Biophysical Journal
Article

Biophysical Society

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

Courtney L. Cooper,[1] Naftali Panitz,[1] Travyse A. Edwards,[2] and Puja Goyal[1,*]
[1]Department of Chemistry and [2]Department of Physics, State University of New York at Binghamton, Binghamton, New York

ABSTRACT The photochemistry of cobalamins has recently been found to have biological importance, with the discovery of bacterial photoreceptor proteins, such as CarH and AerR. CarH and AerR, are involved in the light regulation of carotenoid biosynthesis and bacteriochlorophyll biosynthesis, respectively, in bacteria. Experimental transient absorption spectroscopic studies have indicated unusual photochemical behavior of 5′-deoxy-5′-adenosylcobalamin (AdoCbl) in CarH, with excited-state charge separation between cobalt and adenosyl and possible heterolytic cleavage of the Co-adenosyl bond, as opposed to the homolytic cleavage observed in aqueous solution and in many AdoCbl-based enzymes. We employ molecular dynamics and hybrid quantum mechanical/molecular mechanical calculations to obtain a microscopic understanding of the modulation of the excited electronic states of AdoCbl by the CarH protein environment, in contrast to aqueous solution and AdoCbl-based enzymes. Our results indicate a progressive stabilization of the electronic states involving charge transfer (CT) from cobalt/corrin to adenine on changing the environment from gas phase to water to solvated CarH. The solvent exposure of the adenosyl ligand in CarH, the π-stacking interaction between a tryptophan and the adenine moiety, and the hydrogen-bonding interaction between a glutamate and the lower axial ligand of cobalt are found to contribute to the stabilization of the states involving CT to adenine. The combination of these three factors, the latter two of which can be experimentally tested via mutagenesis studies, is absent in an aqueous solvent environment and in AdoCbl-based enzymes. The favored CT from metal and/or corrin to adenine in CarH may promote heterolytic cleavage of the cobalt-adenosyl bond proposed by experimental studies. Overall, this work provides novel, to our knowledge, physical insights into the mechanism of CarH function and directions for future experimental investigations. The fundamental understanding of the mechanism of CarH functioning will serve the development of optogenetic tools based on the new class of B_{12}-dependent photoreceptors.

SIGNIFICANCE The discovery of B_{12} (cobalamin)-based bacterial photoreceptor proteins, such as CarH, has revealed the biological importance of cobalamin photochemistry, which was earlier unknown. CarH is involved in light regulation of bacterial biosynthesis of carotenoids that prevent photo-oxidative damage. Experiments have indicated that CarH modulates the photochemistry of 5′-deoxy-5′-adenosylcobalamin (AdoCbl) differently compared with bulk water and enzymes, promoting excited-state charge transfer and possibly heterolytic cleavage of the cobalt-adenosyl bond. Our molecular dynamics and quantum mechanical/molecular mechanical calculations reveal that the AdoCbl environment in CarH exhibits unique features contributing to stabilization of excited electronic states involving long range charge transfer. The physical insights obtained provide directions for future experimental investigations and contribute to the fundamental knowledge that will serve development of B_{12}-dependent optogenetic tools.

INTRODUCTION

The discovery of cobalamin-binding photoreceptor proteins, such as CarH and AerR, that regulate transcription in bacteria has revealed the biological importance of cobalamin photochemistry, which was earlier unknown (1–7). For many years, the only known biological function of cobalamins or B_{12} derivatives was as cofactor to enzymes in humans, animals, and microbes, with the Co-C bond undergoing thermally-driven heterolytic (methylcobalamin) or homolytic (5′-deoxy-5′-adenosylcobalamin (AdoCbl)) cleavage to provide methyl cations or free radicals for enzymatic reactions (8,9). Excitation by light was also known to lead to cleavage of the Co-C bond in methylcobalamin and AdoCbl. This photolytic cleavage was, however, not believed to have biological relevance.

Submitted November 25, 2020, and accepted for publication July 20, 2021.
*Correspondence: pgoyal@binghamton.edu
Courtney L. Cooper and Naftali Panitz contributed equally to this work.
Editor: Alan Grossfield.
https://doi.org/10.1016/j.bpj.2021.07.020

CarH and AerR are involved in the light regulation of carotenoid biosynthesis and bacteriochlorophyll biosynthesis, respectively, in bacteria (10). AerR (3,4) shows tight binding for hydroxocobalamin, whereas CarH uses AdoCbl as a chromophore. CarH is a bacterial transcriptional repressor that controls the biosynthesis of carotenoids in response to light (1,2,5–7). Carotenoids protect bacteria from photo-oxidative damage by quenching the singlet oxygen free radicals produced by light absorption. In the dark, upon AdoCbl binding, CarH forms tetramers that bind operator DNA and block transcription. Upon light absorption, the tetramers dissociate, releasing DNA and allowing transcription and the subsequent synthesis of carotenoids that protect the bacteria against damage by oxygen free radicals (Fig. 1) (5,6).

Experimental studies have indicated the modulation of the photochemistry of AdoCbl by the CarH protein environment (6). In contrast to AdoCbl photochemistry in aqueous solution that leads to the cob(II):Ado$^{\cdot}$ radical pair through homolytic cleavage of the Co-C bond, photolysis of CarH-bound AdoCbl yields 4′,5′-anhydroadenosine, a species not observed before as a product of AdoCbl photolysis. Ultrafast spectroscopy experiments on CarH (6), which analyzed the photochemical mechanism on a timescale of femtoseconds to seconds, concluded that photoexcitation of CarH-bound AdoCbl is followed by the formation of a metal-to-ligand charge transfer (CT) state and, possibly, heterolytic cleavage of the Co-C bond. The cob(II):Ado$^{\cdot}$ radical pair formed from a small percentage of homolytic cleavage was found to undergo quantitative recombination. Hydride and proton transfers have also been implicated in the proposed mechanism (Fig. 2).

Despite the experimental studies, several details of the photochemical mechanism remain unknown and not easily accessible to experimental tools. Computational studies are indispensable for a fundamental detailed understanding of CarH photochemistry. CarH is inherently a very flexible biomolecule (5–7,10), and molecular dynamics (MD) simulations can provide information about the structural flexibility of CarH that is not available from the crystal structures. Quantum mechanical (QM) calculations are necessary to understand the photochemical bond breaking and formation reactions in CarH. To the best of our knowledge, this is the first computational study on CarH that involves a systematic investigation of the effect of the environment on the excited electronic states of AdoCbl with the protein (CarH) adequately solvated and its conformational flexibility accounted for (11). The goal of this study is to understand what features of the CarH protein environment contribute to the differences in AdoCbl photochemistry compared with aqueous solution and AdoCbl-based enzymes. The insights obtained will guide new experiments on CarH and contribute to the fundamental understanding of the mechanism of CarH functioning that will serve the development of optogenetic tools based on the new class of B_{12}-dependent photoreceptors. B_{12} absorbs green light, making it suitable for tackling the "blind spot" in optogenetics. Green-light-absorbing B_{12}-binding domains have been used to regulate signaling in human cells and development in zebrafish embryos (12). The development of a B_{12}-based system for conditional expression of essential genes in *Myxococcus xanthus* has also been pursued (13).

MATERIALS AND METHODS

Molecular dynamics simulations

MD simulations were carried out using the Nanoscale Molecular Dynamics (NAMD) software package version 2.13. (14). The initial structure of an AdoCbl-bound CarH monomer was taken from Protein Data Bank (PDB): 5C8E (5). NAMD was used to simulate the CarH monomer bound to AdoCbl in a water box, using periodic boundary conditions. Based on the solvent exposure and the hydrogen-bonding environment, all His residues were treated as neutral, with the N site protonated, and all Lys, Arg, Asp, and Glu residues were treated as being in their standard protonation states at pH 7. Because of the closeness of His142 to Glu141 in the crystal structure, both neutral and protonated states of His142 were studied. The minimum distance between a protein atom and the edges of the water box was 15 Å. Na^+ and Cl^- ions were added to neutralize the charge of the protein and to bring the salt concentration to 0.15 M. These were added using the Visual Molecular Dynamics solvate and autoionize plugins (15). The cutoff, switch, and pairlist distance values used with the switch scheme for treatment of Van der Waals interactions were 14, 12, and 16 Å, respectively. The particle mesh Ewald method (16) was used for electrostatic interactions, with the default particle mesh Ewald parameters in NAMD (14). The CHARMM36 force field parameters for the protein (17) and modified TIP3P model for water (18) were used. Force field parameters for the ions and AdoCbl were taken from (19) and (20), respectively. The force field parameters used for AdoCbl in our MD simulations have been extensively verified by the developers (20). The quality of the parameters was judged based on gas-phase-optimized geometries, the corrin fold angle in MD simulations of cobalamins in aqueous and biomolecular environments, and the stability of cobalamins in the binding sites of proteins and RNA during MD simulations. The QM and molecular mechanical (MM) geometries showed excellent agreement. The corrin fold angle was also reproduced well in MD simulations, and cobalamin remained stable in the binding sites of proteins/RNA.

The initial structure of the wild-type (WT) monomer was energy minimized and equilibrated for 1 ns at 298 K in an NVT ensemble in which

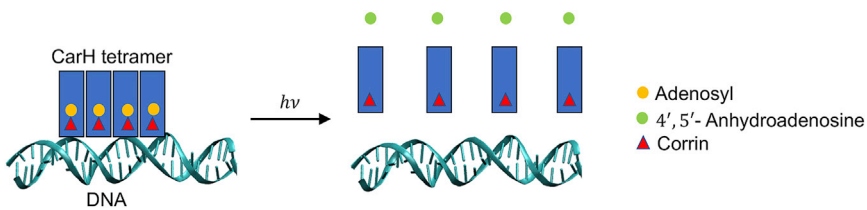

FIGURE 1 Dissociation of DNA-bound CarH tetramer in the presence of light, leading to DNA release and synthesis of carotenoids. Light excitation leads to Co-Ado bond cleavage and release of Ado from the protein. Once released from the protein, Ado gets converted to 4′,5′-anhydroadenosine. To see this figure in color, go online.

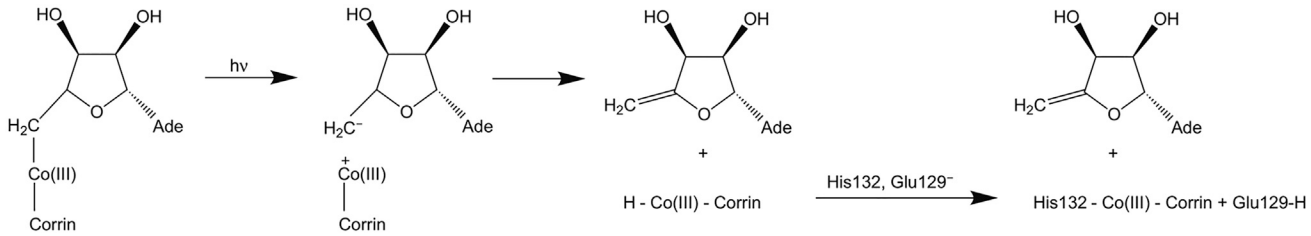

FIGURE 2 Proposed scheme of photochemical reactions in CarH (5–7). Postexcited-state charge transfer, the formation of 4′,5′-anhydroadenosine may be accompanied by hydride coordination to Co(III), with the resulting hydridocobalamin subsequently undergoing oxidation. X-ray crystallography has revealed His132 to be the final upper axial ligand of Co (5), which probably gets deprotonated by the nearby Glu129.

the volume and temperature were held constant. During the production simulations, an NPT ensemble was maintained in which the pressure and temperature were held constant at 1 atm and 298 K, respectively. For the E175Q and E175D mutants, the crystal structure of the AdoCbl-bound CarH monomer was mutated appropriately. Several independent simulations were carried out for both the WT protein and the mutants. Details of these simulations are provided in Table S1. The independent simulations for the WT protein, the E175Q mutant, and the E175D mutant were labeled traj. 1–4, traj. 5 and 6, and traj. 7 and 8, respectively. Unless noted otherwise, these simulations utilized the same protocol as outlined above. In cases where restraints on the N-terminal region, the Gln175-His177 distance or the Asp175-His177 distance were applied during equilibration, the NVT equilibration was followed by an additional 50 ns NPT equilibration.

Hybrid QM/MM calculations

Hybrid QM/MM calculations using a locally modified NAMD-Gaussian interface (21,22) were carried out to investigate the effect of the explicitly solvated protein environment on the excited electronic states of AdoCbl. 25 structures, at intervals of 4 ns, were selected from the first 100 ns of the traj. 1 WT simulation for these calculations. Additionally, six structures were also randomly selected after 100 ns of the traj. 3 WT simulation because of the sampling of shorter Glu141-sugar and Trp131-sugar distances compared with traj. 1. In the QM/MM calculation on each structure, the entire protein and primary water box along with the ions were retained. The water box was centered at AdoCbl, and its longest edge was aligned with the longest axis of the protein. The QM region included the central corrin ring without the substituents, Co, the upper axial adenosyl (Ado) ligand, and the side chain of the lower axial His177 ligand. The region of the corrin ring chosen to be studied quantum mechanically is depicted in Fig. S1. For His177, the boundary between the QM and MM regions was between the C_α and C_β atoms, with all backbone atoms treated as point charges.

All QM atoms at the QM/MM boundary were capped using H link atoms to satisfy their valence and to mirror the nonpolarity of the C-C bond. The redistributed charge scheme (23) implemented in NAMD was used to adjust the charges on the MM atoms at the QM/MM boundary. To maintain a net charge of +1 for the QM region (Co being in the +3 oxidation state), a negligible charge redistribution on the MM atoms was carried out. Based on the partial charges assigned to cobalamin atoms by the force field (20), the net charge of the QM atoms belonging to the corrin ring and Co was 0.954. The protein structure file was edited so that the charges of the 23 heavy atoms in this set were increased by 0.002 each, resulting in a net charge of +1 for the QM region. The charges of the 23 MM atoms farthest away from the QM region (the dimethylbenzimidazole moiety of corrin) were decreased by 0.002 each to conserve the total charge of AdoCbl.

Time-dependent density functional theory (TDDFT) (24) calculation for the QM region was carried out in the presence of the MM point charges, mostly using the BP86 functional and the Def2SVP basis set. A comprehensive benchmarking of density functionals has shown the BP86 functional to best describe the Co-C (upper axial ligand) dissociation energy and structural and spectral features of cobalamins (25–27). Test calculations of AdoCbl excited states in the gas phase and in implicit water with TD-BP86/def2SVP (data provided in Supporting materials and methods, Data S1, Excel sheet) agree well with those in (25) using BP86 and a Pople basis set for C, N, O, and H.

Tests with the B3LYP and CAM-B3LYP functionals and the larger Def2TZVP basis set were also carried out. The B3LYP and CAM-B3LYP functionals were chosen as representatives of the class of hybrid and range-separated functionals, respectively. For each selected structure, TDDFT calculation for the QM region was also carried out in the presence of implicit water using the polarizable continuum model (PCM) (28) and in the gas phase. For calculations in implicit water, linear response nonequilibrium solvation was used for the calculation of vertical excitation energies at the ground state geometries. The default self-consistent reaction field method in Gaussian09 and Gaussian16, PCM using the integral equation formalism PCM, was employed. For some structures, the effect of the net charge of selected charged amino acid residues on the excited electronic states was investigated by setting the partial charges of the side-chain atoms of these residues to 0 and repeating the TDDFT/MM calculations. The procedure adopted for testing the effect of explicit solvent versus implicit solvent on the excited electronic states of AdoCbl is explained in the Supporting materials and methods.

The process of generation of isosurface plots of the natural transition orbitals (NTOs) was automated using Python and Bash scripts, provided as part of the Supporting materials and methods. An isovalue of 0.07 $bohr^{-3/2}$ was used for most of the isosurface plots. In selected cases, an isovalue of 0.06 or 0.05 $bohr^{-3/2}$ was used for better understanding of the electronic transition. The images were visually analyzed to determine the basic nature of the electronic transitions.

Data availability

Images of natural transition orbitals corresponding to all the calculated excited states in all the time-dependent density functional theory calculations performed are available in a Google Drive repository (https://drive.google.com/drive/folders/1cnldtIQksAtH0Lsts6Hn-JueJK3hbSfw?usp=sharing).

RESULTS AND DISCUSSION

Four independent simulations for the WT CarH monomer ranging in length from 155 to 450 ns were carried out (Table S1). The root mean-square deviation (RMSD) of the backbone heavy atoms (Fig. S2, *A and B*) and the root mean-square fluctuation of the amino acid residues (Fig. S3, *A and B*) indicate that the protein is stable and that the N-terminal region exhibits larger flexibility compared with the

C-terminal region. As a test, one of the trajectories (traj. 2) was equilibrated for an additional 50 ns in the NPT ensemble with weak restraints on the N-terminal region residues, after which the restraints were released. The RMSD and root mean-square fluctuation plots were found to have no significant dependence on the application of restraints in the N-terminal region during equilibration. Because Glu141 is close to His142 in the crystal structure, His142 was treated as protonated in traj. 4 to investigate the effect of its protonation on the conformational sampling of Glu141. Although Glu141 remained hydrogen bonded to His142 throughout most of the trajectory (Fig. S4 *B*), unlike the other three trajectories with a neutral His142 (Fig. S4 *A*), Glu141 still showed significant hydrogen-bonding to adenine (Ade), like the other trajectories (Fig. S5, *A and B*). The hydrogen bond between Glu141 and Trp131 was not observed often (Fig. S7, *A and B*), which is similar to the behavior of traj. 1 but in contrast with traj. 2 and 3.

In all four WT trajectories, the Ado ligand is found to sample distinct orientations, as shown using representative snapshots from traj. 1 (Fig. 3). In the first ∼40 ns of this trajectory, the Ado ligand maintains an orientation similar to that in the crystal structure, remaining highly solvent exposed and relatively distant from Glu141 and Trp131. In the last ∼410 ns, it resides closer to Glu141 and Trp131. For all the trajectories, it is observed that in correlation with Ado rotation from a solvent-exposed to a π-stacked conformation (Figs. S8, *A and B* and S9, *A and B*), there is a decrease in the distance between the Glu141 side chain and the NH_2 group of Ade (Fig. S5, *A and B*). The formation of the π-stacked conformation between Trp131 and Ade is monitored by a dihedral angle involving Trp131 and Ade (Fig. S8, *A and B*) and the distance between the centers of mass of Trp131 and Ade (Fig. S9, *A and B*). The ribose moiety of Ado is also found to become more distant from Glu141 (Fig. S10, *A and B*) and Trp131 (Fig. S11, *A and B*), with the formation of the π-stacked conformation between Trp131 and Ade. The radial distribution function (29) plots in Figs. S12–S15 show that the solvent exposure of both the five- and six-membered rings of Ade decreases with formation of the π-stacking interaction between Ade and Trp131.

25 structures from the first 100 ns of traj. 1, at intervals of 4 ns, were selected for QM/MM calculations. Six structures

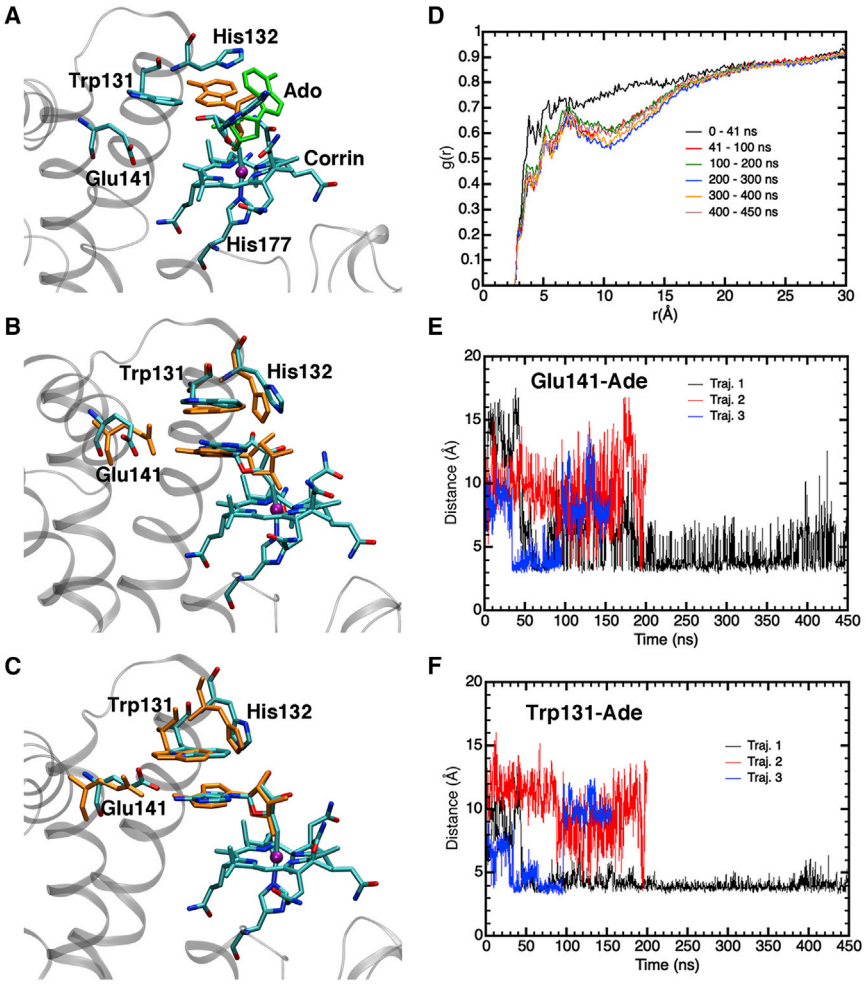

FIGURE 3 (*A–C*) Selected structures from an MD simulation (traj. 1) of the WT CarH monomer to demonstrate the conformational flexibility of the Ado ligand. (*A*) The structure at 28 ns (colored by atom type) is overlaid with that at 4 ns (*orange*) and 44 ns (*green*). Despite its flexibility, Ado remains relatively far from Trp131 and Glu141 and highly solvent exposed. (*B*) The structure at 52 ns (colored by atom type) is overlaid with that at 68 ns (*orange*). The Ade moiety is roughly coplanar with Trp131, and Glu141 forms no or one hydrogen bond to Ade. The solvent exposure of Ade is lower than in the first ~40 ns. (*C*) The structure at 64 ns (colored by atom type) is overlaid with that at 80 ns (*orange*). The difference from (*B*) is the formation of hydrogen bonds between Glu141 and both Trp131 and Ade. (*D*) Radial distribution function for water O atoms around the heavy atoms (excluding the N atom of the $-NH_2$ substituent) of the six-membered ring of the Ade moiety. (*E*) 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 Ade in three independent WT simulations with deprotonated His142. (*F*) Time evolution of the distance between the center of mass of the heavy atoms of the Trp131 side chain and the heavy atoms of the five- and six-membered rings of Ade in three independent WT simulations with deprotonated His142. To see this figure in color, go online.

with short Glu141/Trp131-sugar distances were also randomly selected from traj. 3. TDDFT calculations for the QM region (AdoCbl) were carried out in the presence of the point charges corresponding to the rest of the system. The entire protein and primary water box were retained in these calculations, with the water box centered at cobalamin and its longest edge aligned with the length of the protein. Table 1 shows the effect of the environment on the excitation wavelength of the lowest energy excited states involving CT from metal/corrin to Ade, CT from metal to the σ^* orbital corresponding to the Co-C(Ado) bond, and CT from metal to the corrin ring (Fig. 4). These three categories of excited states were investigated because the first can possibly lead to heterolytic cleavage of the Co-C bond, especially when accompanied by population of the Co-C σ^* orbital, and the latter two have been studied widely in previous investigations of homolytic cleavage of the Co-C bond in enzymes and in CarH (11,30–35). Table S2 shows the presence of the metal/corrin to Ade/σ^* transition in several structures selected from the MD simulation. This transition is absent in most calculations in only implicit solvent.

TABLE 1 Effect of environment on the excitation wavelength in nanometers and index of the lowest energy states involving metal/corrin to Ade, metal to σ^* and metal to corrin transitions

Time (ns)	Type of transition	Solvent	Protein
24[a]	metal/corrin to Ade	387.2, S_{20}	445.3, S_8
	metal to σ^*	481.3, S_7	491.0, S_4
	metal to corrin	568.6, S_2	540.7, S_2
44[a]	metal/corrin to Ade	399.0, S_{20}	519.8, S_4
	metal to σ^*	493.9, S_6	488.0, S_7
	metal to corrin	621.8, S_3	600.0, S_2
52[a]	metal/corrin to Ade	423.4, S_{14}	579.6, S_3
	metal to σ^*	475.0, S_9	486.5, S_8
	metal to corrin	619.3, S_2	595.2, S_2
96[a]	metal/corrin to Ade	473.5, S_{10}	596.9, S_4
	metal to σ^*	601.2, S_2	601.9, S_3
	metal to corrin	593.6, S_3	572.7, S_6
108[b]	metal/corrin to Ade	453.8, S_{14}	476.9, S_{10}
	metal to σ^*	511.2, S_8	516.4, S_6
	metal to corrin	624.2, S_2	604.8, S_2
115[b]	metal/corrin to Ade	439.1, S_{14}	627.0, S_2
	metal to σ^*	544.7, S_4	545.0, S_5
	metal to corrin	590.8, S_2	699.3, S_1
125[b]	metal/corrin to Ade	483.2, S_{11}	493.2, S_8
	metal to σ^*	559.4, S_5	559.9, S_4
	metal to corrin	638.9, S_3	640.5, S_3
137[b]	metal/corrin to Ade	468.0, S_{14}	661.6, S_3
	metal to σ^*	703.8, S_2	695.3, S_2
	metal to corrin	612.7, S_3	594.7, S_4
146[b]	metal/corrin to Ade	458.6, S_{11}	565.3, S_5
	metal to σ^*	558.5, S_4	567.3, S_4
	metal to corrin	598.6, S_2	681.9, S_3
154[b]	metal/corrin to Ade	399.7, S_{23}	475.9, S_{10}
	metal to σ^*	562.1, S_4	559.1, S_3
	metal to corrin	712.1, S_1	668.1, S_1

The Def2SVP basis set was used for all calculations. S_n indicates the n^{th} excited singlet state.

[a]Structures from traj. 1 (Detailed data for the other structures are provided in the Supporting materials and methods).

[b]Structures from traj. 3.

For several structures selected from the MD simulations (Table 1), the replacement of the explicitly solvated protein environment with implicit solvent (water in this case) leads to a significant blue shift in the excitation wavelength of the lowest lying excited state involving CT to Ade (see Table S2 for more details on other structures). In contrast, this change of environment has a small effect on the excitation wavelength of the lowest lying excited states involving CT from metal to the Co-C σ^* orbital and CT from metal to the corrin ring. For the latter, a small red shift is observed upon changing the environment of the chromophore from explicitly solvated protein to implicit solvent in most of the structures. Tests with AdoCbl solvated in implicit versus explicit water show that the effect of treating the solvent implicitly versus explicitly is minimal (Table S3). Changing the basis set from Def2SVP to the larger Def2TZVP basis set has significant effect on the excitation wavelength only if the CT to Ade is accompanied by population of the σ^* orbital corresponding to the Co-C(Ado) bond (Table S4).

The observed red shift for the lowest lying excited state involving CT to Ade upon changing the environment from implicit solvent to explicitly solvated protein was reproducible using a hybrid functional, B3LYP, and a range-separated functional, CAM-B3LYP (Table 2) (36). The magnitude of red shift, however, decreases in the order BP86, B3LYP, CAM-B3LYP. Previous studies have shown that for cobalamins, the BP86 functional best describes the Co-C(upper axial ligand) dissociation energy, and structural and spectral features (25–27). This was explained based on artificial widening of the energy gap between the occupied and virtual orbitals for cobalamins with range-separated functionals such as CAM-B3LYP. Because polarizability is related to the gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) gap (37,38), the large red shifts predicted by BP86 are likely more accurate than the smaller shifts predicted by the B3LYP and CAM-B3LYP functionals.

Whereas the absolute excitation wavelengths of states involving long range CT, like those involving CT from metal/corrin to the Ade ring, may not be well described by a generalized gradient approximation functional, such as BP86 (39), the shift in the excitation wavelength upon change of environment is expected to be more reliable, especially because BP86 describes the HOMO-LUMO gap for cobalamins well. Table 2 shows that the red shift in the excitation wavelength of the lowest lying excited state involving CT to Ade occurs not only upon changing the environment from implicit water to explicitly solvated protein, but also upon changing the environment from gas phase to implicit water.

Additional QM/MM calculations with side-chain charges of selected amino acids zeroed out provided insight into the

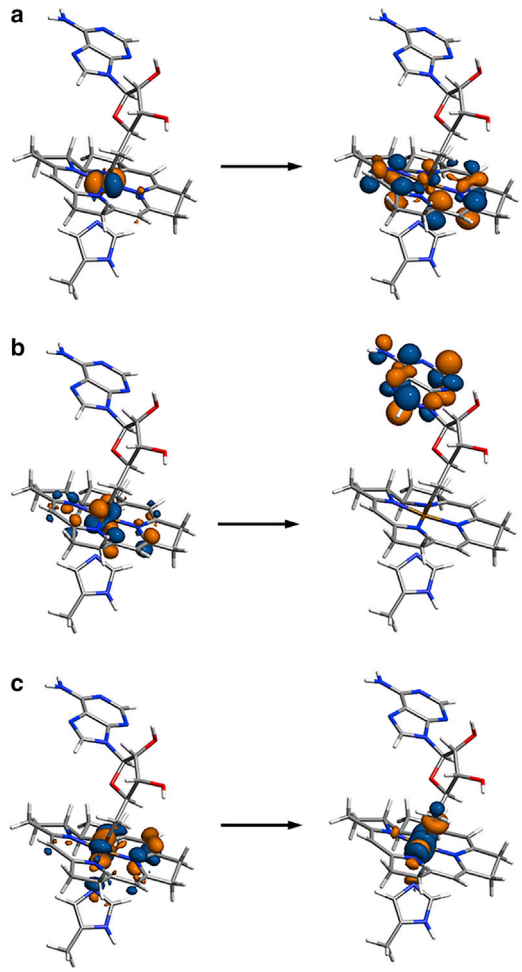

FIGURE 4 Representative natural transition orbitals (NTOs) depicting (*a*) CT from metal to the corrin ring, (*b*) CT from metal/corrin to Ade, (*c*) CT from metal to the σ^* orbital corresponding to the Co-C(Ado) bond. The Co atom is located at the center of the tetrapyrrole ring, the upper axial ligand of Co is the Ado group, and the lower axial ligand of Co is the His177 side chain. An isovalue of 0.05 $bohr^{-3/2}$ was used for these plots. The orange and blue colors of the isosurfaces indicate positive and negative signs of the wavefunction, respectively. To see this figure in color, go online.

aforementioned red shift upon changing the environment from implicit water to explicitly solvated protein. The structures at 32 and 72 ns of traj. 1 were selected because they showed the largest number of transitions involving CT to Ade in the first 60 excited states. Two other randomly selected structures (at 24 and 96 ns) were also investigated. The residues whose side-chain charges were zeroed out one at a time were Arg, Lys, Glu, and Asp residues with any atom within 8 Å of a heavy atom in the corrin central ring or of a heavy atom in the Ado ligand. Tables S5–S8 show that Glu175 is responsible for a large red shift in the excitation wavelength of the lowest lying excited state involving CT to Ade. An examination of the MD trajectory and the crystal structure (PDB: 5C8E) (5) shows that Glu175 is hydrogen bonded to the lower axial ligand of Co, His177 (see Fig. S20). The presence of this negatively charged residue close to the lower axial ligand is expected to promote CT to the upper axial ligand, as is observed in our calculations. The role of Glu175 in modulating the excited electronic states can be experimentally investigated by mutating it to a Gln residue that conserves the hydrogen-bonding interaction with His177 but removes the negative charge, or an Asp residue that possibly removes the hydrogen-bonding interaction but retains the negative charge.

Two independent simulations each for the E175Q and E175D mutant were carried out, and the detailed analysis in Figs. S2–S11, *C and D* and Figs. S16–S19 show similar trends to those in the WT trajectories, except greater sampling of longer distances between Glu141/Trp131 and the ribose O2′ atom in one of the E175Q trajectories. The other E175Q trajectory samples larger RMSD-values, primarily due to greater flexibility of the N-terminal region. Fig. S21 shows that, unlike Glu175, which forms a stable hydrogen bond to the His177 side chain in all independent trajectories for the WT protein, Gln175 and Asp175 do not form stable hydrogen bonds with His177. The rotation of Gln175 and Asp175 away from His177, as well as the rotation of the N_δ atom of His177 away from these residues is observed. Some test QM/MM calculations on structures selected from the E175Q and E175D trajectories show that a very large number of calculations would be required to make a meaningful distinction between the WT and mutant spectra because of the large number of variables that can be different between snapshots selected from the WT and mutant trajectories. The loss of a stable hydrogen bond between residue 175 and His177 in the E175Q and E175D mutants, as observed in MD simulations, besides the removal of a negative charge in the E175Q mutant, provides further motivation for experimental investigations of the E175Q and E175D mutants to understand the role of Glu175 in modulating cobalamin photochemistry in CarH.

Examination of the QM/MM data with the side-chain charges of selected residues zeroed out shows that the small blue shift in the excitation wavelength of the CT from metal to the corrin ring upon changing the environment of the chromophore from implicit solvent to explicitly solvated protein (Table 1) has small contributions from the side-chain charges of residues Arg125, Arg133, and Glu175 (see Supporting materials and methods, Data S1, Excel sheet and NTO isosurface plots). Arg125 and Arg133 are located on the Ado side of the corrin ring. Whereas it is not straightforward to rationalize the effect of these residues on the stability of the excited states with CT from metal to the corrin ring, they are probably responsible for a shift of the α-band to higher energy in the protein environment (10). In the protein environment, the red shift in the wavelength of the excitations involving CT to Ade possibly has little effect on the spectrum because of the low oscillator strength

TABLE 2 Effect of environment and DFT functional on the excitation wavelength in nanometers of the lowest energy state involving CT from metal/corrin to Ade

Time	BP86			B3LYP			CAM-B3LYP		
(ns)									
	Gas phase	Solvent[a]	Protein[b]	Gas phase	Solvent[a]	Protein[b]	Gas phase	Solvent[a]	Protein[b]
24	NA	387.2	445.3 (58.1)	247.0	294.4 (47.4)	319.7 (25.3)	NA	214.3	226.3 (12.0)
44	NA	399.0	519.8 (102.8)	NA	297.8	357.0 (59.2)	NA	218.1	248.4 (30.3)
52	330.1	423.4 (93.3)	579.6 (156.2)	257.2	313.6 (56.4)	385.8 (72.2)	NA	223.9	258.0 (34.1)
96	363.2	473.5 (110.3)	596.9 (123.3)	261.1	316.0 (54.9)	447.3 (131.3)	NA	220.4	277.5 (57.1)

The Def2SVP basis set was used for all calculations. Structures were taken from traj. 1.
NA, not available indicates that the transition could not be located in the first 60 excited states.
[a]Numbers in parentheses indicate the change in excitation wavelength on changing the environment from gas phase to solvent.
[b]Numbers in parentheses indicate the change in excitation wavelength on changing the environment from solvent to protein.

associated with these excitations (Supporting materials and methods, Data S1, NTO isosurface plots).

In addition to Glu175, Trp131, which is often observed to be roughly π-stacked with the Ade ring of AdoCbl, is expected to further stabilize the excited states involving CT to Ade. Calculations with a larger QM region involving Trp131 were not carried out because of concerns with the description of CT to Trp131 by the BP86 functional. However, π-stacking between Ade and Trp131 will most likely lead to stabilization of CT to Ade. An examination of the crystal structures of several AdoCbl-based enzymes, ethanolamine ammonia-lyase (PDB: 3ABS) (40), diol dehydratase (PDB: 5YRV) (41), methylmalonyl-CoA mutase (PDB: 4REQ) (42), and glutamate mutase (PDB: 1I9C) (43) shows lack of a Trp residue close to the cobalamin. This indicates that CarH, in which Trp131 is a highly conserved residue (10), likely uses a special π-stacking mechanism to promote CT to Ade. This can be probed experimentally via mutation of Trp131 to a residue that cannot π-stack with Ade.

Overall, our calculations reveal that the CarH protein environment is associated with three key features that cause extra stabilization of excited states involving CT from metal/corrin to Ade. This extra stabilization of excited states involving CT to Ade in CarH is most likely related to the experimentally observed excited-state charge separation and possibly promotes subsequent heterolytic cleavage of the Co-Ado bond, especially when it is accompanied by population of the Co-C σ^* orbital. Neither an aqueous solution environment nor the AdoCbl-based enzymes referred to above provide a combination of all these features. These three features are as follows: 1) solvent exposure of the Ado ligand, 2) hydrogen bond between a Glu and the lower axial His ligand of Co, and 3) π-stacking between the Ade ring and a Trp side chain. An aqueous solvent environment lacks features 2 and 3, methylmalonyl-CoA mutase (PDB: 4REQ) and glutamate mutase (PDB: 1I9C) lack features 1 and 3, and ethanolamine ammonia-lyase (PDB: 3ABS) and diol dehydratase (PDB: 5YRV) lack all three features. The cobalamin-dependent enzymes have limited conformational flexibility. Their crystal structures do not indicate significant conformational changes upon substrate binding (40,43–46), thus justifying the above comparison between CarH and the enzymes based on the crystal structures of the latter. A comprehensive investigation, however, remains the subject of future studies.

In the crystal structure of the DNA-bound tetramer (Fig. 1) (5), 1) the small dimer interface in the vicinity of AdoCbl, 2) the lack of significant interactions between the monomers at the dimer interface in the vicinity of AdoCbl, and 3) the large solvent exposure of Ado, Trp131 and nearby residues indicate that Ado flexibility will be similar in the monomer and tetramer. Regardless, our observed stabilization of the Co/corrin to Ade CT states in CarH is independent of the protein conformation and Ado orientation. The three factors which promote this stabilization, =solvent exposure of Ado, π-stacking interactions between Ado and Trp131, and hydrogen bonding between His177 and Glu175, are independent of the protein oligomerization state or the rotation of Ado toward Glu141. Parallel displaced and T-shaped π-stacking interactions between Ado and Trp131 are formed in our simulations even without the rotation of the Ado ligand toward Glu141. Hence, this study, based on simulations of the CarH monomer, provides valuable insight into the modulation of AdoCbl excited states by the CarH protein environment. Simulations of the DNA-bound tetramer are a subject of future investigations.

A recent computational study of CarH (11) based on energy-minimized structures of the unsolvated protein monomer assumed the photochemistry to initiate from the Franck-Condon region of the S_1 state, this region being associated with metal to corrin CT. Crossing of the energy barrier between the Franck-Condon region and the region of the S_1 potential energy surface associated with metal to σ^* transition was assumed to be related with homolytic cleavage of the Co-C bond. It should, however, be noted that the photochemical activity of CarH is experimentally observed for a broad range of excitation wavelengths, and the population decay from the excited states does not necessarily lead to population of the Franck-Condon region of the S_1 state. In the absence of experimental evidence of nonreactivity on the higher excited states of CarH, occurrence of heterolytic cleavage of the Co-C bond correlated with

stabilization of the excited states involving CT to Ade in CarH cannot be ruled out. Such heterolytic cleavage may be accompanied by abstraction of the H4′ atom of Ado as hydride by Co (Fig. 2) ; this atom is found to stay relatively close to Co throughout our MD simulation, despite the overall conformational flexibility of Ado (Fig. S22). The nonadiabatic dynamics associated with cobalamin photochemistry in CarH needs to be investigated using a combination of computational and experimental methods. Experimental studies of the wavelength dependence of AdoCbl photolysis in CarH can provide valuable insights in this regard.

CONCLUSIONS

In conclusion, we have investigated the effect of the CarH protein environment on the excited electronic states of AdoCbl, with adequate sampling of the protein conformations using MD simulation, QM/MM calculations with appropriate description of the protein and aqueous environment, and consideration of the role of CT states involving the Ade group in the photochemistry. Our results indicate a progressive red shift in the excitation wavelength and hence stabilization of the electronic states involving CT from metal/corrin to Ade on changing the environment from gas phase to water to solvated protein. The excited states with metal to σ^* transition, which has been widely associated with homolytic cleavage of the Co-C bond (11,30–35), do not exhibit such stabilization in CarH. The solvent exposure of AdoCbl in CarH, the π-stacking interaction between Trp131 and the Ade moiety, and the hydrogen-bonding interaction between Glu175 and the lower axial ligand of Co contribute to the stabilization of the states involving CT to Ade. The combination of these three factors may promote heterolytic cleavage of the cobalt-Ado bond as the protein evolves through excited electronic states, especially those with CT to Ade accompanied by population of the Co-C σ^* orbital. A combination of all these factors is absent in an aqueous solvent environment and in common AdoCbl-based enzymes, which exhibit homolytic cleavage of the Co-C bond upon excitation by light. The role of Trp131 and Glu141 can be experimentally tested through site-directed mutagenesis studies. The involvement of higher excited electronic states in photochemistry can be investigated through studies of the wavelength dependence of AdoCbl photolysis in CarH. Overall, the insights obtained from this study contribute to the fundamental knowledge that will serve the development of optogenetic tools based on the new class of B_{12}-dependent photoreceptors. Future studies using higher level electronic structure methods will allow greater insights into the influence of the CarH protein environment on the excited states of AdoCbl and an understanding of the evolution of the system on the excited states that determines the observed photochemistry.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.07.020.

AUTHOR CONTRIBUTIONS

P.G. designed the research. All authors carried out calculations, analyzed the data, and wrote the article.

ACKNOWLEDGMENTS

We thank Dr. Rafael da Silva and Dr. Amanda Morris for helpful discussions. All figures were made using version 1.9.4 of the Visual Molecular Dynamics program (15), Xmgrace (47), version 2.8 of Gimp (48), and ImageMagick (49).

This study was supported by funding from the Chemistry Department at the State University of New York at Binghamton and the Research Foundation of the State University of New York.

SUPPORTING CITATIONS

Reference (50) appears in the supporting material.

REFERENCES

1. Pérez-Marín, M. C., S. Padmanabhan, …, M. Elías-Arnanz. 2008. Vitamin B12 partners the CarH repressor to downregulate a photoinducible promoter in Myxococcus xanthus. *Mol. Microbiol.* 67:804–819.
2. Ortiz-Guerrero, J. M., M. C. Polanco, …, M. Elías-Arnanz. 2011. Light-dependent gene regulation by a coenzyme B12-based photoreceptor. *Proc. Natl. Acad. Sci. USA.* 108:7565–7570.
3. Cheng, Z., K. Li, …, C. E. Bauer. 2014. Vitamin B12 regulates photosystem gene expression via the CrtJ antirepressor AerR in Rhodobacter capsulatus. *Mol. Microbiol.* 91:649–664.
4. Vermeulen, A. J., and C. E. Bauer. 2015. Members of the PpaA/AerR antirepressor family bind cobalamin. *J. Bacteriol.* 197:2694–2703.
5. Jost, M., J. Fernández-Zapata, …, C. L. Drennan. 2015. Structural basis for gene regulation by a B12-dependent photoreceptor. *Nature.* 526:536–541.
6. Kutta, R. J., S. J. O. Hardman, …, A. R. Jones. 2015. The photochemical mechanism of a B12-dependent photoreceptor protein. *Nat. Commun.* 6:7907.
7. Chemaly, S. M. 2016. New light on vitamin B12: the adenosylcobalamin-dependent photoreceptor protein CarH. *S. Afr. J. Sci.* 112:1–9.
8. Banerjee, R., and S. W. Ragsdale. 2003. The many faces of vitamin B12: catalysis by cobalamin-dependent enzymes. *Annu. Rev. Biochem.* 72:209–247.
9. Jones, A. R. 2017. The photochemistry and photobiology of vitamin B_{12}. *Photochem. Photobiol. Sci.* 16:820–834.
10. Padmanabhan, S., M. Jost, …, M. Elías-Arnanz. 2017. A new facet of vitamin B_{12}: gene regulation by cobalamin-based photoreceptors. *Annu. Rev. Biochem.* 86:485–514.
11. Toda, M. J., A. A. Mamun, …, P. M. Kozlowski. 2020. Why is CarH photolytically active in comparison to other B_{12}-dependent enzymes? *J. Photochem. Photobiol. B.* 209:111919.
12. Kainrath, S., M. Stadler, …, H. Janovjak. 2017. Green-light-induced inactivation of receptor signaling using cobalamin-binding domains. *Angew. Chem. Int.Engl.* 56:4608–4611.
13. García-Moreno, D., M. C. Polanco, …, M. Elías-Arnanz. 2009. A vitamin B12-based system for conditional expression reveals dksA to

be an essential gene in Myxococcus xanthus. *J. Bacteriol.* 191:3108–3119.

14. Phillips, J. C., D. J. Hardy, …, E. Tajkhorshid. 2020. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* 153:044130.
15. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28..
16. Darden, T., D. York, and L. Pedersen. 1993. Particle Mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
17. Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.
18. MacKerell, A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
19. Beglov, D., and B. Roux. 1994. Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* 100:9050–9063.
20. Pavlova, A., J. M. Parks, and J. C. Gumbart. 2018. Development of CHARMM-compatible force-field parameters for cobalamin and related cofactors from quantum mechanical calculations. *J. Chem. Theory Comput.* 14:784–798.
21. Melo, M. C. R., R. C. Bernardi, …, Z. Luthey-Schulten. 2018. NAMD goes quantum: an integrative suite for hybrid simulations. *Nat. Methods.* 15:351–354.
22. Frisch, M. J., G. W. Trucks, …, D. J. Fox. 2009. Gaussian09 Revision E.01. Gaussian Inc., Wallingford, CT.
23. Lin, H., and D. G. Truhlar. 2005. Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations. *J. Phys. Chem. A.* 109:3991–4004.
24. Burke, K., J. Werschnik, and E. K. U. Gross. 2005. Time-dependent density functional theory: past, present, and future. *J. Chem. Phys.* 123:62206.
25. Andruniów, T., M. Jaworska, …, P. M. Kozlowski. 2009. Time-dependent density functional theory study of cobalt corrinoids: electronically excited states of coenzyme B_{12}. *J. Chem. Phys.* 131:105105.
26. Kornobis, K., N. Kumar, …, P. M. Kozlowski. 2011. Electronically excited states of vitamin B12: benchmark calculations including time-dependent density functional theory and correlated ab initio methods. *J. Phys. Chem. A.* 115:1280–1292.
27. Toda, M. J., P. Lodowski, …, P. M. Kozlowski. 2019. Photolytic properties of the biologically active forms of vitamin B_{12}. *Coord. Chem. Rev.* 385:20–43.
28. Miertuš, S., E. Scrocco, and J. Tomasi. 1981. Electrostatic interaction of a solute with a continuum. A direct utilizaion of *ab initio* molecular potentials for the prevision of solvent effects. *Chem. Phys.* 55:117–129.
29. Levine, B. G., J. E. Stone, and A. Kohlmeyer. 2011. Fast analysis of molecular dynamics trajectories with graphics processing units-radial distribution function histogramming. *J. Comput. Phys.* 230:3556–3569.
30. Mamun, A. A., M. J. Toda, …, P. M. Kozlowski. 2018. Mechanism of light induced radical pair formation in coenzyme B_{12}-dependent ethanolamine ammonia-lyase. *ACS Catal.* 8:7164–7178.
31. Ghosh, A. P., A. A. Mamun, …, P. M. Kozlowski. 2018. Mechanism of the photo-induced activation of CoC bond in methylcobalamin-dependent methionine synthase. *J. Photochem. Photobiol. B.* 189:306–317.
32. Ghosh, A. P., A. A. Mamun, and P. M. Kozlowski. 2019. How does the mutation in the cap domain of methylcobalamin-dependent methionine synthase influence the photoactivation of the Co-C bond? *Phys. Chem. Chem. Phys.* 21:20628–20640.
33. Mamun, A. A., M. J. Toda, and P. M. Kozlowski. 2019. Can photolysis of the CoC bond in coenzyme B_{12}-dependent enzymes be used to mimic the native reaction? *J. Photochem. Photobiol. B.* 191:175–184.
34. Kumar, N., D. Bucher, and P. M. Kozlowski. 2019. Mechanistic implications of reductive Co-C bond cleavage in B_{12}-dependent methylmalonyl CoA mutase. *J. Phys. Chem. B.* 123:2210–2216.
35. Mamun, A. A., M. J. Toda, …, P. M. Kozlowski. 2019. Photolytic cleavage of Co-C bond in coenzyme B_{12}-dependent glutamate mutase. *J. Phys. Chem. B.* 123:2585–2598.
36. Mardirossian, N., and M. Head-Gordon. 2017. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Mol. Phys.* 115:2315–2372.
37. Pearson, R. G. 1986. Absolute electronegativity and hardness correlated with molecular orbital theory. *Proc. Natl. Acad. Sci. USA.* 83:8440–8441.
38. Janjua, M. R. S. A., A. Mahmood, …, S. Pan. 2014. Electronic absorption spectra and nonlinear optical properties of ruthenium acetylide complexes: a DFT study toward the designing of new high NLO response compounds. *Acta Chim. Slov.* 61:382–390.
39. Kümmel, S. 2017. Charge-transfer excitations: a challenge for time-dependent density functional theory that has been met. *Adv. Energy Mater.* 7:1700440.
40. Shibata, N., H. Tamagaki, …, T. Toraya. 2010. Crystal structures of ethanolamine ammonia-lyase complexed with coenzyme B12 analogs and substrates. *J. Biol. Chem.* 285:26484–26493.
41. Shibata, N., Y. Sueyoshi, …, T. Toraya. 2018. Direct participation of a peripheral side chain of a corrin ring in coenzyme B_{12} catalysis. *Angew. Chem. Int.Engl.* 57:7830–7835.
42. Mancia, F., and P. R. Evans. 1998. Conformational changes on substrate binding to methylmalonyl CoA mutase and new insights into the free radical mechanism. *Structure.* 6:711–720.
43. Gruber, K., R. Reitzer, and C. Kratky. 2001. Radical shuttling in a protein: ribose pseudorotation controls alkyl-radical transfer in the coenzyme B(12) dependent enzyme glutamate mutase This work was supported by the Österreichische Akademie der Wissenschaften (APART fellowship 614), the Österreichische Fonds zur Förderung der wissenschaftlichen Forschung (FWF-project 11599), and the European Commission (TMR project number ERB 4061 PL 95-0307). Crystallographic data were collected at the EMBL-beamline BW7B at DESY in Hamburg, Germany. We thank the beamline scientists for their assistance, and Ingrid Dreveny, Günter Gartler, Gerwald Jogl, and Oliver Sauer for their help during data collection. This research emerged from a collaboration with Prof. W. Buckel (Marburg) who supplied us with clones of the glutamate mutase proteins. *Angew. Chem. Int. Ed. Engl.* 40:3377–3380.
44. Yamanishi, M., M. Yunoki, …, T. Toraya. 2002. The crystal structure of coenzyme B12-dependent glycerol dehydratase in complex with cobalamin and propane-1,2-diol. *Eur. J. Biochem.* 269:4484–4494.
45. Gruber, K., and C. Kratky. 2002. Coenzyme B(12) dependent glutamate mutase. *Curr. Opin. Chem. Biol.* 6:598–603.
46. Toraya, T., S. Honda, and K. Mori. 2010. Coenzyme B12-dependent diol dehydratase is a potassium ion-requiring calcium metalloenzyme: evidence that the substrate-coordinated metal ion is calcium. *Biochemistry.* 49:7210–7217.
47. Stambulchik, E.. Grace. https://plasma-gate.weizmann.ac.il/Grace/.
48. The GIMP Development Team.. GIMP. https://www.gimp.org.
49. The ImageMagick Development Team.. ImageMagick. https://imagemagick.org.
50. Miertuš, S., E. Scrocco, and J. Tomasi. 1981. Electrostatic interaction of a solute with a continuum. A direct utilization of Ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* 55:117–129.

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

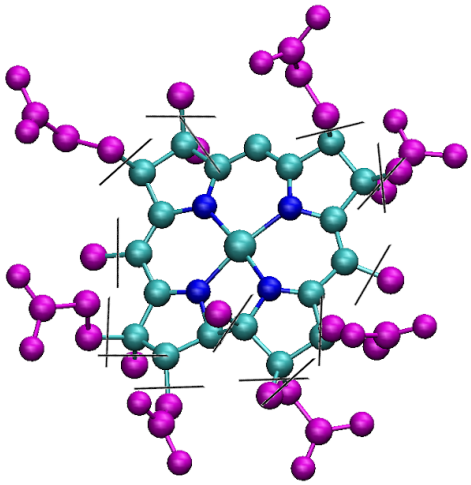

Figure S1: The black lines indicate the location of the QM/MM[3] boundary for the corrin ring. All atoms colored in purple, along with their covalently bonded H atoms were treated as point charges, while all other atoms, along with their covalently bonded H atoms were treated as QM.

Analysis of MD simulation data

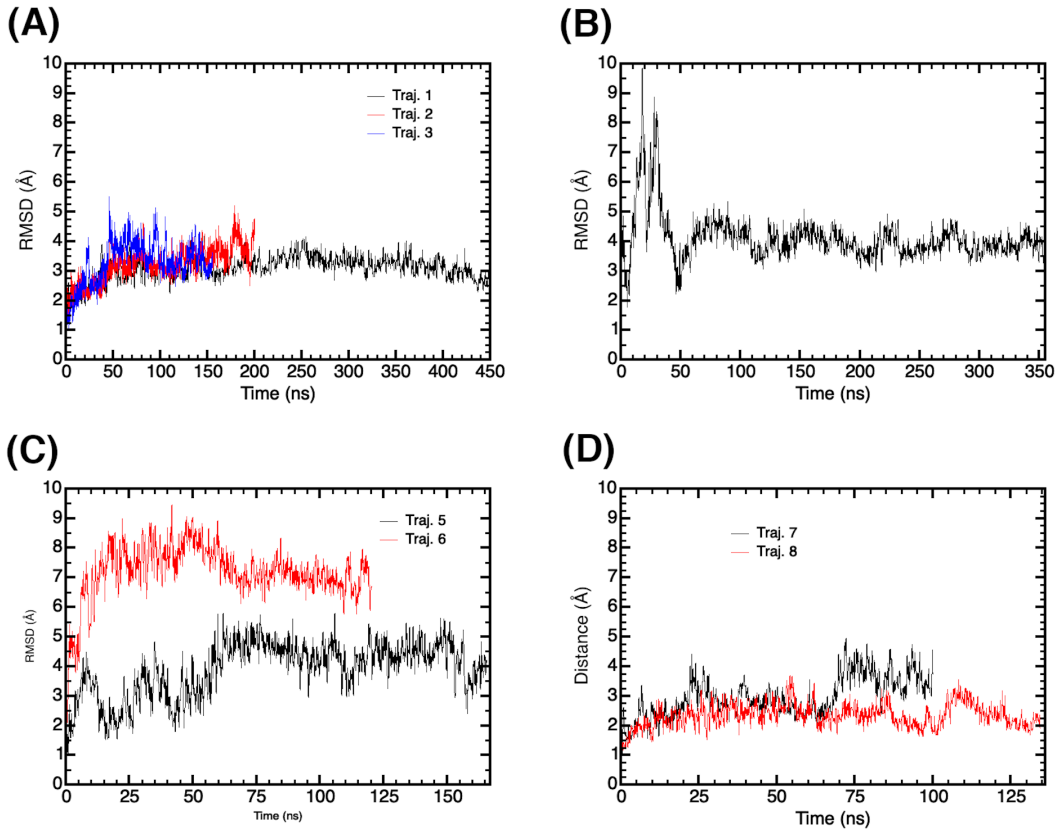

Figure S2: Root mean square deviation (RMSD) of the protein backbone heavy atoms with respect to the first frame in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations 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.

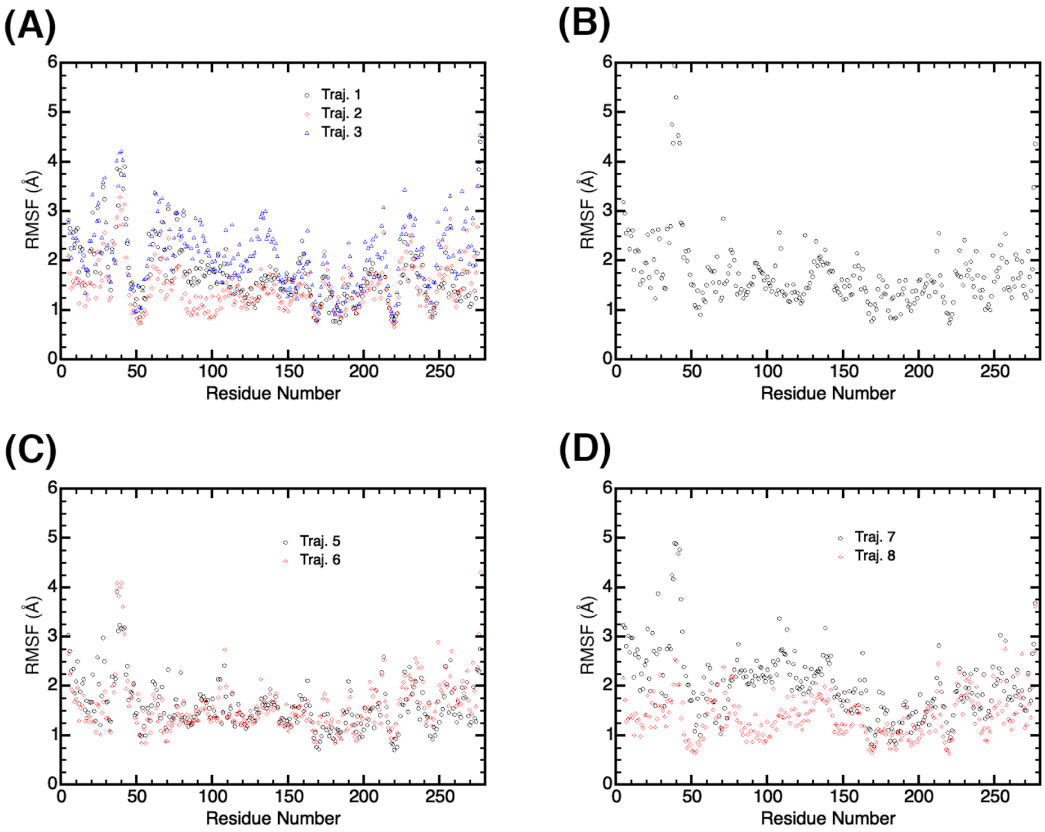

Figure S3: Root mean square fluctuation (RMSF) of the amino acid residues in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D mutant with traj. 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.

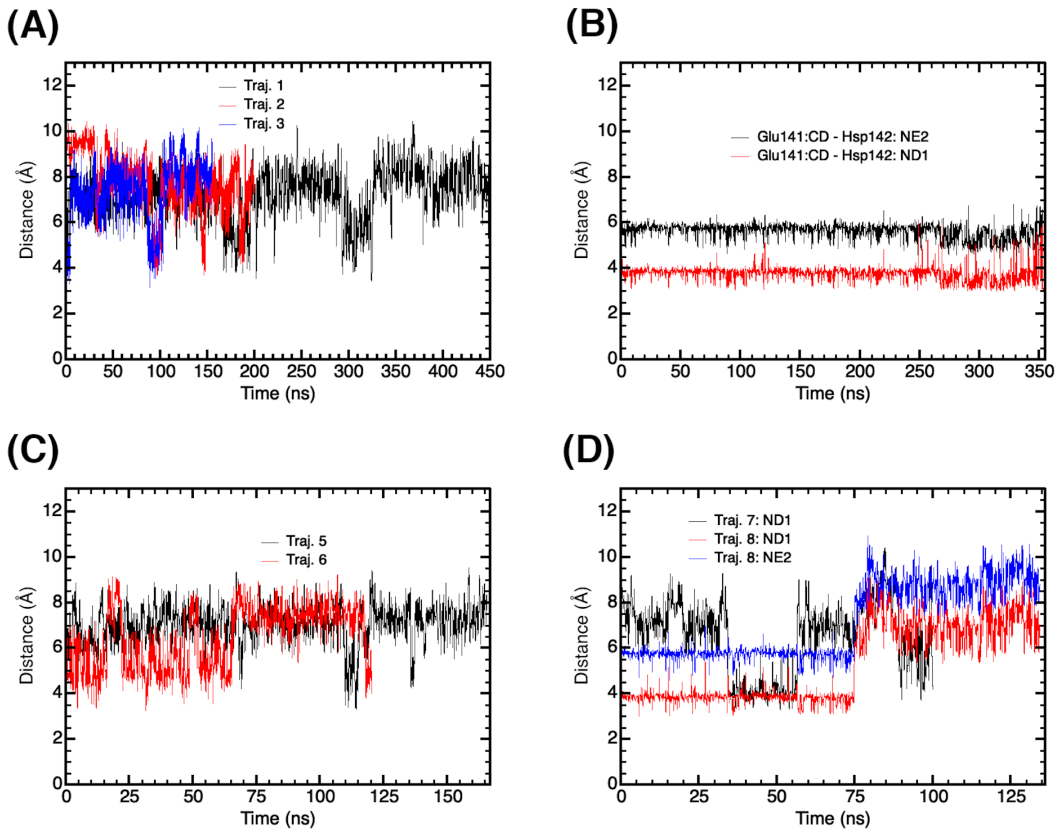

Figure S4: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the sidechain N atom(s) of His142 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) 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.

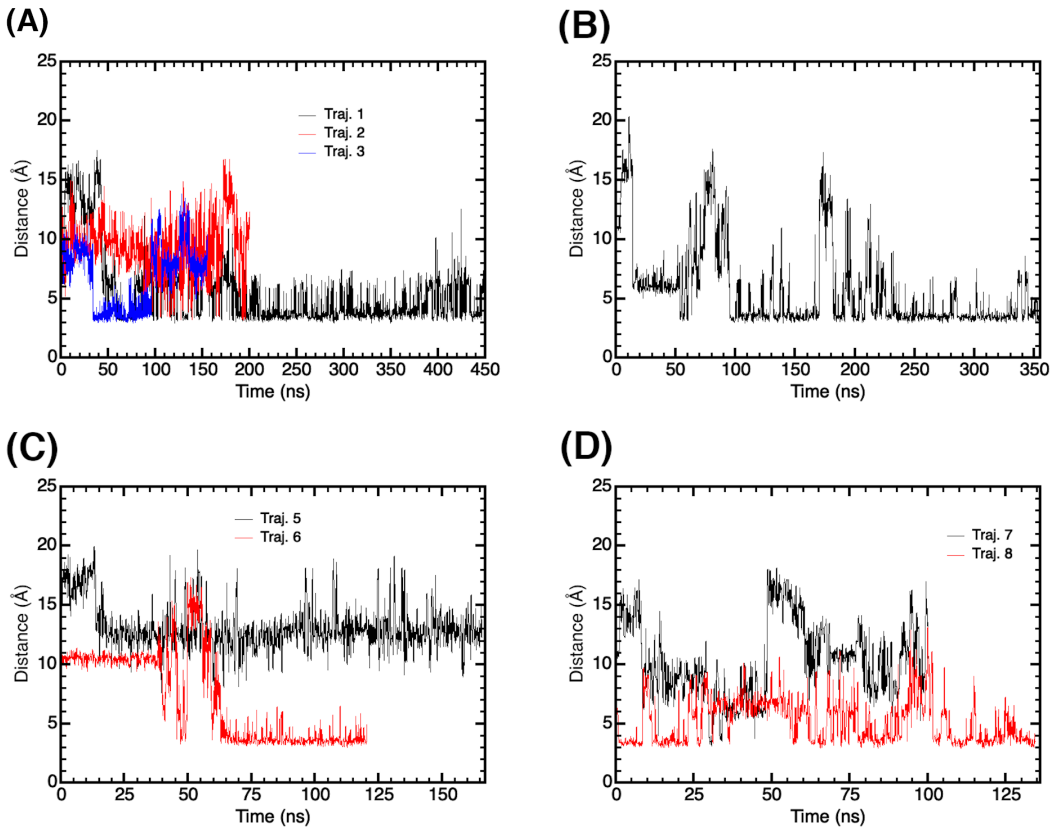

Figure S5: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the N6 atom (N atom of -NH_2 group, also see Fig. S6) of adenine in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations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

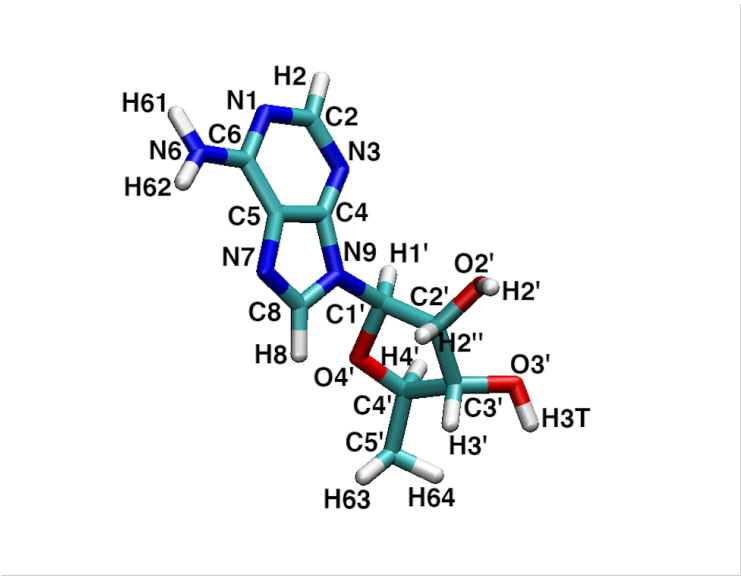

Figure S6: A depiction of the atom names in the adenosyl ligand, also referred to as 5AD.

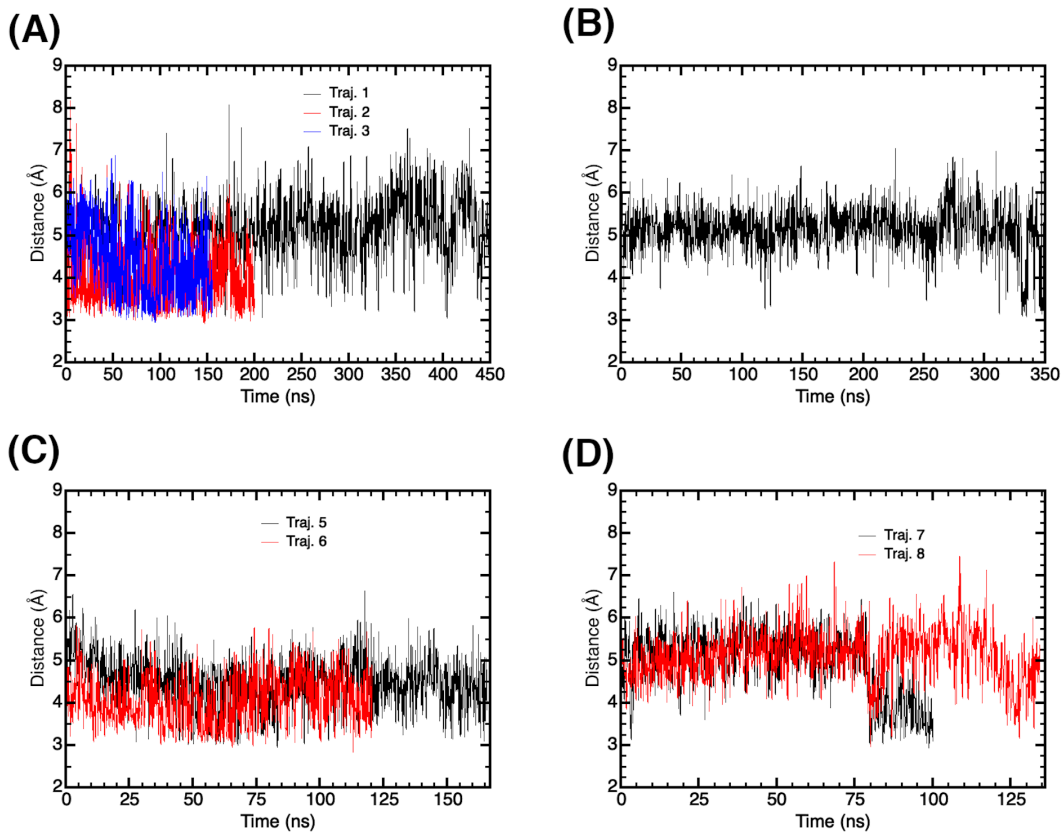

Figure S7: Time evolution of the distance between the C_δ atom (carboxylate C atom) of Glu141 and the sidechain N atom of Trp131 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two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

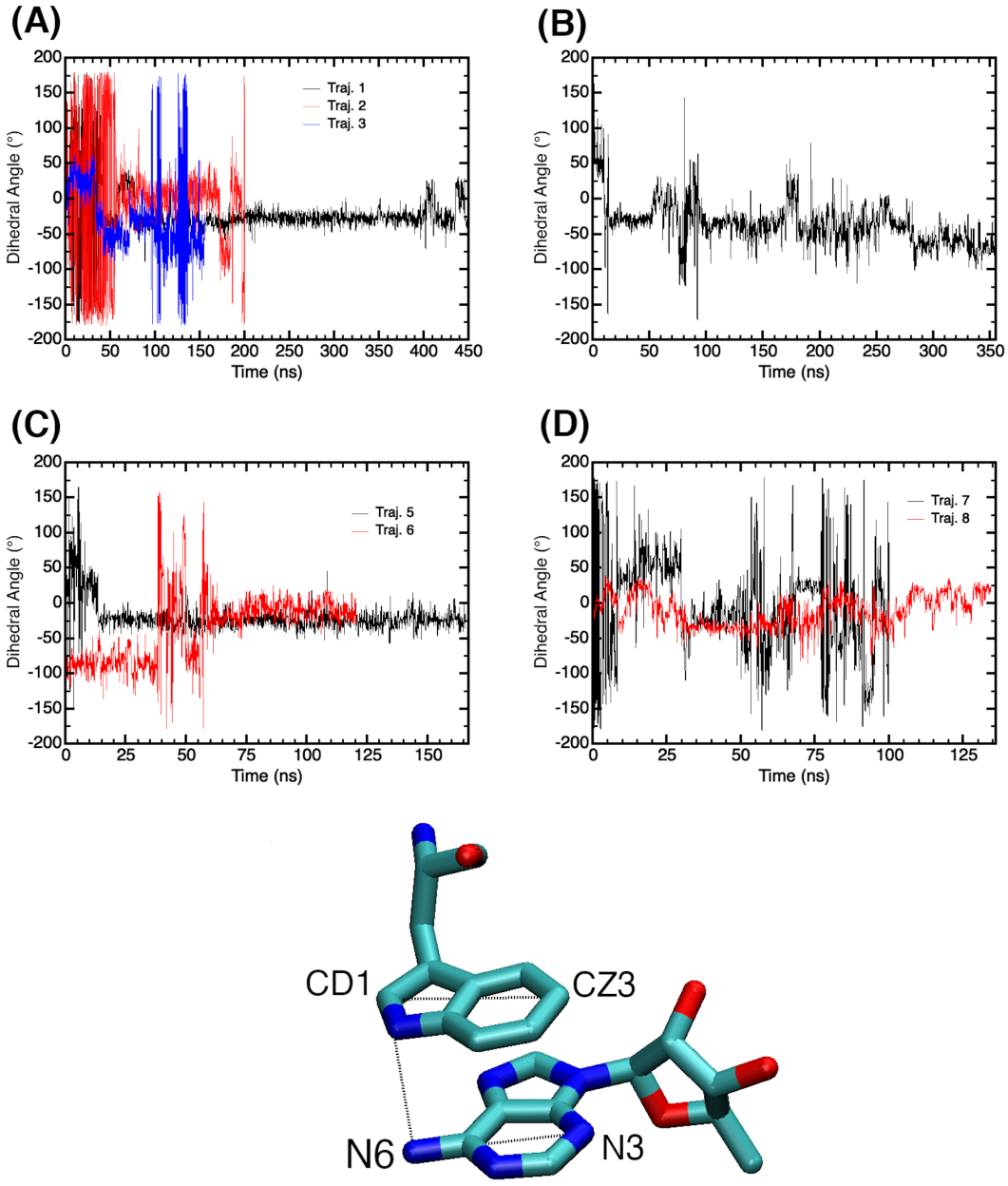

Figure S8: Time evolution of the CZ3-CD1-N6-N3 dihedral angle between Trp131 and adenosyl in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the E175D 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.

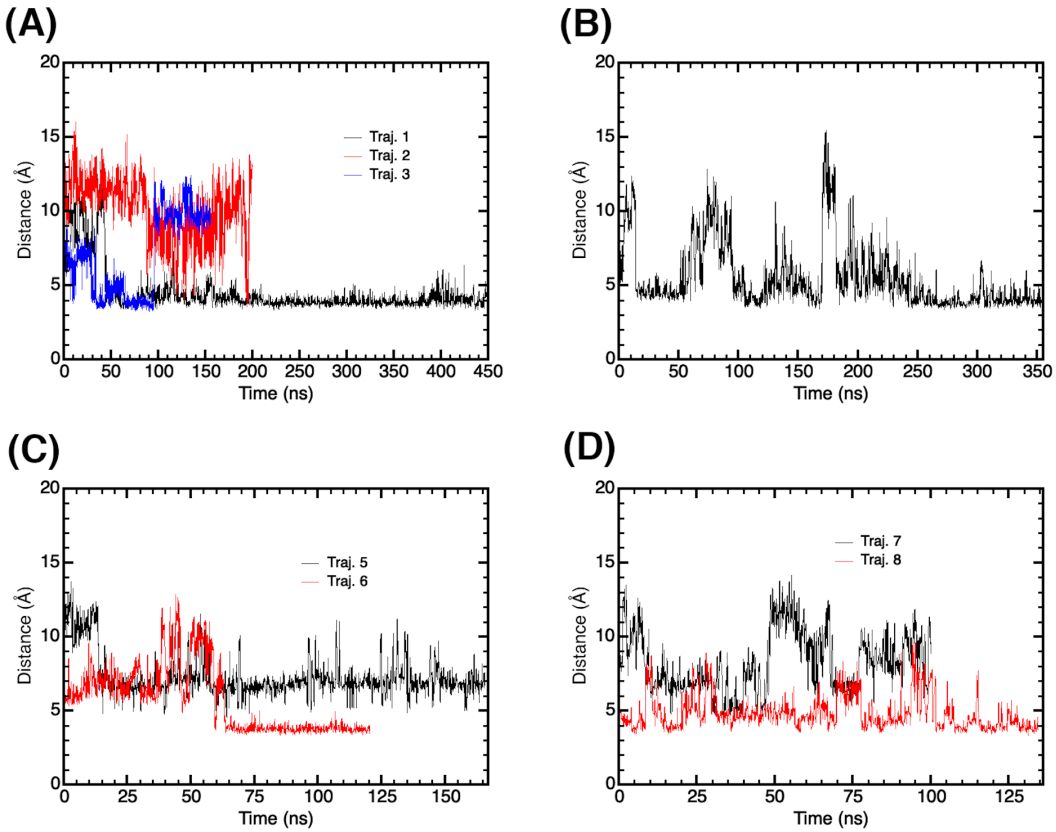

Figure S9: Time evolution of the distance between the center of mass of the heavy atoms of the Trp131 sidechain and the heavy atoms of the five- and six-membered rings of adenine in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent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

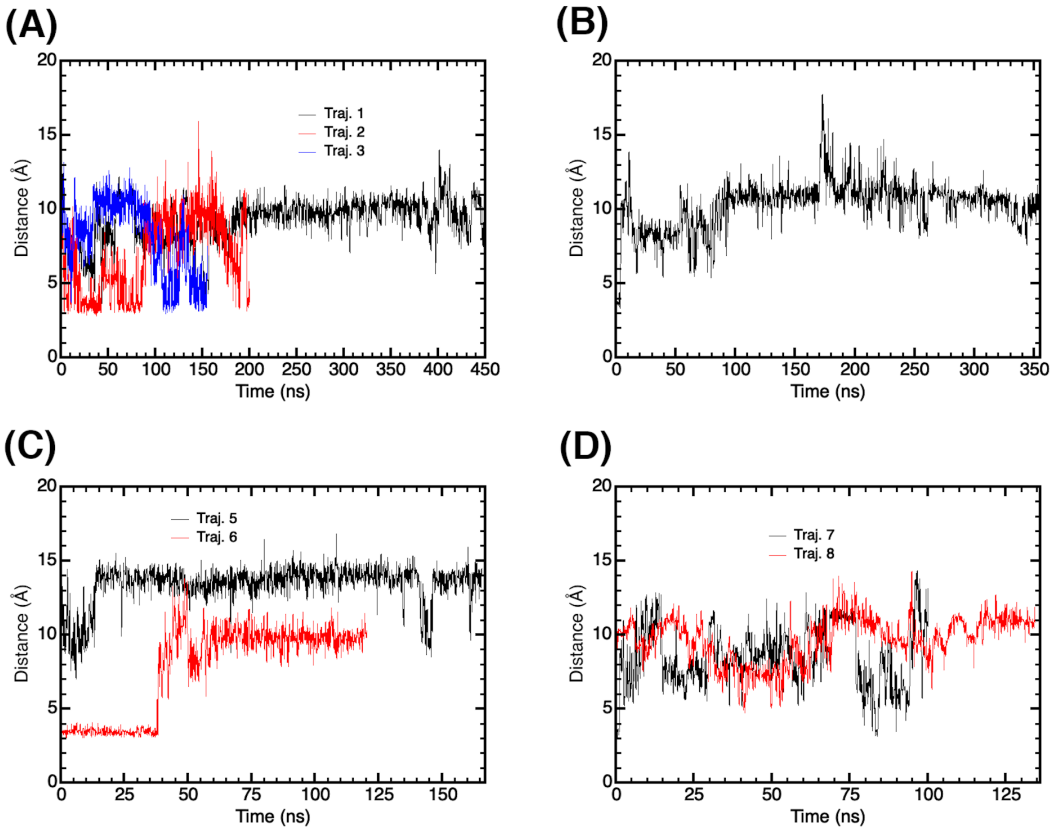

Figure S10: Time evolution of the distance between the ribose O2′ atom of adenosyl (see Fig. S6) and the carboxylate C atom (C_δ) of Glu141 for (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant,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

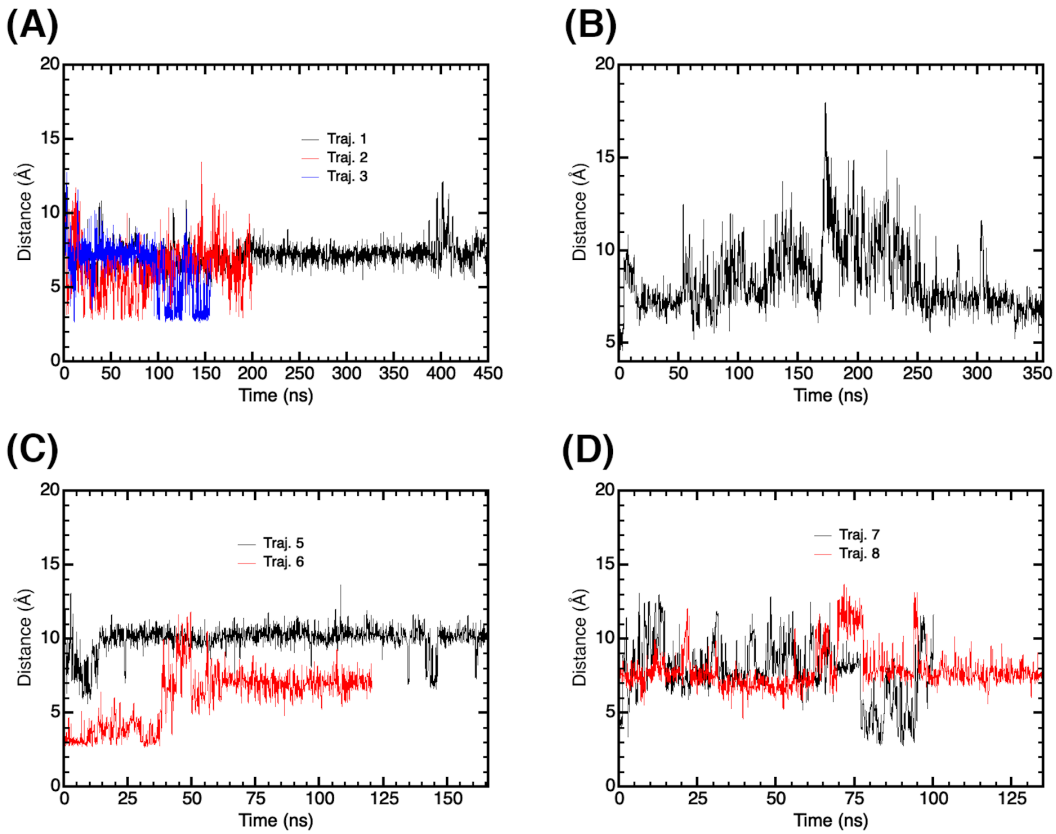

Figure S11: Time evolution of the distance between the ribose O2′ atom of adenosyl (see Fig. S6) and the side chain N atom of Trp131 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D)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

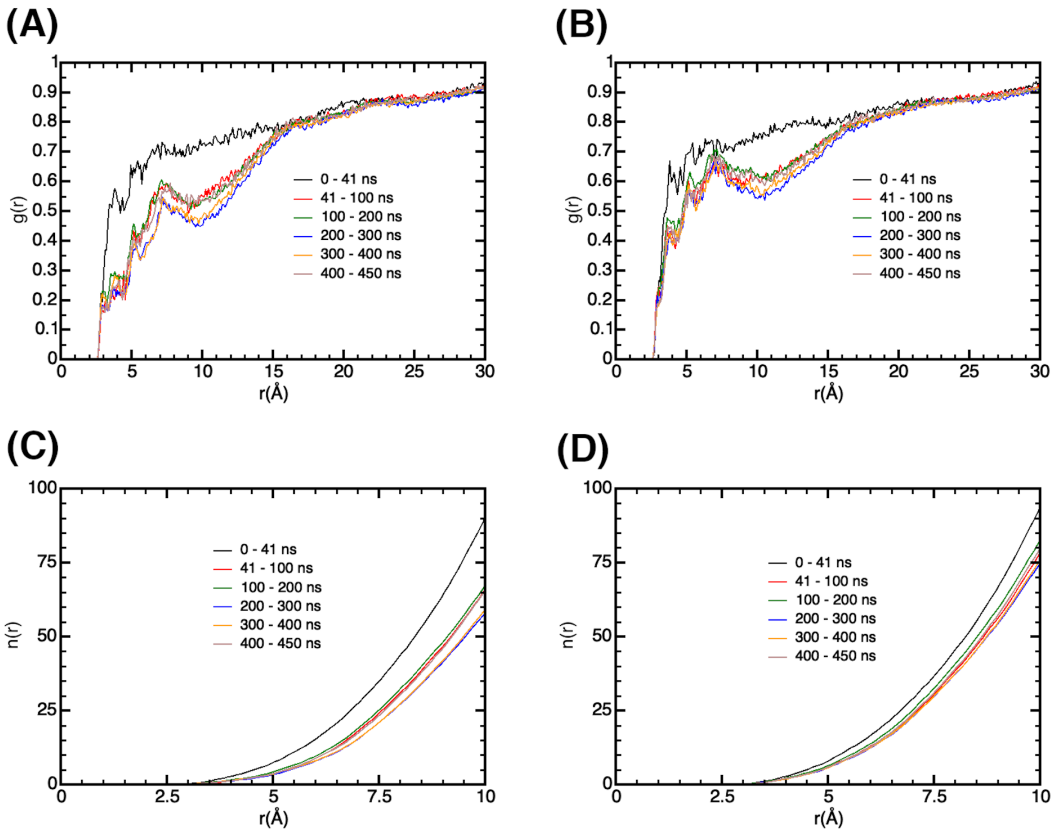

Figure S12: Radial distribution function (RDF) for water O atoms around the heavy atoms (excluding the N atom of the -NH_2 substituent) of the five-membered ring (A) and the six-membered ring (B) of the adenine moiety in traj. 1. The RDFs were examined for separate time intervals based on the time evolution 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]

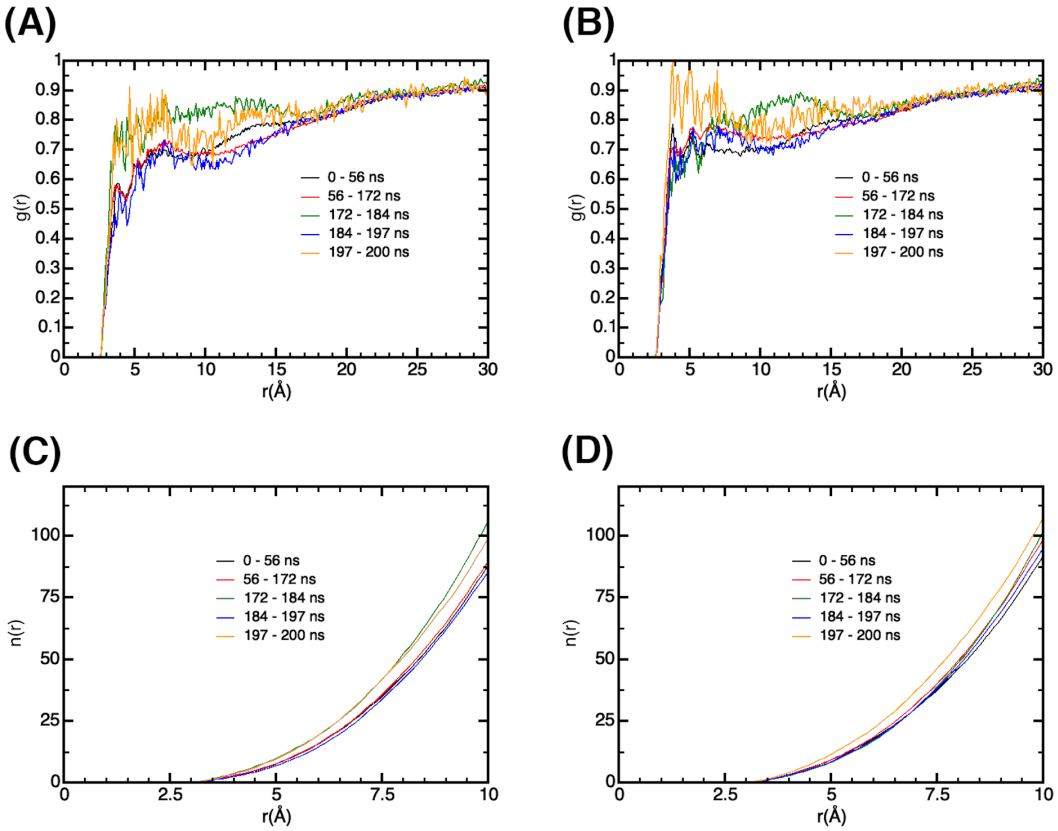

Figure S13: RDF for traj. 2 calculated using the same procedure as for Fig. S12.

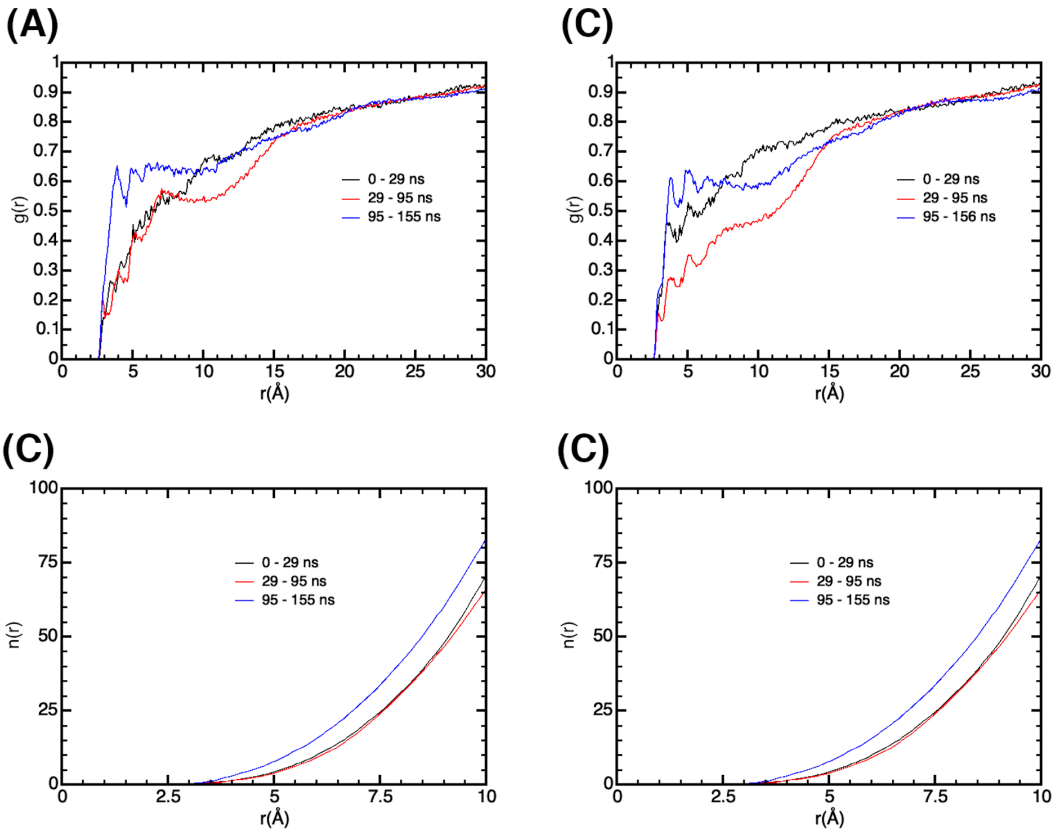

Figure S14: RDF for traj. 3 calculated using the same procedure as for Fig. S12.

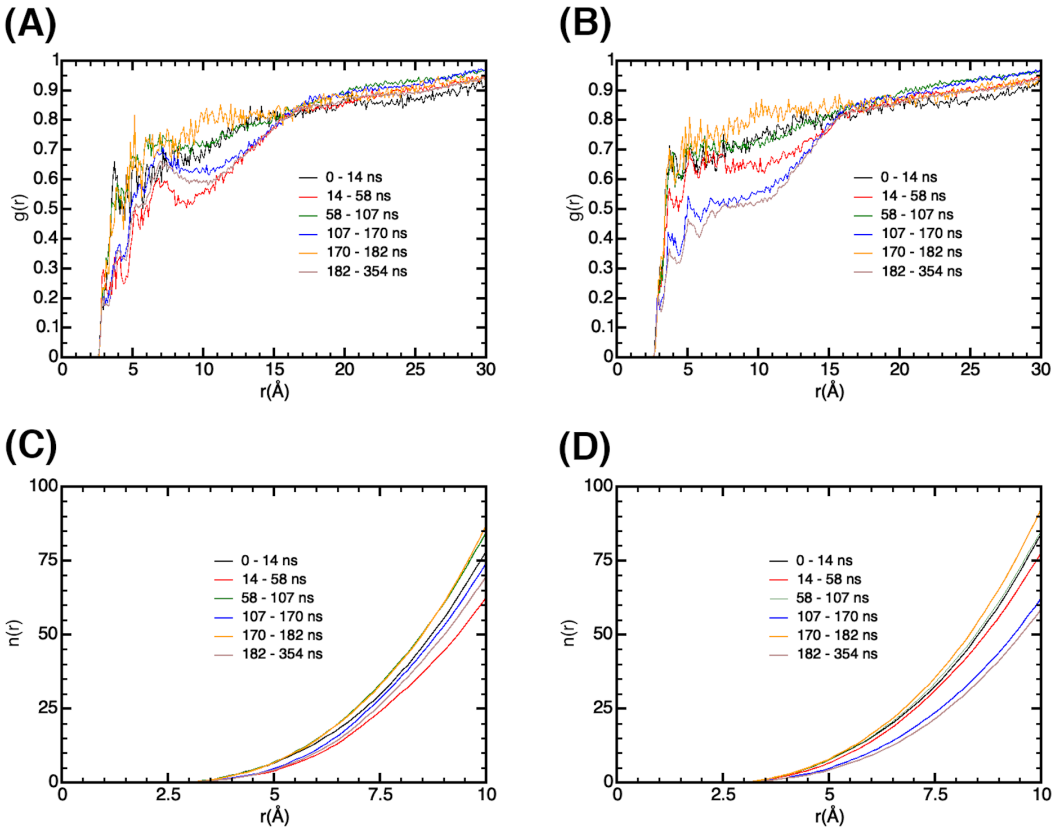

Figure S15: RDF for traj. 4 calculated using the same procedure as for Fig. S12.

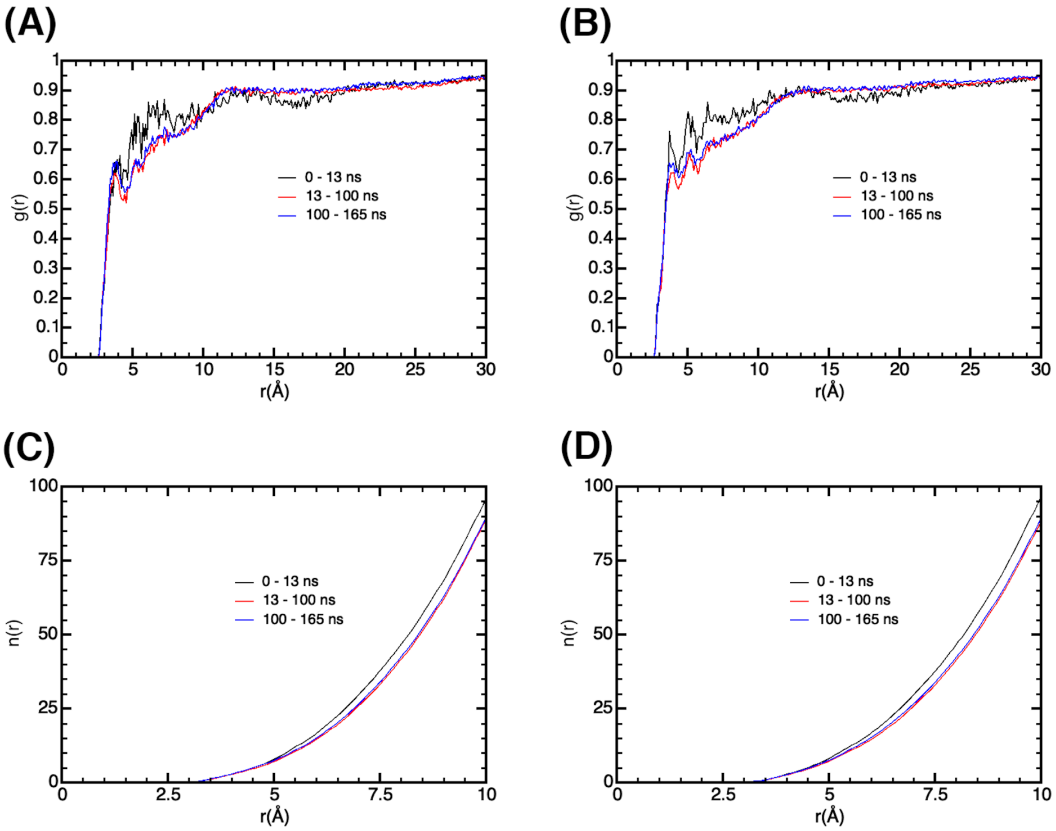

Figure S16: RDF for traj. 5 calculated using the same procedure as for Fig. S12.

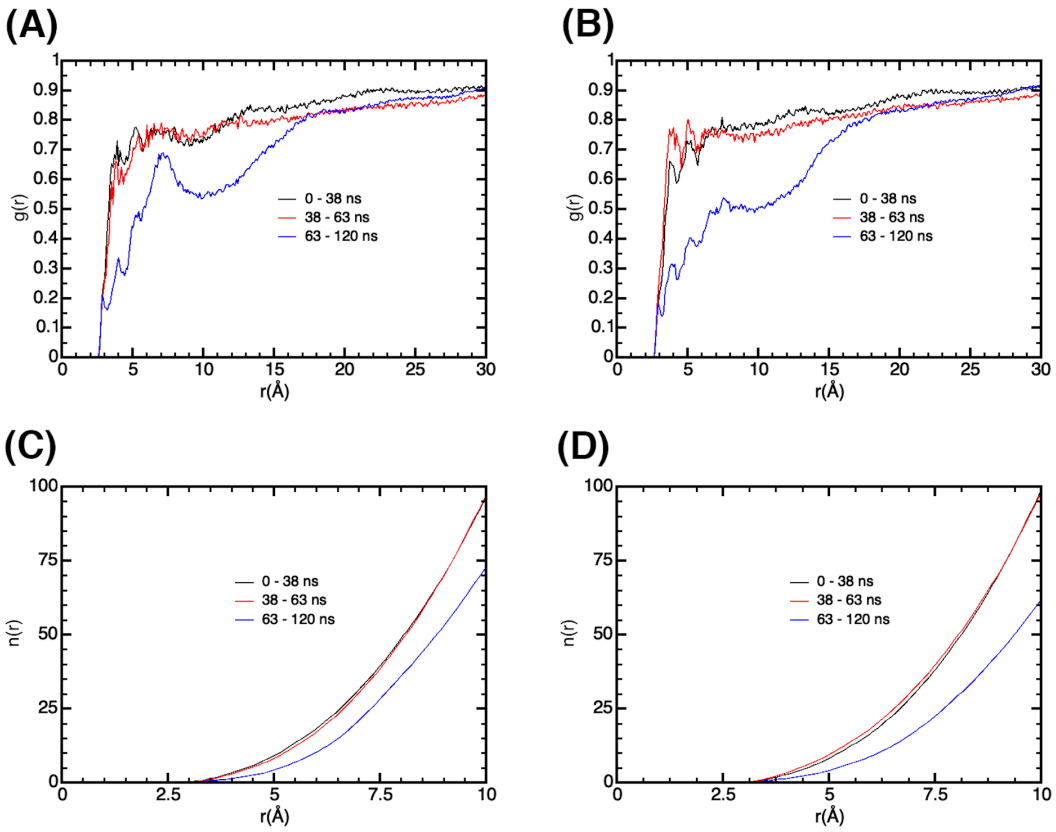

Figure S17: RDF for traj. 6 calculated using the same procedure as for Fig. S12.

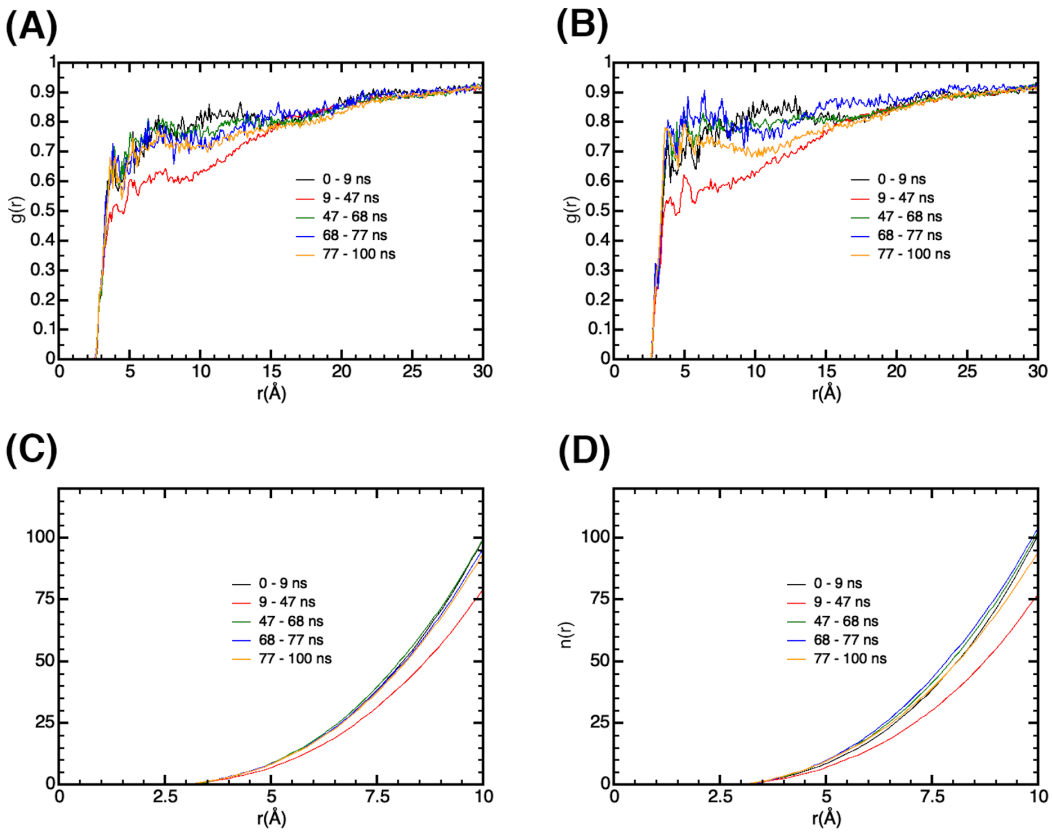

Figure S18: RDF for traj. 7 calculated using the same procedure as for Fig. S12.

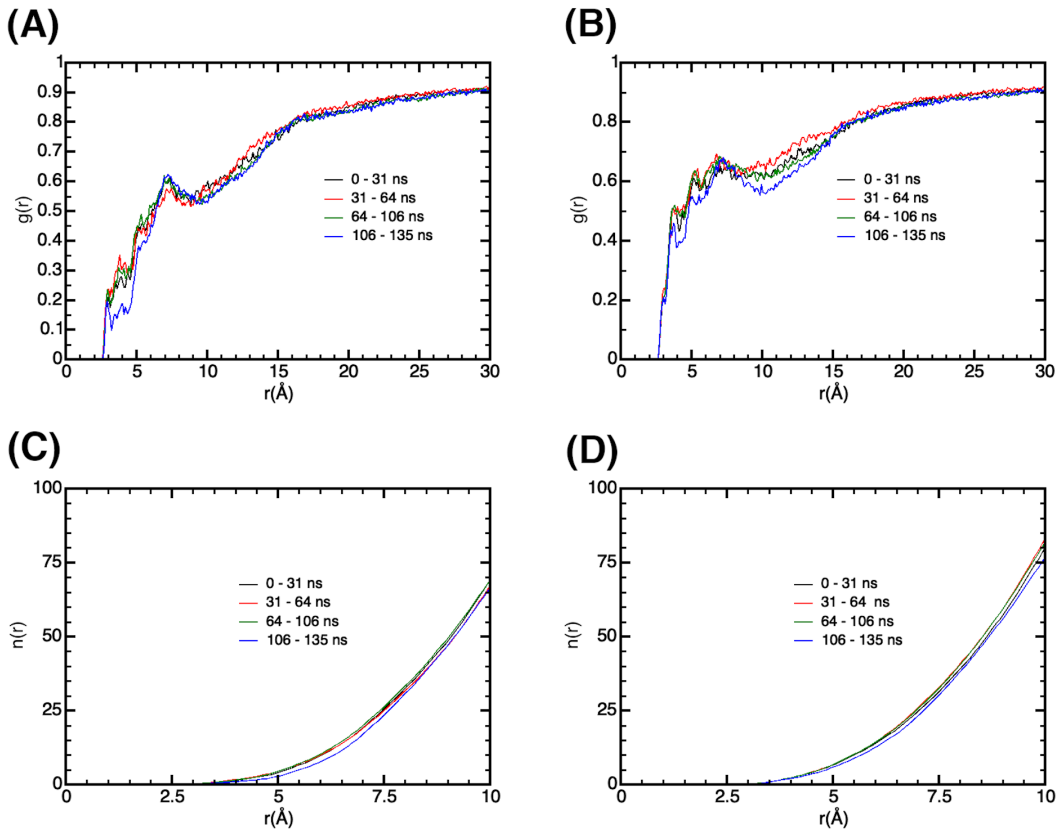

Figure S19: RDF for traj. 8 calculated using the same procedure as for Fig. S12. Adenine and Trp131 remain π-stacked for the majority of the 135 ns simulation.

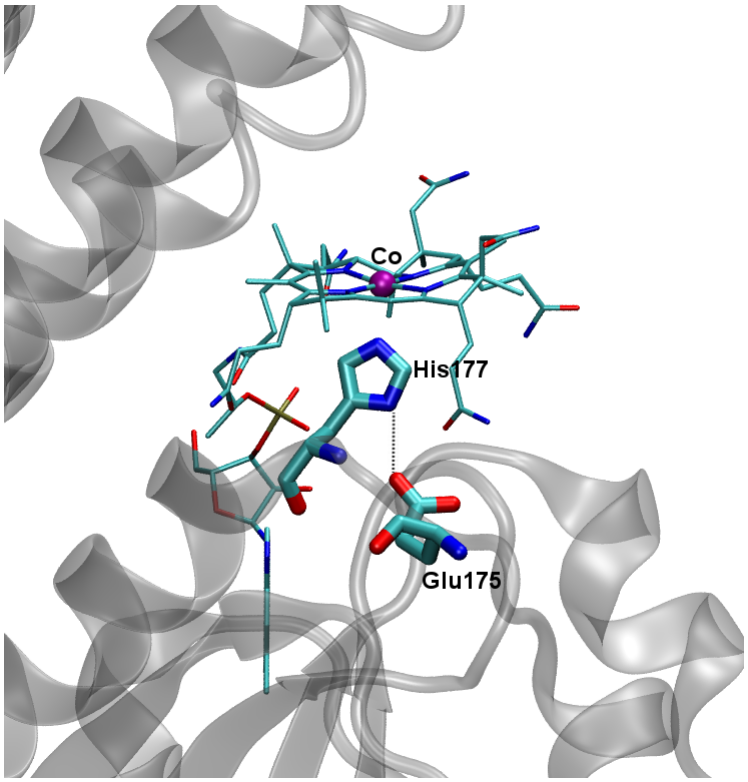

Figure S20: An illustration of the hydrogen bond between His177 (the lower axial ligand of Co) and Glu175 in CarH. In chain A of the crystal structure (PDB ID 5C8E),[5] the distance depicted as the dotted line is 2.92 Å. The hydrogen atoms have been omitted for clarity.

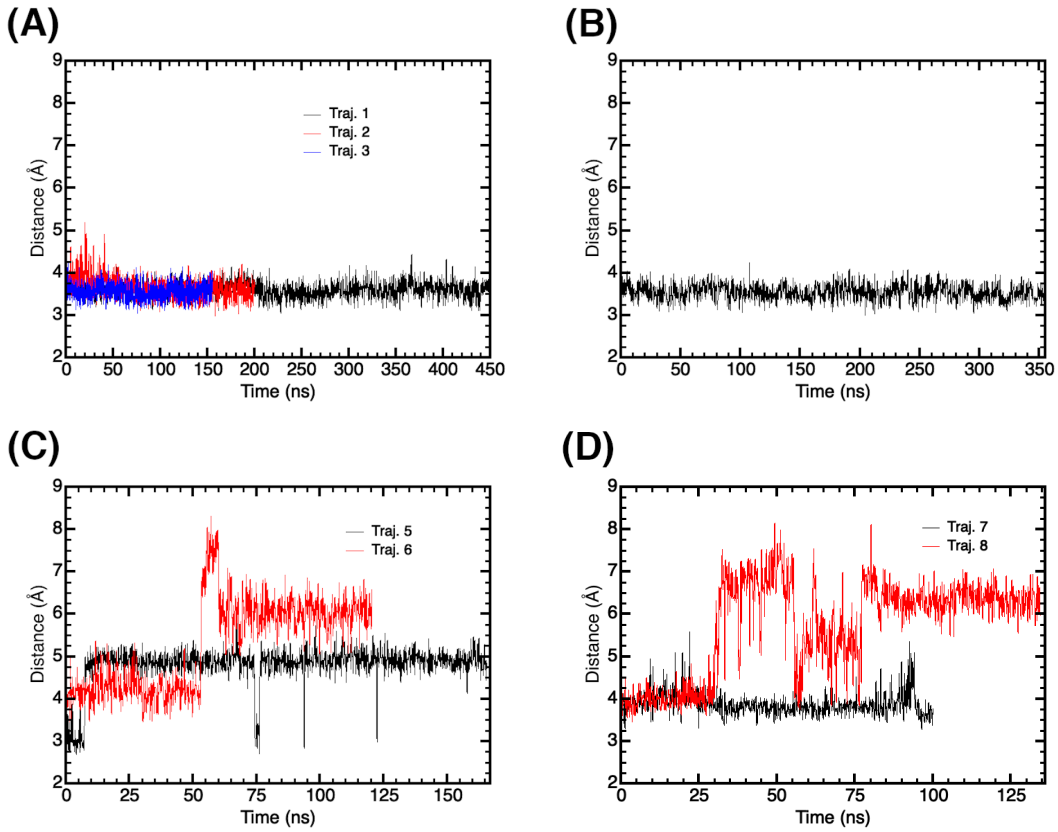

Figure S21: Time evolution of the distance between the N_δ (ND1) atom of His177 and a selected sidechain atom of residue number 175 in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) 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.

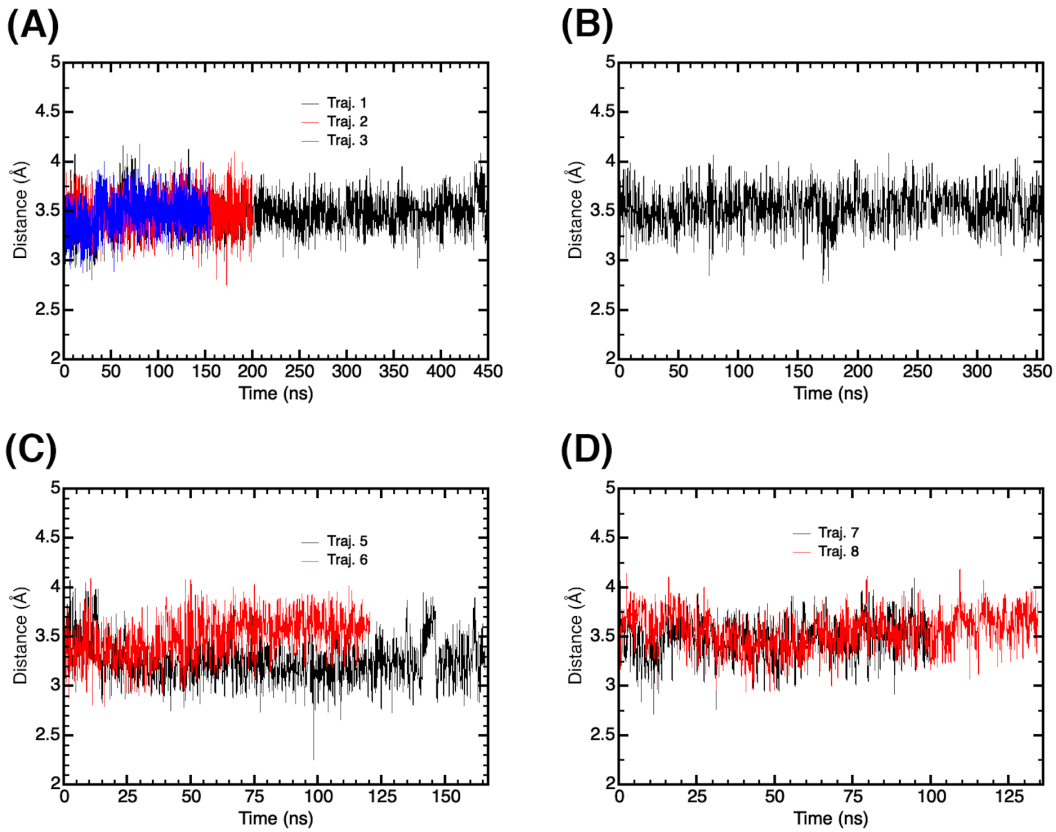

Figure S22: Time evolution of the distance between the H4′ atom of adenosyl (see Fig. S11) and Co in (A) three independent WT simulations with deprotonated His142, (B) a WT simulation with protonated His142, (C) two independent simulations of the E175Q mutant, and (D) two independent simulations of the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