0Lsts6Hn-JueJK3hbSfw?usp=sharing`

Supporting References

(1) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(2) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of *Ab Initio* Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117–129.

(3) Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. B.; Ribeiro, J. V.; Stone, J. E. et al. NAMD Goes Quantum: An Integrative Suite for Hybrid Simulations. *Nat. Methods* **2018**, *15*, 351–354.

(4) Levine, B. G.; Stone, J. E.; Kohlmeyer, A. Fast Analysis of Molecular Dynamics Trajectories with Graphics Processing Units - Radial Distribution Function Histogramming. *J. Comput. Phys.* **2011**, *230*, 3556–3569.

(5) Jost, M.; Fernández-Zapata, J.; Polanco, M. C.; Ortiz-Guerrero, J. M.; Chen, P. Y.-T.; Kang, G.; Padmanabhan, S.; Elías-Arnanz, M.; Drennan, C. L. Structural Basis for Gene Regulation by a B_{12}-dependent Photoreceptor. *Nature* **2015**, *526*, 536–541.

(6) Burke, K.; Werschnik, J.; Gross, E. K. U. Time-dependent Density Functional Theory: Past, Present, and Future. *J. Chem. Phys.* **2005**, *123*, 062206.