Supporting Information for “Protonation of the Biliverdin IXα Chromophore in the Red and Far-red Photoactive States of Bacteriophytochrome”

Vaibhav Modi,1 Serena Donnini,2 Gerrit Groenhof,1 Dmitry Morozov1*

1Department of Chemistry and NanoScience Center, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Jyväskylä, Finland

2Department of Biological and Environmental Science and NanoScience Center, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Jyväskylä, Finland.

Corresponding Author
*dmitry.morozov@jyu.fi
Contents

1. Molecular modelling
2. Force field parameterization of BV chromophore
3. Additional results from MD simulations
4. Additional results from QM/MM calculations
5. Proton release/uptake pathway
6. The effect of the including Range-Separation on the chromophore force field
**List of Figures**

S1. Representation of water molecules added to the starting structure with DOWSER program
S2. Root mean square deviation of phytochrome backbone and chromophore
S3. Stick representations of the QM subsystems
S4. Distance distribution plots for the polar contact network involving BV, Asp207, His260 interactions in Pr state
S5. Distance distribution plots for non-covalent interaction between BV, hydrophobic cavity residues (Tyr 176, Phe203, Tyr263), Tyr 216, Arg254, and Ser274; and salt bridge between Asp207-Arg466 in Pr state
S6. Distance distribution plots for the polar contact network involving BV, Asp207, His260 interactions in Pfr state.
S7. Distance distribution plots for non-covalent interactions involving BV, Asp207, Tyr 216, Arg254, Ser274, Tyr263 and Ser468 in Pfr state.
S8. Q-band and Soret-band maxima using XMCQDPT2 method
S9. UV/Vis absorption spectra comparison for different sizes of the QM subsystem.
S10. CASSCF (12,12) active space orbitals generated for the XMCQDPT2 excited state energy calculations.
S11. Excited-state relaxed energy scans of the most important biliverdin torsions.
S12. The contribution of different excited state transitions (S_0 → S_n) to the absorption spectra computes with xMCQDPT2 method for the Pr and Pfr states with BV deprotonated at D-ring.
S13. Snapshot from MD simulation of Pr state showing the speculated proton-entry and -release wires.
S14. Comparison of CAM-B3LYP and B3LYP optimized BV geometry.
List of Tables

S1. Missing residues in DrBphP crystal structures 4O0P (Pr) and 4O01(Pfr).
S2. BV chromophore parameters generated for DOWSER program.
S3. Histidine residue tautomeric forms defined using pKa estimation for DrBphP MM simulations.
S4. Atomic charges for the BV chromophore in Pr state derived using RESP procedure for AMBER03 force fields
S5. Atomic charges for BV chromophore in Pfr state derived using RESP procedure for AMBER03 force fields
S6. Atomic charges used for Pr state Cys24 residue in AMBER03 force fields for MD simulations.
S7. Atomic charges used for Pfr state Cys24 residue in AMBER03 force fields for MD simulations.
S8. Comparison of DFT functionals based on the bond lengths for optimized BV chromophore for Pr and Pfr state full protonated chromophore.
**Molecular modelling**

The residues unresolved from X-ray crystallography in PDB models 4O0P and 4O01 were added using loop modelling, which are listed in Table S1. The longest range of missing residue (17 amino acids) lies at the N-terminal end, and the other 5 unresolved loop regions in the protein complex vary within length of 1-6 residues. A total of 100 new models were generated from loop optimization step of modelling using simulated annealing technique and sorted again based on DOPE score to retrieve an optimal phytochrome conformation.

**Table S1.** Missing residue regions in DrBphP crystal structures 4O0P (Pr) and 4O01 (Pfr). The amino acid residues are denoted by single letter abbreviations and the two chains of protein homodimer are named A and B.

| PDB Model    | Residues                                      |
|--------------|-----------------------------------------------|
| 4O0P         | A 1-17 (MASMTGGQQMGRGSMSR)                    |
|              | A 121-123 (AAG)                               |
|              | A 146-151 (WDSTGP)                            |
|              | A 415 (A)                                     |
|              | A 445-447 (GEG)                               |
|              | A 472-476 (QAKDD)                             |
|              | A 521-523 (HHH)                               |
|              | B 1-17 (MASMTGGQQMGRGSMSR)                    |
|              | B 121-123 (AAG)                               |
|              | B 145-151 (AWDSTGP)                           |
|              | B 415 (A)                                     |
|              | B 472-476 (QAKDD)                             |
|              | B 518-523 (HHHHHH)                            |
| 4O01         | A 1-19 (MASMTGGQQMGRGSMRDP)                   |
|              | A 220-221 (PA)                                |
|              | A 145-149 (AWDST)                             |
|              | A 465-475 (WGGATPDQAKD)                       |
|              | A 521-523 (HHH)                               |
|              | B 1-19 (MASMTGGQQMGRGSMRDP)                   |
|              | B 220-221 (PA)                                |
|              | B 145 (A)                                     |
|              | B 149 (T)                                     |
|              | B 465-475 (WGGATPDQAKD)                       |
|              | B 521-523 (HHH)                               |
A total of 81 and 84 water molecules were added in the Pr and Pfr phytochrome cavities. The buried water molecules placed by DOWSER were refined by energy optimization and short equilibration simulations. An environment comparable to the high resolution (1.45 Å) crystal structure of DrBphP CBD domain (PDB: 2O9C) was obtained (Figure S1). After the energy minimization and short equilibration runs the water molecules added with DOWSER form a network of non-covalent interactions with the chromophore and adjacent residues analogous to those in the experimentally resolved CBD structure.

**Table S2.** DOWSER parameters for BV chromophore generated using the atomic charges, bond lengths, angles, and dihedrals parameters derived for AMBER03 force field.

| Atom name | Type | Backward | Forward | Bond  | Angle | Dihedral |
|-----------|------|----------|---------|-------|-------|----------|
| CBA       | CH2  | NOT      | CAA     | 0     | 0     | 0        |
| CAA       | CH1  | CBA      | C3A     | 1.510 | 0     | 0        |
| C3A       | CR   | CAA      | C2A     | 1.343 | 125.9 | 0        |
| C4A       | CR   | C3A      | NOT     | 1.39  | 120   | 180      |
| C2A       | CH1  | C3A      | C1A     | 1.510 | 120   | 0        |
| CMA       | CH3  | C2A      | NOT     | 1.541 | 114   | 60       |
| C1A       | CR   | C2A      | N_A     | 1.604 | 112.1 | 0        |
| O_A       | O    | C1A      | NOT     | 1.226 | 124   | 180      |
| N_A       | N    | C1A      | NOT     | 1.375 | 112.1 | 0        |
| HNA       | H    | N_A      | NOT     | 1.026 | 126.9 | 180      |
| CHB       | CH1  | C4A      | C1B     | 1.393 | 122   | 60       |
| C1B       | CR   | CHB      | C2B     | 1.406 | 125.1 | 180      |
| N_B       | N    | C1B      | NOT     | 1.381 | 125.3 | 0        |
| HNB       | H    | N_B      | NOT     | 1.026 | 122.3 | 0        |
| C2B       | CR   | C1B      | C3B     | 1.439 | 126.5 | 180      |
| CMB       | CH3  | C2B      | NOT     | 1.497 | 123.5 | 0        |
| C3B       | CR   | C2B      | C4B     | 1.402 | 106.9 | 180      |
| CAB       | CH2  | C3B      | CBB     | 1.497 | 129   | 180      |
| CBB       | CH2  | CAB      | CGB     | 1.526 | 114   | 60       |
| CGB       | CR   | CBB      | NOT     | 1.522 | 116.6 | 180      |
| O1B       | O    | CGB      | NOT     | 1.222 | 12.5  | 60       |
| O2B       | O    | CGB      | NOT     | 1.363 | 111.5 | 60       |
| C4B       | CR   | C3B      | NOT     | 1.456 | 107.1 | 0        |
| Atom 1 | Atom 2 | Atom 3 | Bond Length | Angle 1 | Angle 2 |
|-------|-------|-------|-------------|---------|---------|
| CHC   | CH1   | C4B   | C1C         | 1.404   | 125     | 180     |
| C1C   | CR    | CHC   | C2C         | 1.406   | 126     | 180     |
| N_C   | N     | C1C   | NOT         | 1.381   | 125.3   | 0       |
| HNC   | H     | N_C   | NOT         | 1.026   | 122.3   | 0       |
| C2C   | CR    | C1C   | C3C         | 1.439   | 126.5   | 180     |
| CAC   | CH2   | C2C   | CBC         | 1.497   | 123.5   | 0       |
| CBC   | CH2   | CAC   | CGC         | 1.526   | 114     | 60      |
| CGC   | CR    | CBC   | NOT         | 1.522   | 116.6   | 180     |
| O1C   | O     | CGC   | NOT         | 1.222   | 126.5   | 60      |
| O2C   | O     | CGC   | NOT         | 1.363   | 111.5   | 60      |
| C3C   | CR    | C2C   | C4C         | 1.402   | 106.9   | 180     |
| CMC   | CH3   | C3C   | NOT         | 1.497   | 129     | 180     |
| C4C   | CR    | C3C   | NOT         | 1.456   | 107.1   | 0       |
| CHD   | CH1   | C4C   | C1D         | 1.404   | 125     | 180     |
| C1D   | CR    | CHD   | C2D         | 1.406   | 126     | 60      |
| N_D   | N     | C1D   | NOT         | 1.381   | 125.3   | 0       |
| HND   | H     | N_D   | NOT         | 1.026   | 122.3   | 0       |
| C2D   | CR    | C1D   | C3D         | 1.439   | 126.5   | 180     |
| CMD   | CH3   | C2D   | NOT         | 1.497   | 123.5   | 0       |
| C3D   | CR    | C2D   | C4D         | 1.402   | 106.9   | 180     |
| C3D   | CAD   | CBD   | CH1         | 1.404   | 120     | 180     |
| CBD   | CH2   | CAD   | NOT         | 1.363   | 120     | 180     |
| C4D   | CR    | C3D   | O_D         | 1.455   | 120     | 0       |
| O_D   | O     | C4D   | NOT         | 1.226   | 130.7   | 180     |

* The atom names refer to BV representation in Figure 1a.
Figure S1. A dotted sphere representation of water molecules added in the vicinity of biliverdin using DOWSER program and equilibrated using MD simulations, labelled as DOWSER water. The arrangement of DOWSER water molecules is identical to the crystal water arrangement around chromophore in high resolution (1.45 Å) structure of CBD domain from DrBphP in dark form.

The tautomeric form of Histidine (His) residues and protonation state of amino-acids with ionizable side chains are summarized in Table S3. The protonation form for other residues of photo-sensory core were assigned based on their standard pK\textsubscript{a} values in solution.

Table S3. Histidine residue tautomeric forms defined using pKa estimation for DrBphP MM simulations.

| Residue | Protonated nitrogen atom(s) |
|---------|---------------------------|
|         | Pr            | Pfr            |
| His44   | $\epsilon_2$ | $\epsilon_2$ |
| His52   | $\epsilon_2$ | $\epsilon_2$ |
| His61   | $\epsilon_2$ | $\epsilon_2$ |
| His124  | $\epsilon_2$ and $\delta_1$ | $\epsilon_2$ and $\delta_1$ |
| His130  | $\epsilon_2$ | $\epsilon_2$ |
| His152  | $\delta_1$ | $\delta_1$ |
| His210  | $\epsilon_2$ | $\epsilon_2$ |
| His215  | $\epsilon_2$ and $\delta_1$ | $\delta_1$ |
| His233  | $\epsilon_2$ | $\epsilon_2$ |
| His274  | $\epsilon_2$ | $\delta_1$ |
| His304  | $\epsilon_2$ | $\delta_1$ |
| His305 | ε2 | ε2 |
|------|-----|-----|
| His348 | ε2 | ε2 |
| His349 | ε2 | ε2 and δ1 |
| His358 | ε2 and δ1 | ε2 |
| His363 | δ1 | ε2 |
| His481 | ε2 | ε2 and δ1 |
| His498 | ε2 | ε2 |
| His518 | ε2 | ε2 |
| His519 | ε2 | ε2 and δ1 |
| His520 | ε2 | ε2 and δ1 |
| His521 | δ1 | ε2 |
| His522 | ε2 and δ1 | ε2 |
| His523 | ε2 and δ1 | ε2 and δ1 |

**Force field parameters for BV chromophore**

Table S4. Atomic charges and atom types for Pr state of BV chromophore used in the AMBER03 force field for the MM calculations. The hydrogen atoms are denoted Ax, where the letter A is the index of heavy atom connected to the corresponding hydrogen atom and x = {a, b, c} denotes the equivalent hydrogen atoms.

| Atom name** | Atom | Atom Type | Allprot | Deprot-A | Deprot-B | Deprot-C | Deprot-D |
|-------------|------|-----------|---------|----------|----------|----------|----------|
| CBA         | C    | CT        | -0.125784 | -0.203762 | -0.103050 | -0.229543 | -0.124648 |
| HBA1        | H    | HC        | 0.141099  | 0.136623  | 0.090903  | 0.147943  | 0.114963  |
| HBA2        | H    | HC        | 0.141099  | 0.136623  | 0.090903  | 0.147943  | 0.114963  |
| CAA         | C    | CM        | -0.128990 | -0.114071 | -0.138173 | -0.214065 | -0.212512 |
| HAA         | H    | HC        | 0.197467  | 0.183367  | 0.197327  | 0.211276  | 0.200575  |
| C3A         | C    | CM        | -0.222884 | -0.256168 | -0.140716 | -0.117059 | -0.209827 |
| C2A         | C    | CT        | 0.150926  | 0.197132  | 0.179136  | 0.323745  | 0.495323  |
| H2A         | H    | HC        | 0.078532  | 0.029831  | 0.122425  | 0.003098  | -0.046148 |
| CMA         | C    | CT        | -0.232203 | -0.063887 | -0.137302 | -0.100033 | -0.306863 |
| HMA1        | H    | HC        | 0.071541  | 0.011709  | 0.040031  | 0.022898  | 0.070351  |
| HMA2        | H    | HC        | 0.071541  | 0.011709  | 0.040031  | 0.022898  | 0.070351  |
|       |     |     |       |       |       |       |
|-------|-----|-----|-------|-------|-------|-------|
| HMA3  | H   | HC  | 0.071541 | 0.011709 | 0.040031 | 0.022898 | 0.070351 |
| C1A   | C   | C   | 0.494423 | 0.653508 | 0.496733 | 0.424996 | 0.454889 |
| OA    | O   | O   | -0.566941 | -0.712171 | -0.627838 | -0.583277 | -0.578586 |
| NA    | N   | NA  | -0.371493 | -0.581040 | -0.176933 | -0.209379 | -0.406888 |
| HNA   | H   | H   | 0.312913 | ----- | 0.283875 | 0.271741 | 0.0353203 |
| C4A   | C   | CC  | 0.293610 | 0.444239 | 0.048805 | 0.224996 | 0.454889 |
| CHB   | C   | C*  | -0.508717 | -0.620033 | -0.326311 | -0.170602 | -0.236868 |
| HHB   | H   | HC  | 0.180791 | 0.165923 | 0.136282 | 0.106955 | 0.121715 |
| C1B   | C   | CC  | 0.299924 | 0.326981 | 0.321287 | 0.000552 | -0.040423 |
| C2B   | C   | C*  | 0.096990 | 0.075598 | 0.132417 | 0.137246 | 0.197569 |
| CMB   | C   | CT  | -0.407671 | -0.218837 | -0.324154 | -0.324154 | -0.368699 |
| HMB1  | H   | HC  | 0.137772 | 0.068741 | 0.076523 | 0.096788 | 0.109631 |
| HMB2  | H   | HC  | 0.137772 | 0.068741 | 0.076523 | 0.096788 | 0.109631 |
| HMB3  | H   | HC  | 0.137772 | 0.068741 | 0.076523 | 0.096788 | 0.109631 |
| C3B   | C   | C*  | -0.223157 | -0.161872 | -0.236870 | -0.359789 | -0.385011 |
| C4B   | C   | CC  | 0.204458 | 0.023029 | 0.082417 | 0.137246 | 0.197569 |
| NB    | N   | NA  | -0.461834 | -0.226986 | -0.548085 | -0.239061 | -0.277691 |
| HNB   | H   | H   | 0.294567 | 0.183123 | ----- | 0.237684 | 0.264059 |
| CAB   | C   | CT  | -0.049467 | 0.026304 | -0.056134 | 0.256050 | 0.258543 |
| HAB1  | H   | HC  | 0.059532 | 0.016007 | 0.028238 | -0.023859 | -0.019262 |
| HAB2  | H   | HC  | 0.059532 | 0.016007 | 0.028238 | -0.023859 | -0.019262 |
| CBB   | C   | CT  | -0.147753 | 0.073027 | 0.009740 | -0.097219 | -0.123706 |
| HBB1  | H   | HC  | 0.032167 | -0.049630 | -0.014053 | -0.003895 | 0.000576 |
| HBB2  | H   | HC  | 0.032167 | -0.049630 | -0.014053 | -0.003895 | 0.000576 |
| CGB   | C   | C   | 0.868953 | 0.837699 | 0.843738 | 0.837507 | 0.839885 |
| O1B   | O   | O2  | -0.799804 | -0.815523 | -0.824408 | -0.833546 | -0.830977 |
| O2B   | O   | O2  | -0.799804 | -0.815523 | -0.824408 | -0.833546 | -0.830977 |
| CHC   | C   | CC  | -0.300568 | -0.186229 | -0.217611 | -0.238604 | -0.345144 |
| HHC   | H   | HC  | 0.260852 | 0.228768 | 0.229372 | 0.235788 | 0.247287 |
| C1C   | C   | C*  | 0.338657 | 0.130317 | 0.126218 | 0.235670 | 0.240732 |
| NC    | N   | NA  | -0.715401 | -0.369563 | -0.363737 | -0.563945 | -0.520438 |
| HNC   | H   | H   | 0.429492 | 0.343208 | 0.219782 | ----- | 0.338459 |
|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| **C4C** | C | CC | 0.575162 | 0.122641 | 0.132139 | 0.374510 | 0.388743 |
| **C3C** | C | CC | -0.059559 | 0.081143 | 0.053854 | 0.037294 | 0.148245 |
| **CMC** | C | CT | -0.354912 | -0.408173 | -0.296450 | -0.491414 | -0.332064 |
| **HMC1** | H | HC | 0.120436 | 0.124811 | 0.097032 | 0.147399 | 0.118294 |
| **HMC2** | H | HC | 0.120436 | 0.124811 | 0.097032 | 0.147399 | 0.118294 |
| **HMC3** | H | HC | 0.120436 | 0.124811 | 0.097032 | 0.147399 | 0.118294 |
| **C2C** | C | C* | -0.098782 | -0.173870 | -0.179305 | -0.174327 | -0.139621 |
| **CAC** | C | CT | 0.098471 | -0.017298 | -0.045656 | -0.001461 | -0.152533 |
| **HAC1** | H | HC | 0.018879 | 0.016670 | 0.020068 | 0.004290 | 0.065392 |
| **HAC2** | H | HC | 0.018879 | 0.016670 | 0.020068 | 0.004290 | 0.065392 |
| **CBC** | C | CT | -0.201235 | 0.118531 | -0.015955 | 0.145156 | 0.092259 |
| **HBC1** | H | HC | 0.032777 | -0.076538 | -0.014929 | -0.084033 | -0.052608 |
| **HBC2** | H | HC | 0.032777 | -0.076538 | -0.014929 | -0.084033 | -0.052608 |
| **CGC** | C | C | 0.928626 | 0.852904 | 0.852455 | 0.887954 |
| **O1C** | O | O2 | -0.830879 | -0.808738 | -0.809695 | -0.834242 | -0.828726 |
| **O2C** | O | O2 | -0.830879 | -0.808738 | -0.809695 | -0.834242 | -0.828726 |
| **CHD** | C | CC | -0.591351 | -0.281211 | -0.297969 | -0.299437 | -0.623872 |
| **HHD** | H | HC | 0.186294 | 0.194954 | 0.187372 | 0.125432 | 0.166525 |
| **C1D** | C | C* | 0.429679 | 0.189137 | 0.183425 | 0.210773 | 0.525233 |
| **ND** | N | NA | -0.756385 | -0.596745 | -0.584351 | -0.635954 | -0.769332 |
| **HND** | H | H | 0.492973 | 0.416667 | 0.414284 | 0.408138 | ----- |
| **C4D** | C | C | 0.802707 | 0.700539 | 0.695758 | 0.736142 | 0.849571 |
| **OD** | O | O | -0.574791 | -0.621307 | -0.620724 | -0.618986 | -0.619391 |
| **C3D** | C | CC | -0.088126 | -0.039825 | -0.033711 | -0.059290 | -0.153860 |
| **C2D** | C | C* | 0.009894 | 0.000393 | -0.008482 | 0.015008 | -0.019325 |
| **CMD** | C | CT | -0.304626 | -0.180547 | -0.216321 | -0.125136 | -0.187728 |
| **HMD1** | H | HC | 0.107849 | 0.072109 | 0.066567 | 0.052565 | 0.065375 |
| **HMD2** | H | HC | 0.107849 | 0.072109 | 0.066567 | 0.052562 | 0.065375 |
| **HMD3** | H | HC | 0.107849 | 0.072109 | 0.066567 | 0.052565 | 0.065375 |
| **CAD** | C | CM | -0.168161 | -0.172837 | -0.177418 | -0.182681 | -0.116367 |
| **HAD** | H | HC | 0.160099 | 0.155533 | 0.145617 | 0.164358 | 0.139860 |
| **CBD** | C | CM | -0.373148 | -0.462792 | -0.399914 | -0.429720 | -0.415734 |
| Atom name** | Atom Type | BV charges (e) – Pfr Illuminated state |
|------------|-----------|---------------------------------------|
|            |           | Allprot | Deprot-A | Deprot-B | Deprot-C | Deprot-D |
| CBA        | C         | CT      | -0.160213 | -0.189742 | -0.149204 | -0.134423 | -0.136000 |
| HBA1       | H         | HC      | 0.142406  | 0.139983  | 0.127506  | 0.124373  | 0.125055   |
| HBA2       | H         | HC      | 0.142406  | 0.139983  | 0.127506  | 0.124373  | 0.125055   |
| CAA        | C         | CM      | -0.150339 | -0.116493 | -0.143944 | -0.183544 | -0.128359  |
| HAA        | H         | HC      | 0.206208  | 0.168066  | 0.169436  | 0.177800  | 0.161937   |
| C3A        | C         | CM      | -0.190843 | -0.284840 | -0.184399 | -0.158934 | -0.246548  |
| C2A        | C         | CT      | 0.288527  | 0.325334  | 0.338244  | 0.301547  | 0.379720   |
| H2A        | H         | HC      | 0.035952  | -0.022687 | -0.007789 | 0.010866  | -0.008850  |
| CMA        | C         | CT      | -0.257769 | -0.269885 | -0.283010 | -0.391525 | -0.337182  |
| HMA1       | H         | HC      | 0.073841  | 0.061502  | 0.072656  | 0.112922  | 0.083724   |
| HMA2       | H         | HC      | 0.073841  | 0.061502  | 0.072656  | 0.112922  | 0.083724   |
| HMA3       | H         | HC      | 0.073841  | 0.061502  | 0.072656  | 0.112922  | 0.083724   |
| C1A        | C         | C       | 0.431185  | 0.626008  | 0.491165  | 0.509974  | 0.514265   |
| OA         | O         | O       | -0.575833 | -0.689026 | -0.606700 | -0.628127 | -0.587148  |
| NA         | N         | NA      | -0.213914 | -0.622246 | -0.314624 | -0.288515 | -0.441656  |
| HNA        | H         | H       | 0.212987  | ----      | 0.282468  | 0.257195  | 0.341992   |
| C4A        | C         | CC      | 0.095123  | 0.502234  | 0.146009  | 0.065336  | 0.174924   |
| CHB        | C         | C*      | -0.334989 | -0.666589 | -0.412996 | -0.247561 | -0.321444  |
| HHB        | H         | HC      | 0.166613  | 0.191091  | 0.175175  | 0.154362  | 0.170634   |
| C1B        | C         | CC      | 0.152106  | 0.349842  | 0.345132  | 0.036036  | 0.038104   |
| C2B        | C         | C*      | 0.090712  | 0.072554  | 0.040966  | 0.091563  | 0.125958   |
| CMB        | C         | CT      | -0.258848 | -0.207441 | -0.319834 | -0.306687 | -0.316116  |
| HMB1       | H         | HC      | 0.085754  | 0.065913  | 0.084576  | 0.085749  | 0.091067   |
| HMB2       | H         | HC      | 0.085754  | 0.065913  | 0.084576  | 0.085749  | 0.091067   |

** The atom names refer to BV representation in Figure 1a.

Table S5. Atomic charges and atom types for Pr state of BV chromophore used in the AMBER03 force field for the MM calculations. The hydrogen atoms are denoted Ax, where A is the index of heavy atom connected to the corresponding hydrogen atom and x = {a, b, c} denotes the equivalent hydrogen atoms.
|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| HMB3 | H   | HC  | 0.085754 | 0.065913 | 0.084576 | 0.085749 | 0.091067 |
| C3B  | C   | C*  | -0.107975 | -0.153443 | -0.254932 | -0.274516 | -0.322487 |
| C4B  | C   | CC  | 0.019292 | 0.036231 | 0.185116 | 0.093966 | 0.244521 |
| NB   | N   | NA  | -0.330574 | -0.250914 | -0.545432 | -0.261497 | -0.378494 |
| HNB  | H   | H   | 0.281417 | 0.206126 | -----     | 0.295026 | 0.305018 |
| CAB  | C   | CT  | 0.011177 | 0.017465 | 0.168369 | 0.236018 | 0.228751 |
| HAB1 | H   | HC  | 0.037059 | 0.016123 | 0.014037 | -0.005678 | 0.004231 |
| HAB2 | H   | HC  | 0.037059 | 0.016123 | 0.014037 | -0.005678 | 0.004231 |
| CBB  | C   | CT  | 0.005539 | 0.081418 | 0.836315 | 0.838509 | 0.869845 |
| HBB1 | H   | HC  | -0.024865 | -0.051371 | 0.024860 | 0.020123 | 0.034541 |
| HBB2 | H   | HC  | -0.024865 | -0.051371 | 0.024860 | 0.020123 | 0.034541 |
| CGB  | C   | C   | 0.386315 | 0.838509 | 0.869845 | 0.869237 | 0.869446 |
| O1B  | O   | O2  | -0.804384 | -0.816067 | -0.838478 | -0.839682 | -0.836738 |
| O2B  | O   | O2  | -0.804384 | -0.816067 | -0.838478 | -0.839682 | -0.836738 |
| CHC  | C   | CC  | -0.131675 | -0.209971 | -0.167397 | -0.183526 | -0.462604 |
| HHC  | H   | HC  | 0.208094 | 0.235189 | 0.225111 | 0.222066 | 0.255630 |
| C1C  | C   | C*  | 0.083961 | 0.132181 | 0.081733 | 0.200414 | 0.378707 |
| NC   | N   | NA  | -0.519627 | -0.316647 | -0.266306 | -0.522486 | -0.670893 |
| HNC  | H   | H   | 0.392200 | 0.326444 | 0.291588 | -----     | 0.382728 |
| C4C  | C   | CC  | 0.258241 | 0.014106 | -0.001728 | 0.287223 | 0.446492 |
| C3C  | C   | CC  | 0.012488 | 0.128103 | 0.087547 | 0.097066 | 0.095018 |
| CMC  | C   | CT  | -0.425555 | -0.388585 | -0.232960 | -0.304866 | -0.489485 |
| HMC1 | H   | HC  | 0.133515 | 0.115665 | 0.070119 | 0.083725 | 0.152595 |
| HMC2 | H   | HC  | 0.133515 | 0.115665 | 0.070119 | 0.083725 | 0.152595 |
| HMC3 | H   | HC  | 0.133515 | 0.115665 | 0.070119 | 0.083725 | 0.152595 |
| C2C  | C   | C*  | 0.125036 | -0.168090 | -0.170476 | -0.219102 | -0.147368 |
| CAC  | C   | CT  | -0.064168 | -0.015186 | 0.079050 | 0.042756 | -0.056536 |
| HAC1 | H   | HC  | 0.041743 | 0.019364 | -0.012050 | 0.003489 | 0.038338 |
| HAC2 | H   | HC  | 0.041743 | 0.019364 | -0.012050 | 0.003489 | 0.038338 |
| CBC  | C   | CT  | -0.030445 | 0.059306 | 0.07455 | 0.104223 | 0.040694 |
| HBC1 | H   | HC  | -0.007975 | -0.062378 | -0.091380 | -0.067907 | -0.037704 |
| HBC2 | H   | HC  | -0.007975 | -0.062378 | -0.091380 | -0.067907 | -0.037704 |
|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| CGC | C | C | 0.850120 | 0.868532 | 0.854408 | 0.866815 | 0.866541 |
| O1C | O | O2 | -0.749769 | -0.821857 | -0.821538 | -0.827200 | -0.817653 |
| O2C | O | O2 | -0.749769 | -0.821857 | -0.821538 | -0.827200 | -0.817653 |
| CHD | C | CC | -0.399765 | -0.272894 | -0.228179 | -0.335205 | -0.722267 |
| HHD | H | HC | 0.187392 | 0.231610 | 0.182002 | 0.204315 | 0.224194 |
| C1D | C | C* | 0.346411 | 0.261188 | 0.205031 | 0.294617 | 0.719153 |
| ND  | N | NA | -0.790471 | -0.818128 | -0.787264 | -0.838345 | -0.847725 |
| HND | H | H  | 0.441957 | 0.444510 | 0.413430 | 0.455579 | -----     |
| C4D | C | C  | 0.814382 | 0.841195 | 0.840706 | 0.834491 | 0.875858 |
| OD  | O | O  | -0.621731 | -0.676482 | -0.668571 | -0.656471 | -0.615953 |
| C3D | C | CC | -0.068393 | -0.133935 | -0.139937 | -0.067046 | -0.131867 |
| C2D | C | C* | 0.010889 | 0.052690 | 0.130575 | 0.013815 | -0.117855 |
| CMD | C | CT | -0.244646 | -0.262942 | -0.312858 | -0.272259 | -0.262124 |
| HMD1| H | HC | 0.110513 | 0.118988 | 0.123464 | 0.116000 | 0.112664 |
| HMD2| H | HC | 0.110513 | 0.118988 | 0.123464 | 0.116000 | 0.112664 |
| HMD3| H | HC | 0.110513 | 0.118988 | 0.123464 | 0.116000 | 0.112664 |
| CAD | C | CM | -0.173485 | -0.120615 | -0.135393 | -0.163882 | -0.097059 |
| HAD | H | HC | 0.173013 | 0.153962 | 0.166629 | 0.173063 | 0.152614 |
| CBD | C | CM | -0.418324 | -0.483019 | -0.472113 | -0.475402 | -0.460241 |
| HBD1| H | HC | 0.191871 | 0.190620 | 0.189349 | 0.189567 | 0.182800 |
| HBD2| H | HC | 0.191871 | 0.190620 | 0.189349 | 0.189567 | 0.182800 |

** The atom names refer to BV representation in Figure 1a.
Table S6. Atomic charges used for Pr state Cys24 residue in the AMBER03 force field for the MD simulations.

| Atom | Charges (e) – Pr Dark state |
|------|----------------------------|
|      | Pr | Deprot-A | Deprot-B | Deprot-C | Deprot-D |
| N    | -0.406250 | -0.406250 | -0.406250 | -0.406250 | -0.406250 |
| H    | 0.276625  | 0.276625  | 0.276625  | 0.276625  | 0.276625  |
| CA   | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  |
| HA   | 0.081325  | 0.081325  | 0.081325  | 0.081325  | 0.081325  |
| CB   | -0.209278 | -0.199800 | -0.180751 | -0.201389 | -0.221453 |
| HB1  | 0.163720  | 0.176122  | 0.176122  | 0.159492  | 0.166868  |
| HB2  | 0.163720  | 0.176122  | 0.176122  | 0.159492  | 0.166868  |
| SG   | -0.251045 | -0.299206 | -0.287515 | -0.257517 | -0.253368 |
| C    | -0.606750 | -0.606750 | -0.606750 | -0.606750 | -0.606750 |
| O    | -0.558450 | -0.558450 | -0.558450 | -0.558450 | -0.558450 |

Table S7. Atomic charges used for Pfr state Cys24 residue in the AMBER03 force field for the MD simulations.

| Atom | Charges (e) – Pfr Illuminated state |
|------|-----------------------------------|
|      | Pr | Deprot-A | Deprot-B | Deprot-C | Deprot-D |
| N    | -0.406250 | -0.406250 | -0.406250 | -0.406250 | -0.406250 |
| H    | 0.276625  | 0.276625  | 0.276625  | 0.276625  | 0.276625  |
| CA   | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  |
| HA   | 0.081325  | 0.081325  | 0.081325  | 0.081325  | 0.081325  |
| CB   | -0.199281 | -0.213535 | -0.210651 | -0.202524 | -0.199082 |
| HB1  | 0.155258  | 0.166013  | 0.162773  | 0.159005  | 0.155258  |
| HB2  | 0.155258  | 0.166013  | 0.162773  | 0.159005  | 0.155258  |
| SG   | -0.268582 | -0.249284 | -0.270392 | -0.275985 | -0.268582 |
| C    | -0.606750 | -0.606750 | -0.606750 | -0.606750 | -0.606750 |
| O    | -0.558450 | -0.558450 | -0.558450 | -0.558450 | -0.558450 |
Additional results from MD simulations

Figure S2. Root mean-square deviation (rmsd) for the protein backbone and the chromophore as a function of time over 100ns of MD trajectories for all 5 protonation models of DrBphP in dark state.

Figure S3. BV chromophore atoms included in the QM-full and QM-small subsystem for the single state excited state energy calculations using TD-DFT and XMCQDPT2 level of theory. In the QM-small subsystem the propionate side chains and methyl groups were replaced by hydrogen atoms.
Figure S4. Distance distribution plots of the polar contacts between the BV A, B, and C pyrrole ring N-atoms, His260, and Asp207 in the chromophore binding pocket (shown earlier in Figure 3) for the 5 Pr state models with different protonation form of BV.
Figure S5: Distance distribution plots for the 5 Pr state models with different protonation form of BV. The electrostatic interactions formed by salt-bridge Asp207-Arg466, hydrophobic cavity residues (Tyr176, Phe203, Tyr263) and D-ring of chromophore, chromophore B/C-ring side-chain carboxylate group and Tyr216, Arg254, and Ser274 which stabilize the chromophore in the CBD domain in Pr state.
Figure S6. Distance distribution plots of the polar contacts between the BV A, B, and C pyrrole ring N-atoms, His260, and Asp207 in the chromophore binding pocket (shown earlier in Figure 3) for the 5 Pfr state models with different protonation form of BV.
Figure S7: Distance distribution plots for salt-bridge Asp207-Arg466, Tyr263-Ser468, chromophore B/C-ring side-chain carboxylate group and Tyr 216, Arg254, and Ser274 which are the conserved non-covalent interaction in active site of Pfr state.
Additional results from QM/MM calculations

Figure S8: Inclusion of more states (nstates=10) in the xMCQDPT2 calculations reproduces the Soret-band for the Pr and Pfr fully protonated BV models.
Effect of QM subsystem size and protein environment: Figure S9 compares the UV/Vis spectra for the different QM subsystem sizes for the Pr state with a fully protonated BV. The inclusion of pyrrole water or protein environment (using point charges) in the QM subsystem does not show significant change in the Q band maxima, with the corresponding peak observed at ~690 nm. This suggests that the point charge representation of the pyrrole water is sufficiently accurate. Excluding the protein environment and the propionate side chains of the chromophore do not have any effect on the Q-band maxima, which indicates that future computations could be carried out with the QM-small subsystem to save computational effort without any qualitative loss.

Figure S9. UV/Vis absorption spectra for the Pr state with fully protonated chromophore calculated using BLYP/6-31G(d,p) level of theory for the QM-full sub-system (see Figure S3) with the pyrrole water (blue) and without the pyrrole water (black). The experimental spectrum is shown as brown dotted line.
Figure S10: CASSCF (12,12) active space orbitals used in the XMCQDPT2 excited state energy calculations.
Figure S11: Excited-state relaxed energy scans of the most important biliverdin torsions. Excited $S_1$ state is shown with red line while ground $S_0$ state are blue line. The scans has been done at the SA(5)-CASSCF(12,12)/cc-pVDZ level of theory on the model chromophore in vacuo. The results show that (12,12) active space used in the work is stable with respect to the chromophore distortions.
Figure S12. The contribution of different excited state transitions ($S_0 \rightarrow S_n$) to the absorption spectra computes with xMCQDPT2 method for the Pr and Pfr states with chromophore deprotonated at D-ring.

Proton release/uptake pathway

In our simulations we observe occasionally an exchange of the pyrrole water with bulk solvent molecules in both Pr anf Pfr. This exchange was observed by monitoring the hydrogen bond distance between BV and nearest water molecule. The pyrrole water exchange involving the residue His260 might provide a pathway for the proton from the chromophore D-ring.

Figure S13: Snapshot from MD simulation of Pr state with a fully protonated chromophore illustrating a potential proton wire involving the chromophore, residue His260, pyrrole water and several solvent molecules. Colour code: carbon atoms of BV in turquoise and of the amino acid side chains involved in the proton wires in orange, oxygen atoms in red, nitrogen atoms in blue. Aliphatic hydrogen atoms are not shown.
**The effect of the including Range-Separation on the chromophore force field**

While the CAM-B3LYP functional has shown good performance for charge transfer excitations and excited-state dynamics in conjugated systems, it may not be the optimal choice for describing the ground state properties in such systems (M. A. Rohrdanz, J. Herbert: Simultaneous benchmarking of ground- and excited-state properties with long-range-corrected density functional theory, *J. Chem. Phys.* **129** (2008) 034107), we show below (Fig. S14 and Table S8) that the most critical force field parameters, namely bond lengths, angles and charges are essentially the same, irrespective of whether the geometry has been optimized with B3LYP or CAM-B3LYP.”
Figure S14: A superimposed stick representation of the optimized BV chromophore geometry using B3LYP (cyan) and CAM-B3LYP (green) functionals and basis set 6-31G* for (a) Pr and (d) Pfr state with fully protonated chromophore. A linear regression fit of the optimized geometry bond lengths (b)-Pr, (e)-Pfr and partial charges (c)-Pr, (f)-Pfr to compare the effect of using B3LYP instead of CAM-B3LYP for geometry optimization step in the RESP procedure.
Table S8: Comparison of DFT functionals based on the bond lengths for optimized BV chromophore for Pr and Pfr state full protonated chromophore.

| Bonds         | Pr (fully protonated BV) | Pfr (fully protonated BV) |
|---------------|--------------------------|----------------------------|
|               | CAM-B3LYP | B3LYP | Std. dev. | CAM-B3LYP | B3LYP | Std. dev. |
| CBA-S (CYS24) | 1.832     | 1.858 | -0.026    | 1.834 | 1.857 | -0.023 |
| CBA-CAA       | 1.494     | 1.494 | 0.000     | 1.439 | 1.495 | -0.056 |
| CAA-C3A       | 1.335     | 1.346 | -0.011    | 1.341 | 1.352 | -0.011 |
| C3A-C2A       | 1.518     | 1.524 | -0.006    | 1.491 | 1.491 | 0.000  |
| C2A-CMA       | 1.535     | 1.543 | -0.008    | 1.535 | 1.544 | -0.009 |
| C2A-C1A       | 1.526     | 1.534 | -0.008    | 1.497 | 1.497 | 0.000  |
| C1A-OA        | 1.213     | 1.219 | -0.006    | 1.230 | 1.236 | -0.006 |
| C1A-NA        | 1.372     | 1.378 | -0.006    | 1.339 | 1.339 | 0.000  |
| NA-C4A        | 1.400     | 1.403 | -0.003    | 1.449 | 1.449 | 0.000  |
| C4A-C3A       | 1.478     | 1.476 | 0.002     | 1.488 | 1.488 | 0.000  |
| C4A-CHB       | 1.350     | 1.363 | -0.013    | 1.344 | 1.357 | -0.013 |
| CHB-C1B       | 1.438     | 1.432 | 0.006     | 1.439 | 1.434 | 0.005  |
| C1B-NB        | 1.375     | 1.388 | -0.013    | 1.382 | 1.389 | -0.007 |
| C1B-C2B       | 1.396     | 1.405 | -0.009    | 1.388 | 1.403 | -0.015 |
| C2B-CMB       | 1.497     | 1.502 | -0.005    | 1.497 | 1.499 | -0.002 |
| C2B-C3B       | 1.415     | 1.417 | -0.002    | 1.430 | 1.429 | 0.001  |
| C3B-CAB       | 1.489     | 1.489 | 0.000     | 1.499 | 1.505 | -0.006 |
| CAB-CBB       | 1.544     | 1.559 | -0.015    | 1.541 | 1.545 | -0.004 |
| CBB-CGB       | 1.572     | 1.583 | -0.011    | 1.570 | 1.575 | -0.005 |
| CGB-O1B       | 1.240     | 1.247 | -0.007    | 1.242 | 1.247 | -0.005 |
| CGB-O2B       | 1.261     | 1.261 | 0.000     | 1.260 | 1.267 | -0.007 |
| C3B-C4B       | 1.411     | 1.422 | -0.011    | 1.386 | 1.400 | -0.014 |
| C4B-NB        | 1.401     | 1.406 | -0.005    | 1.389 | 1.394 | -0.005 |
| C4B-CHC | 1.398 | 1.400 | -0.002 | 1.447 | 1.441 | 0.006 |
| CHC-C1C | 1.381 | 1.391 | -0.010 | 1.345 | 1.357 | -0.012 |
| C1C-NC | 1.410 | 1.414 | -0.004 | 1.422 | 1.423 | -0.001 |
| C1C-C2C | 1.415 | 1.419 | -0.004 | 1.521 | 1.523 | -0.002 |
| C2C-CAC | 1.491 | 1.495 | -0.003 | 1.545 | 1.554 | -0.009 |
| CAC-CBC | 1.568 | 1.579 | -0.011 | 1.525 | 1.529 | -0.004 |
| CBC-CGC | 1.554 | 1.562 | -0.008 | 1.517 | 1.524 | -0.007 |
| CGC-O1C | 1.240 | 1.245 | -0.005 | 1.200 | 1.206 | -0.006 |
| CGC-O2C | 1.270 | 1.277 | -0.007 | 1.351 | 1.357 | -0.006 |
| C2C-C3C | 1.398 | 1.405 | -0.007 | 1.514 | 1.518 | -0.004 |
| C3C-CMC | 1.497 | 1.500 | -0.003 | 1.491 | 1.495 | -0.004 |
| C3C-C4C | 1.400 | 1.407 | -0.007 | 1.343 | 1.356 | -0.013 |
| C4C-NC | 1.369 | 1.385 | -0.016 | 1.410 | 1.415 | -0.005 |
| C4C-CHD | 1.437 | 1.431 | 0.006 | 1.463 | 1.458 | 0.005 |
| CHD-C1D | 1.357 | 1.372 | -0.015 | 1.344 | 1.357 | -0.013 |
| C1D-ND | 1.375 | 1.379 | -0.004 | 1.397 | 1.399 | -0.002 |
| C1D-C2D | 1.479 | 1.475 | 0.004 | 1.477 | 1.473 | 0.004 |
| C2D-CMD | 1.494 | 1.499 | -0.005 | 1.494 | 1.498 | -0.004 |
| C2D-C3D | 1.349 | 1.362 | -0.013 | 1.354 | 1.369 | -0.015 |
| C3D-CAD | 1.454 | 1.451 | 0.003 | 1.455 | 1.452 | 0.003 |
| CAD-CBD | 1.336 | 1.343 | -0.007 | 1.334 | 1.343 | -0.009 |
| C3D-C4D | 1.507 | 1.509 | -0.002 | 1.500 | 1.501 | 0.001 |
| C4D-OD | 1.214 | 1.221 | -0.007 | 1.217 | 1.224 | -0.007 |
| C4D-ND | 1.385 | 1.393 | -0.008 | 1.375 | 1.383 | -0.008 |