Structural Determinants of Improved Fluorescence in a Family of Bacteriophytochrome-based Infrared Fluorescent Proteins: Insights from Continuum Electrostatic Calculations and Molecular Dynamics Simulations.

Supplementary Information

Mikolaj Feliks,[1,2,3] Céline Lafaye,[1,2,3] Xiaokun Shu,[4,5]
Antoine Royant,[1,2,3,6,†] Martin Field[1,2,3,†]

[1)] Université Grenoble Alpes, IBS, F-38044 Grenoble, France

[2)] CNRS, IBS, F-38044 Grenoble, France

[3)] CEA, IBS, F-38044 Grenoble, France

[4)] Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158, USA

[5)] Cardiovascular Research Institute, University of California, San Francisco, California 94158, USA

[6)] European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, Grenoble, 38043, France

[†)] to whom correspondence should be addressed; e-mail: antoine.royant@ibs.fr, martin.field@ibs.fr

July 18, 2016

Contents

List of Figures

List of Tables

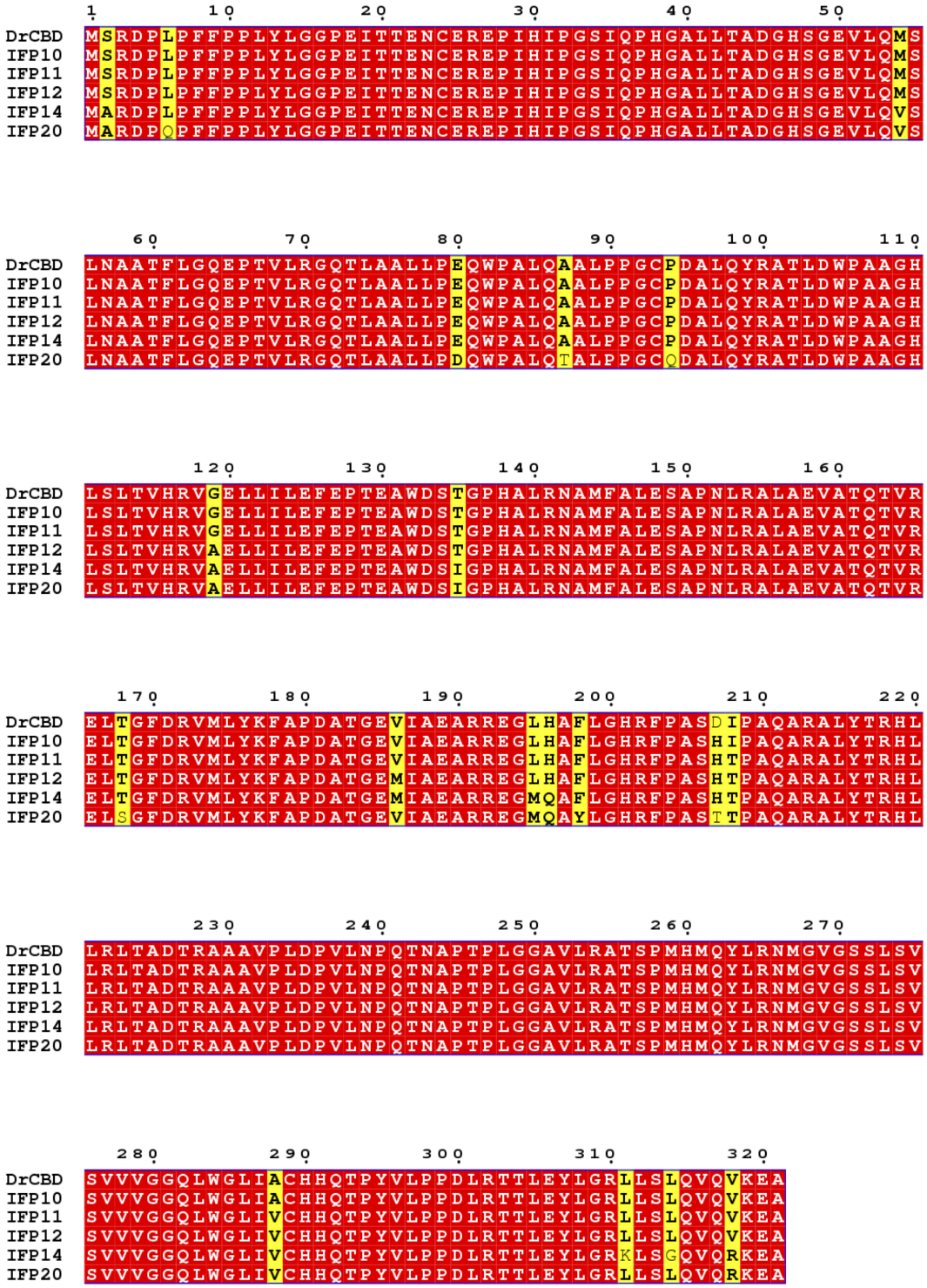


Figure S1: Sequence alignment between the chromophore-binding domain of the bacteriophytochrome from *Deinococcus radiodurans Dr*CBD and infrared fluorescent proteins IFP1.0, IFP1.1, IFP1.2, IFP1.4 and IFP2.0. The alignment was performed via the ClustalW server[1] and the figure was prepared with ESPript.[2]

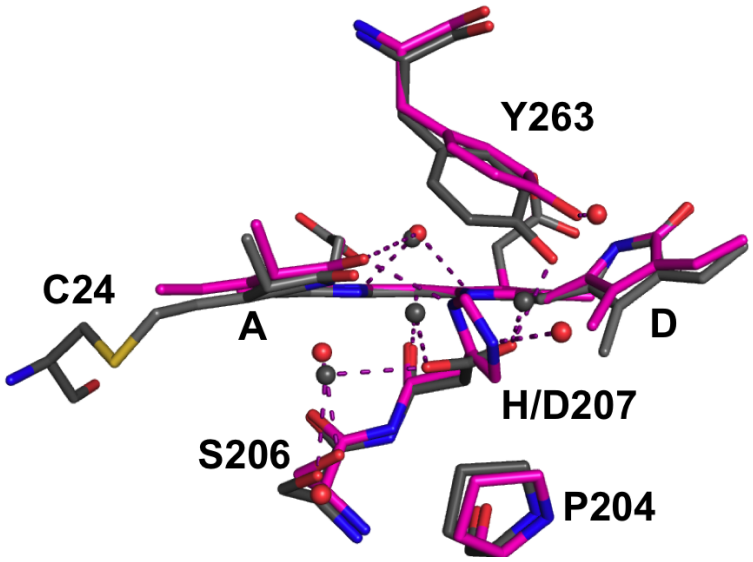


Figure S2: Comparison between the chromophore environment in *Dr*CBD (PDB entry code: 2O9C, dark grey) and IFP1.0 (PDB entry code: 3S7O, magenta). The figure was prepared with PyMOL.[3]

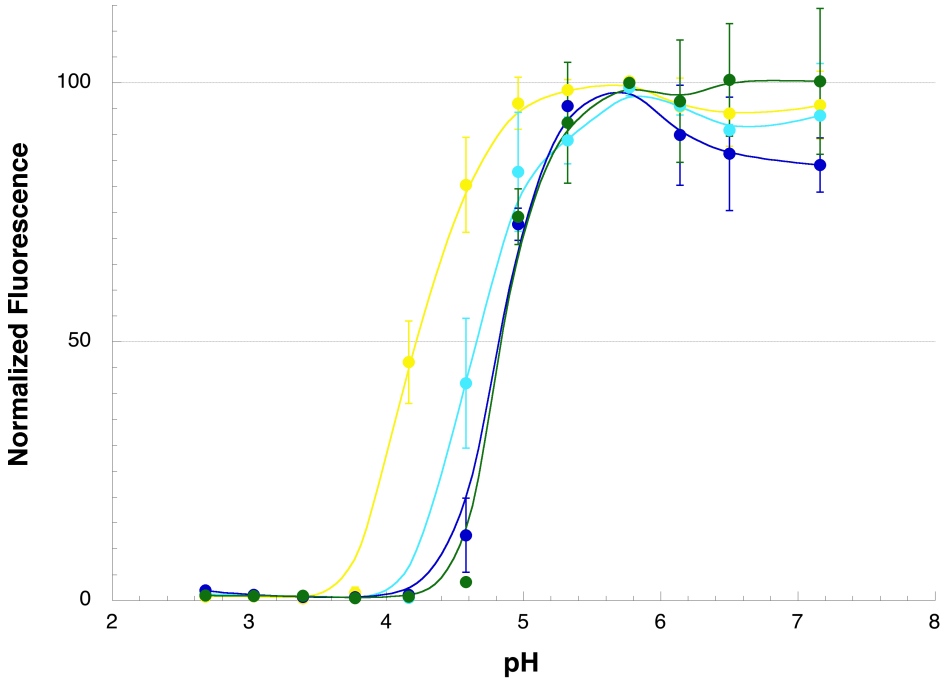


Figure S3: pH dependence of fluorescence for IFP1.1 (blue), IFP1.2 (green), IFP1.4 (cyan) and IFP2.0 (yellow). Fluorescence excitation was set at 640 nm and fluorescence emission was monitored between 660 and 800 nm.

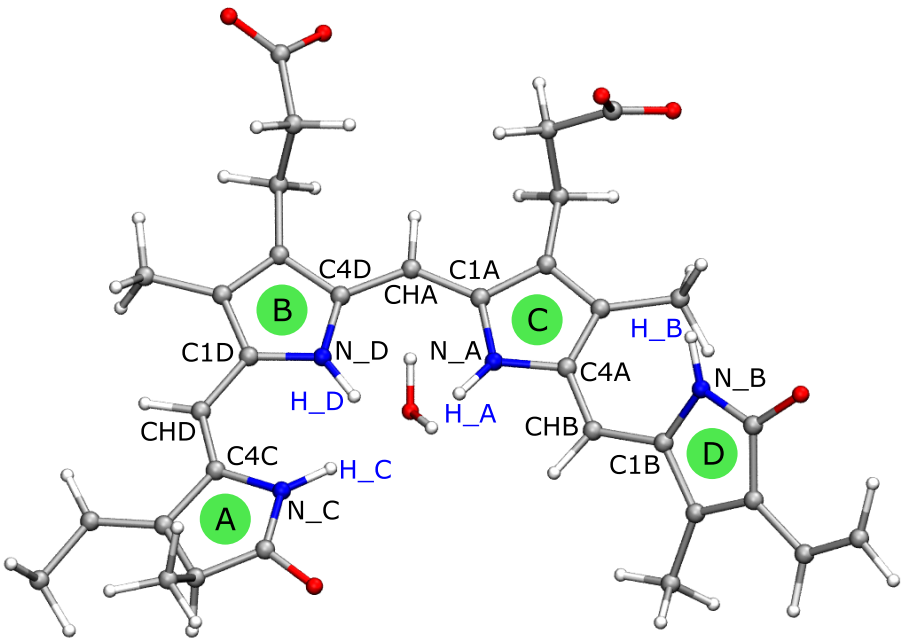
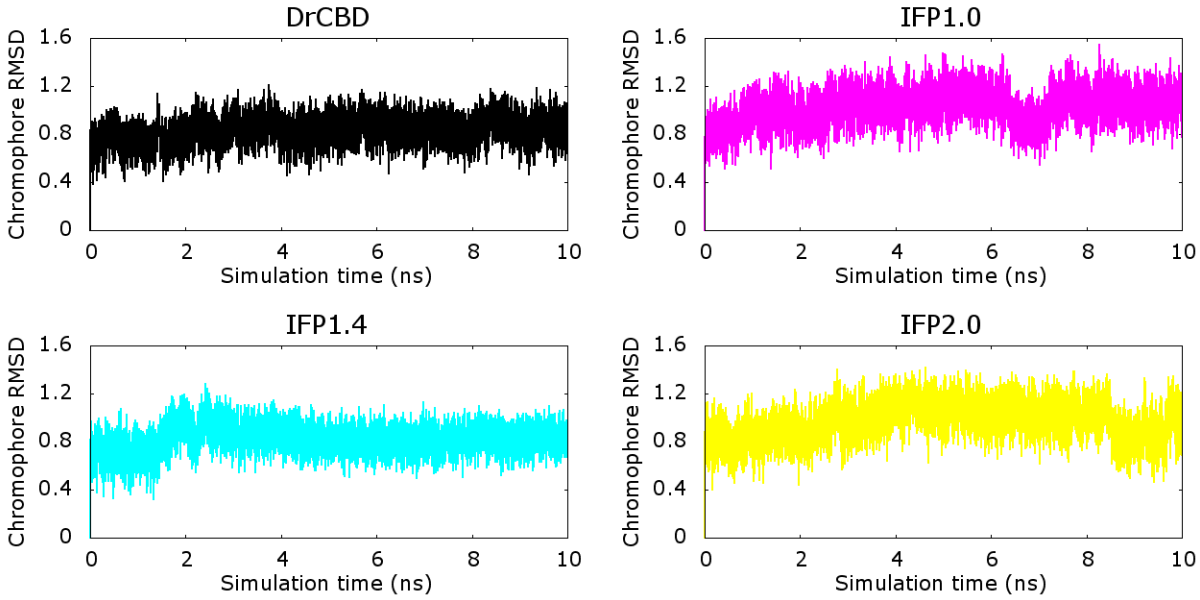


Figure S4: Fully protonated biliverdin chromophore as seen in the binding pocket of *Dr*CBD. Important atoms discussed in the text are labeled. Hydrogens at titratable positions are labeled in blue. The other atoms are labeled in black. The dihedral angles discussed in the text are numbered in the counterclockwise direction starting from ring A: **1**: N_C, C4C, CHD, C1D; **2**: C4C, CHD, C1D, N_D; **3**: N_D, C4D, CHA, C1A; **4**: C4D, CHA, C1A, N_A; **5**: N_A, C4A, CHB, C1B; **6**: C4A, CHB, C1B, N_B;

(a) Neutral pH



(b) Low pH

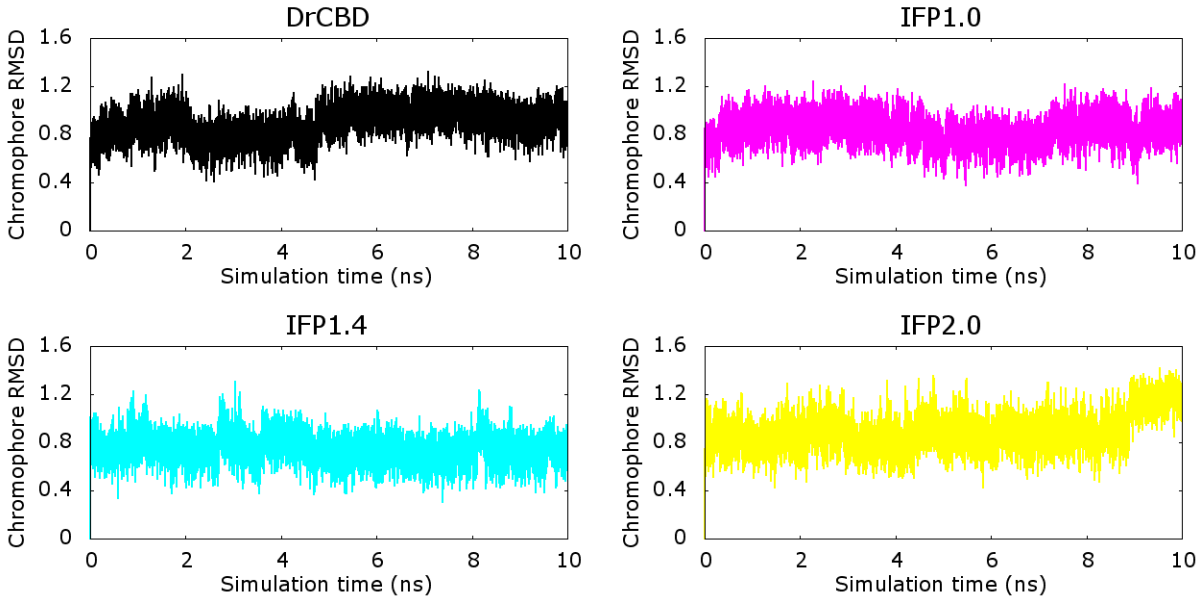


Figure S5: Time evolution of the root-mean-square deviations of the chromophore atoms in different proteins with respect to the final geometry from the heating stage. Two situations are shown: a) neutral pH; and b) low pH.

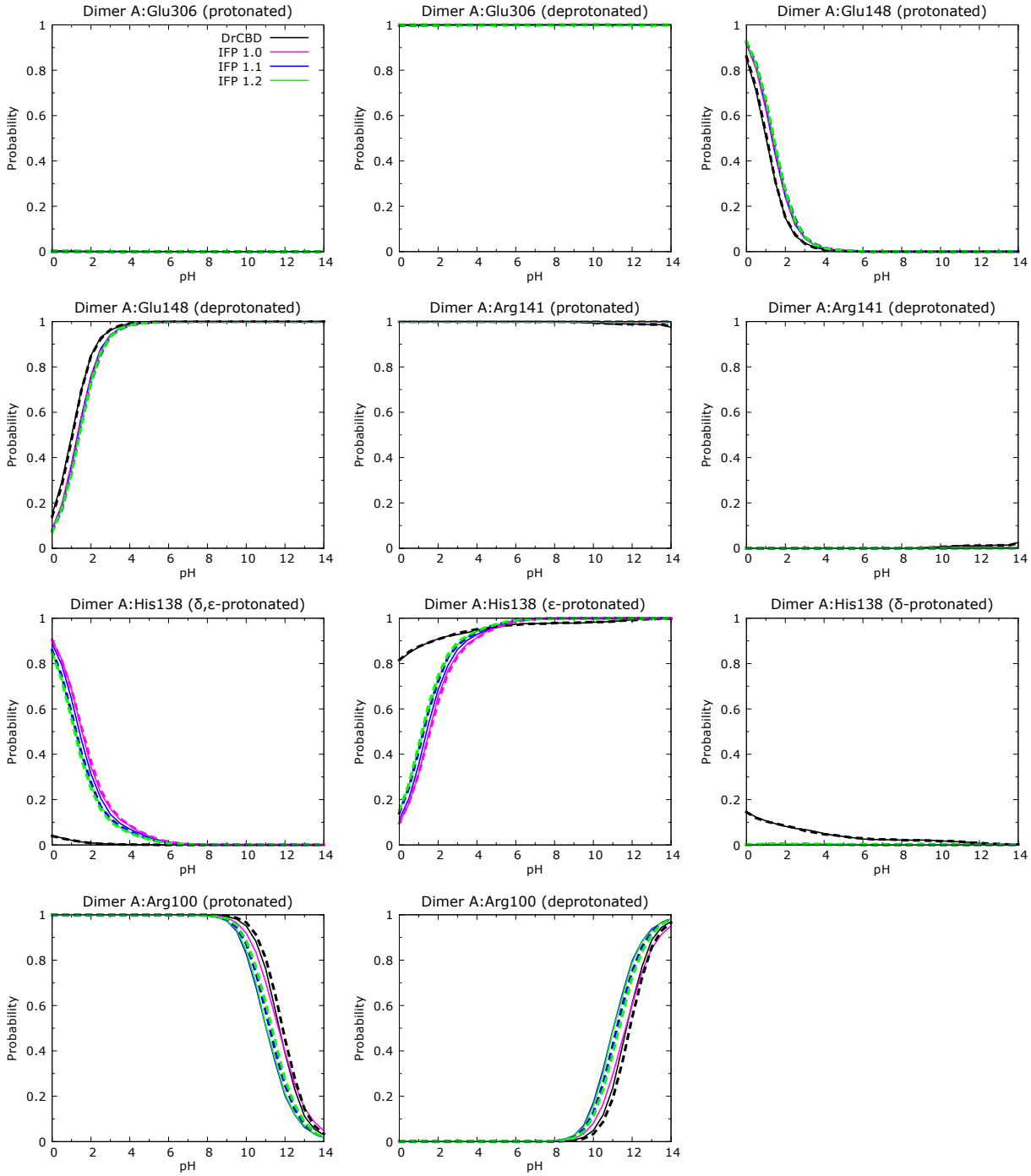


Figure S6: Titration curves of different protonation forms of titratable residues at the interface between the monomers (Arg100, His138, Arg141, Glu148, and Glu306). The calculations were performed on the structure of the complete dimer; only the curves for monomer A are shown. Solid and dashed lines represent the situations with the fully protonated and B-deprotonated chromophore, respectively.

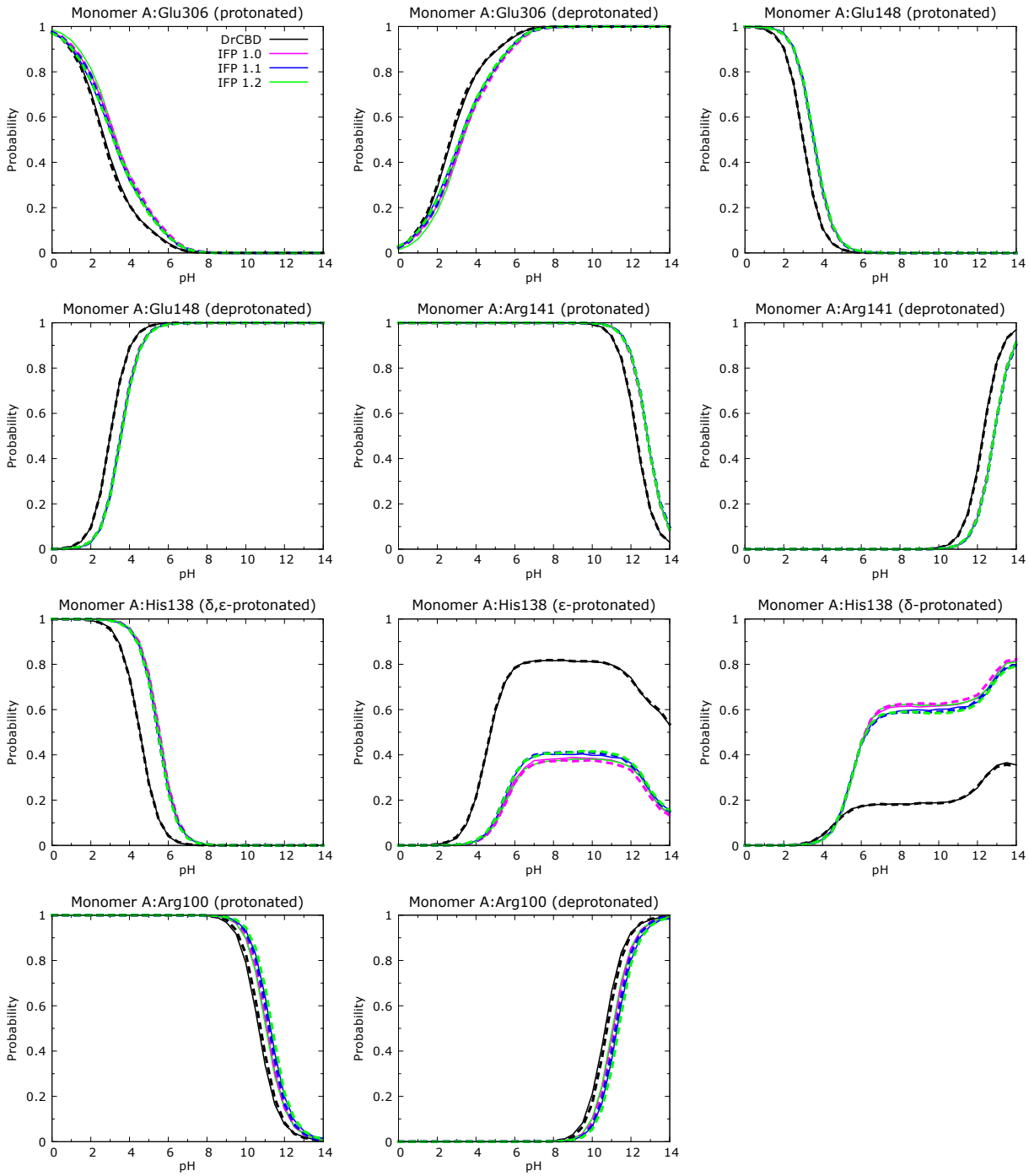


Figure S7: Titration curves of different protonation forms of titratable residues at the interface between the monomers (Arg100, His138, Arg141, Glu148, and Glu306). The calculations were performed on the structure of the isolated monomer A; monomer B has been removed from the model. Solid and dashed lines represent the situations with the fully protonated and B-deprotonated chromophore, respectively.

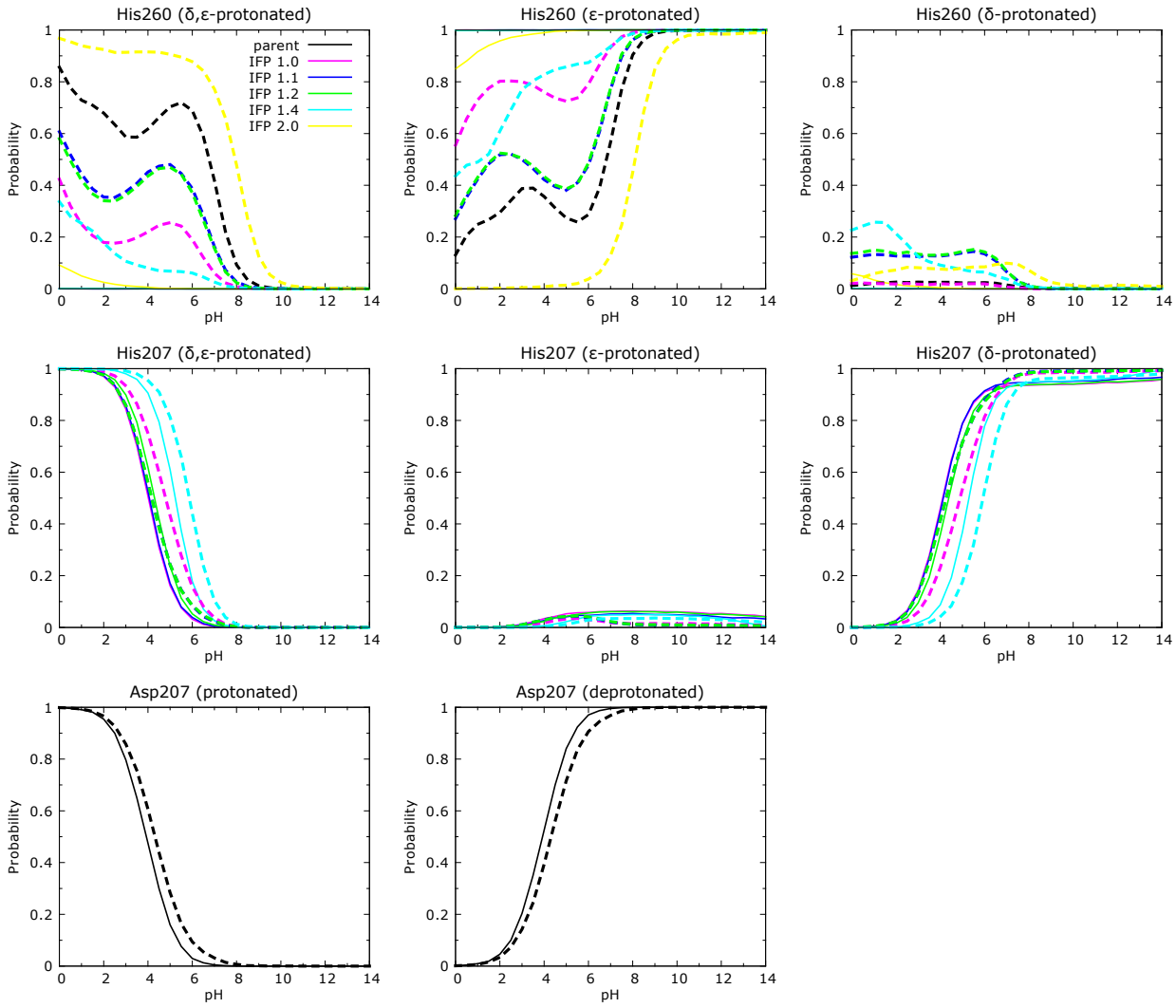


Figure S8: Probabilities of occurrence, calculated as a function of pH, of different protonation forms of His260, Asp207 and His207. The plots of the fully deprotonated forms of the histidines were omitted for clarity, since their probabilities of occurrence were calculated to be always zero. Solid and dashed lines represent the situations with the chromophore fixed in its fully protonated and B-deprotonated form, respectively. If the chromophore is fully protonated, His260 remains ϵ-protonated. Otherwise, His260 shows an irregular protonation behavior, where two protonation forms dominate, ϵ- or fully protonated, depending on the pH and the particular protein. Asp207 displays a typical protonation behavior. His207 has its midpoint pK_a in the range of 4<pH<6, where the fluorescence is known to rapidly increase. The probability curves for Asp207 and His207 show only small displacements depending on the protonation state of the chromophore. The probability curves of IFP 1.1 and IFP 1.2 align closely, because these mutant proteins differ by only a single mutation in a region remote from the chromophore.

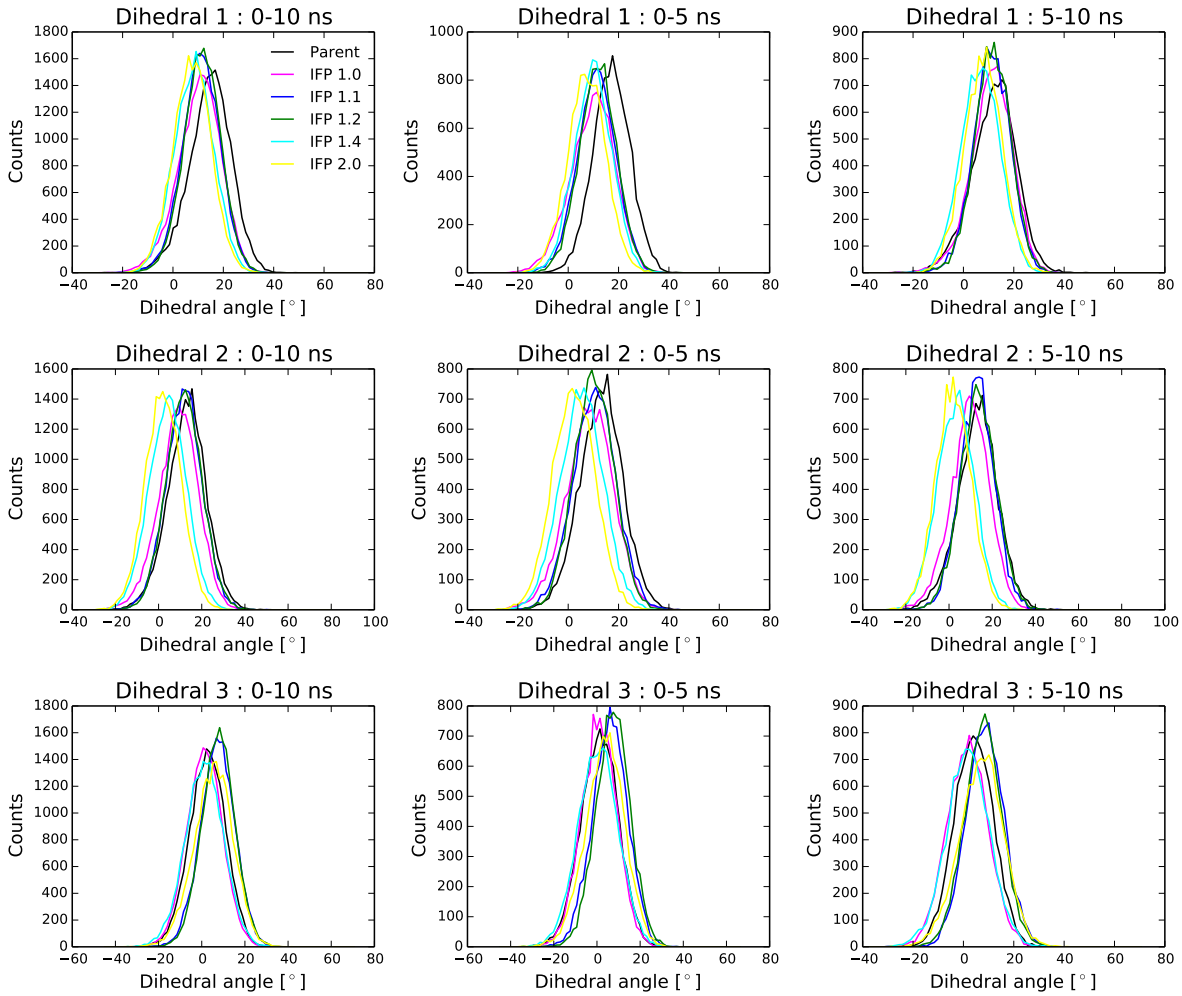


Figure S9: Statistical distributions of dihedral angles 1–3 of the chromophore for all proteins. The simulations were performed at neutral pH and with the chromophore in its fully protonated form. Three distributions are shown for each dihedral, calculated from the full (0–10 ns) simulation, and from the first (0–5 ns) and second (5–10 ns) halves. See Fig. 3 in the paper for the definitions of the dihedrals.

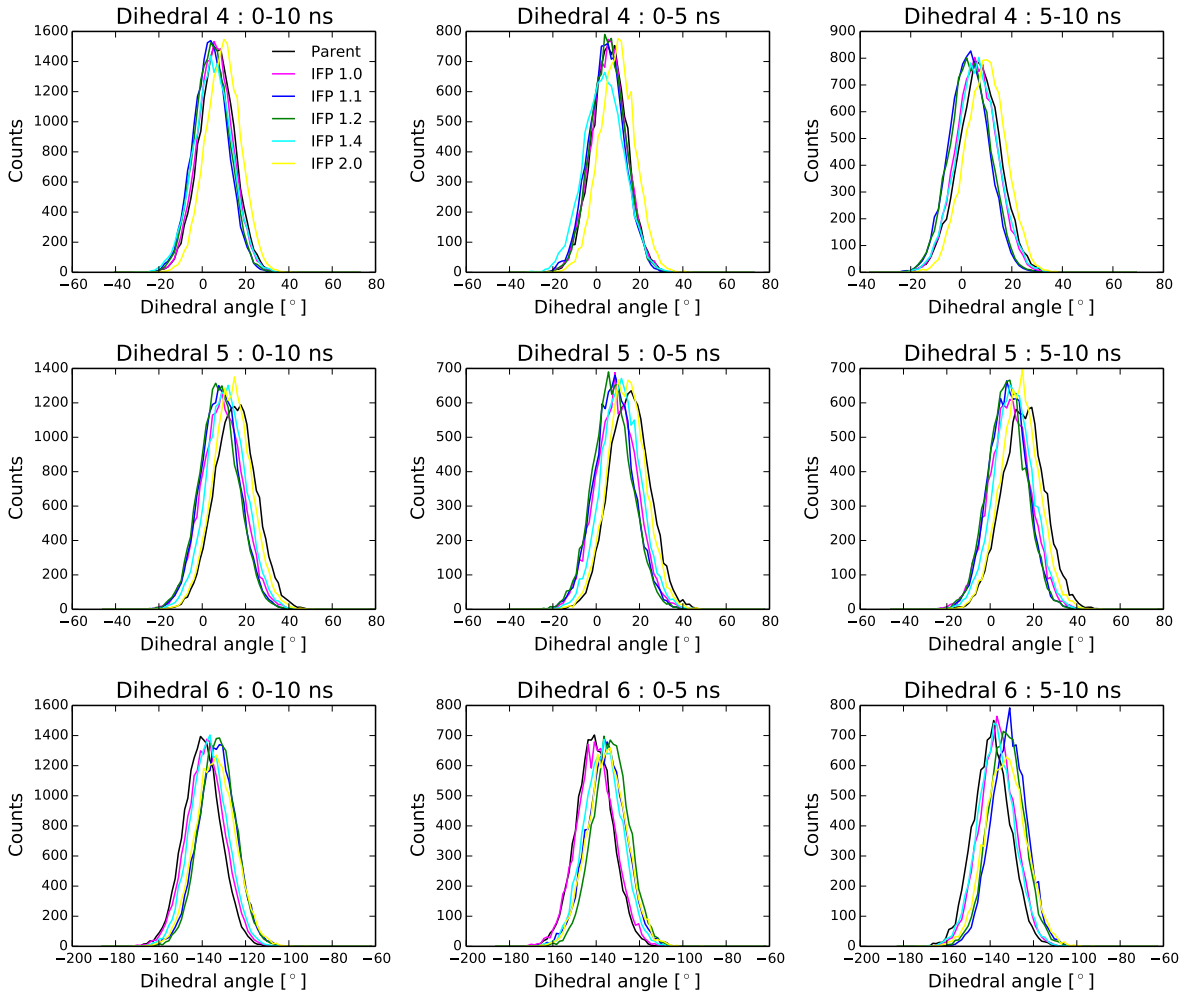


Figure S10: Statistical distributions of dihedral angles 4–6 of the chromophore for all proteins. The simulations were performed at neutral pH and with the chromophore in its fully protonated form. Three distributions are shown for each dihedral, calculated from the full (0–10 ns) simulation, and from the first (0–5 ns) and second (5–10 ns) halves. See Fig. 3 in the paper for the definitions of the dihedrals.

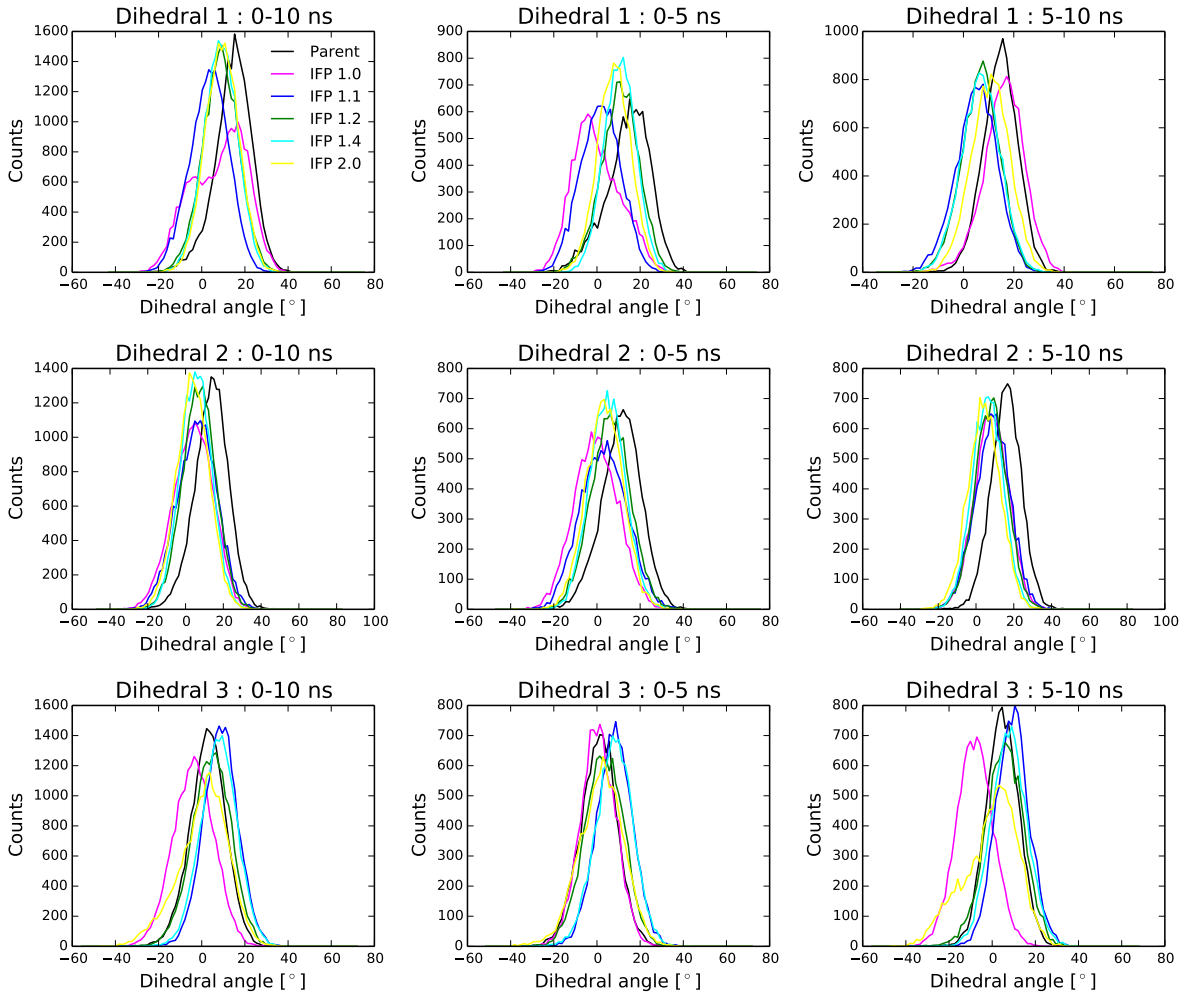


Figure S11: Statistical distributions of dihedral angles 1–3 of the chromophore for all proteins. The simulations were performed at low pH and with the chromophore in its fully protonated form. Three distributions are shown for each dihedral, calculated from the full (0–10 ns) simulation, and from the first (0–5 ns) and second (5–10 ns) halves. See Fig. 3 in the paper for the definitions of the dihedrals.

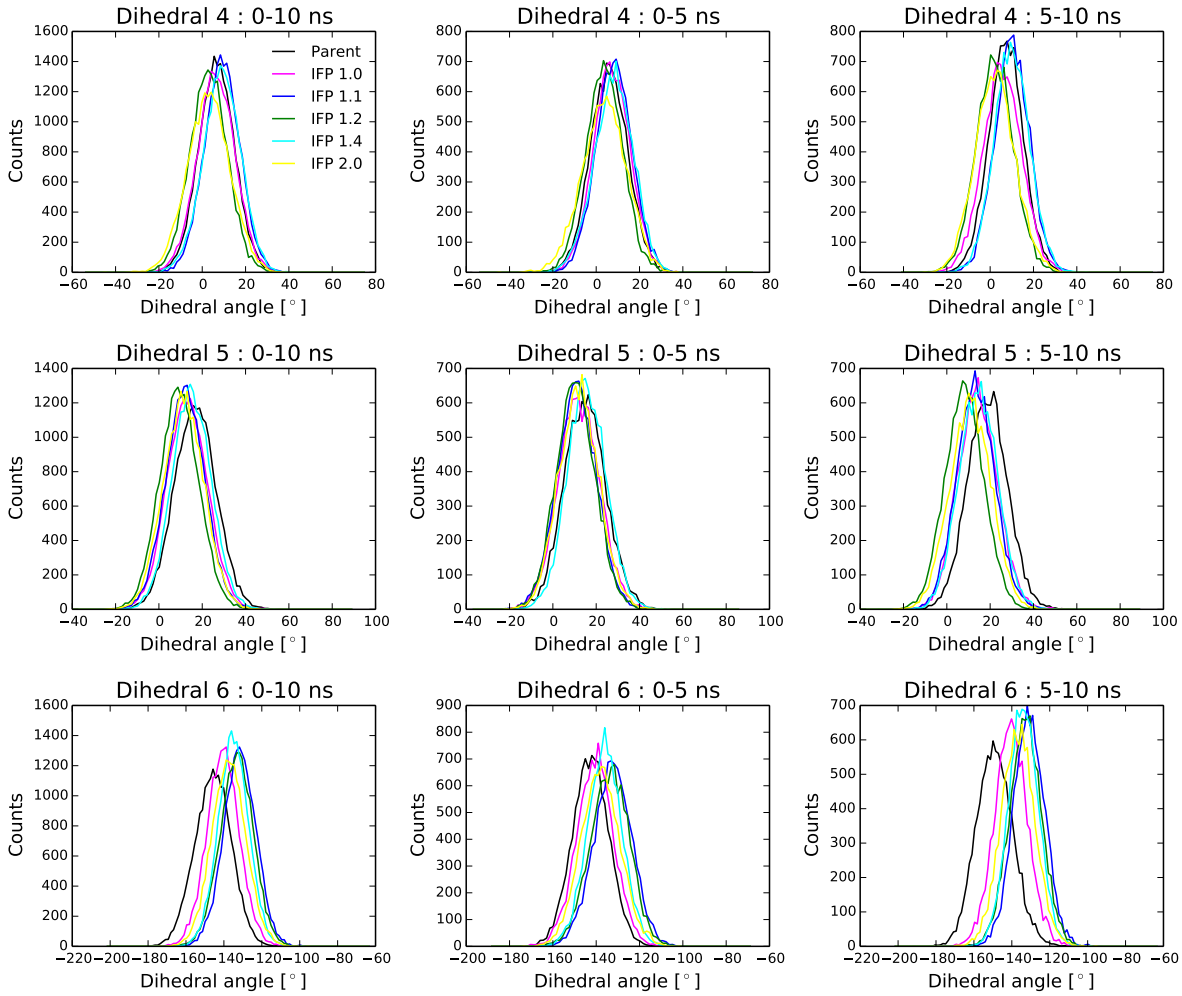


Figure S12: Statistical distributions of dihedral angles 4–6 of the chromophore for all proteins. The simulations were performed at low pH and with the chromophore in its fully protonated form. Three distributions are shown for each dihedral, calculated from the full (0–10 ns) simulation, and from the first (0–5 ns) and second (5–10 ns) halves. See Fig. 3 in the paper for the definitions of the dihedrals.

Table S1: Distributions of the six dihedral angles of the chromophore, collected from the molecular dynamics simulations over the time span of 10 ns. The simulations were done at neutral pH. There are three lines for each protein. These correspond to the full (0–10 ns), first half (0–5 ns), and second half (5–10 ns) of the simulations, respectively. The color codes to indicate which μ values are nearest planarity and which σ values are nearest zero, compared to the simulations at low pH, are green (0–10 ns), blue (0–5 ns), and red (5–10 ns), respectively. See the caption to Tab. 3 in the main text for further description and Fig. 3 for the numbering of the dihedrals.

Dihedral→	**1**		**2**		**3**		**4**		**5**		**6**	
Protein↓	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
*Dr*CBD	14.6	8.5	13.1	9.0	2.7	8.3	6.6	8.2	15.3	9.5	−140.0	8.5
	17.1	7.3	13.0	8.6	1.2	8.6	6.1	8.0	15.3	9.2	−140.9	8.4
	12.0	8.9	13.1	9.4	4.1	7.7	7.1	8.3	15.2	9.7	−139.2	8.5
IFP1.0	10.4	8.5	9.0	9.2	1.3	8.2	5.8	8.0	9.1	9.5	−138.5	8.8
	9.6	8.8	8.5	9.4	1.0	8.2	5.8	8.1	9.3	9.5	−140.3	8.9
	11.1	8.2	9.6	9.0	1.7	8.2	5.9	8.0	9.0	9.5	−136.7	8.3
IFP1.1	11.1	7.5	11.7	8.4	7.3	7.9	3.8	7.9	8.2	8.9	−133.4	8.8
	11.2	7.6	10.4	8.5	6.0	8.2	4.7	8.1	8.3	9.1	−135.4	9.0
	11.1	7.3	13.0	8.1	8.6	7.5	3.0	7.7	8.2	8.8	−131.4	8.2
IFP1.2	11.6	7.3	11.8	8.4	7.6	7.7	4.5	8.0	7.6	9.0	−133.0	8.3
	11.7	7.3	10.1	7.9	7.4	7.9	5.5	8.0	7.5	9.1	−133.4	8.5
	11.4	7.3	13.5	8.5	7.8	7.5	3.4	7.8	7.7	8.8	−132.7	8.2
IFP1.4	8.3	7.7	4.2	8.6	1.2	8.9	4.8	8.8	11.0	9.1	−137.0	8.6
	9.8	7.5	5.3	8.5	0.5	9.2	3.5	9.2	11.3	9.1	−136.8	8.7
	6.8	7.7	3.1	8.5	1.8	8.5	6.2	8.2	10.7	9.0	−137.3	8.6
IFP2.0	7.5	7.5	2.2	8.3	5.3	9.1	9.6	8.1	13.8	9.0	−134.0	9.2
	7.4	7.6	2.3	8.3	3.6	9.1	9.7	8.2	14.0	8.9	−135.0	9.0
	7.7	7.5	2.0	8.2	7.1	8.8	9.6	8.0	13.6	9.0	−132.9	9.3

Table S2: Distributions of the six dihedral angles of the chromophore, collected from the molecular dynamics simulations over the time span of 10 ns. The simulations were done at low pH. There are three lines for each protein. These correspond to the full (0–10 ns), first half (0–5 ns), and second half (5–10 ns) of the simulation, respectively. The color codes to indicate which μ values are nearest planarity and which σ values are nearest zero, compared to the simulations at neutral pH, are green (0–10 ns), blue (0–5 ns), and red (5–10 ns), respectively. See the caption to Tab. 3 in the main text for further description and Fig. 3 for the numbering of the dihedrals.

Dihedral→	**1**		**2**		**3**		**4**		**5**		**6**	
Protein↓	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
*Dr*CBD	14.4	8.7	13.5	9.0	2.8	8.2	6.6	8.2	16.8	9.6	−145.7	11.1
	14.2	10.2	11.2	9.2	0.7	8.5	6.0	8.5	14.8	9.6	−142.6	8.4
	14.6	6.9	15.9	8.1	4.8	7.5	7.3	7.9	18.7	9.2	−148.8	12.5
IFP1.0	7.9	12.5	3.9	10.4	−3.8	9.4	6.4	8.7	13.1	9.3	−140.4	8.7
	0.0	10.9	−0.4	10.1	0.5	8.2	7.2	8.4	12.1	9.5	−140.5	8.6
	15.9	8.2	8.2	8.9	−8.1	8.5	5.6	8.9	14.2	9.1	−140.4	8.9
IFP1.1	3.4	8.8	5.9	10.4	8.7	8.0	9.1	8.2	12.2	9.0	−131.7	8.6
	1.7	9.2	3.1	10.6	8.1	8.3	8.2	8.3	11.0	9.1	−132.4	8.9
	5.2	8.1	8.7	9.3	9.4	7.7	10.1	7.9	13.4	8.8	−130.9	8.4
IFP1.2	8.7	8.4	7.0	8.9	4.4	9.0	3.0	8.6	9.4	9.0	−133.3	8.8
	10.6	8.7	6.2	9.2	3.2	9.0	3.8	8.6	10.6	8.9	−133.8	9.0
	6.8	7.5	7.8	8.5	5.6	8.9	2.2	8.6	8.2	8.9	−132.8	8.6
IFP1.4	8.8	7.7	5.4	8.2	7.8	8.5	9.0	8.6	15.1	9.0	−135.0	8.1
	10.6	7.4	4.8	8.2	7.9	8.6	8.0	9.0	15.6	9.0	−136.1	8.1
	6.9	7.6	6.0	8.2	7.8	8.3	9.9	8.1	14.5	8.9	−133.9	8.0
IFP2.0	9.3	7.8	3.9	8.6	0.6	11.5	3.0	9.7	11.5	9.3	−137.4	9.1
	8.0	7.7	3.5	8.6	1.9	10.3	3.7	10.2	12.0	9.2	−138.0	9.1
	10.7	7.7	4.3	8.7	−0.8	12.5	2.3	9.1	10.9	9.4	−136.8	9.0

S1 Statistical distribution of the important dihedral angles of the chromophore.

In the present section, we analyze in detail the intrinsic dynamics of the protein-bound biliverdin in different IFPs by focussing on the six dihedral angles of the chromophore located between the pyrrole rings defined in Fig. 3.

In the main text, Fig. 4 and Tab. 3 show the plotted distributions and fitted normal distribution parameters, μ and σ, for each of the proteins calculated from the last 5 ns of the 10 ns molecular dynamics simulations. By contrast, data from the complete simulations, together with separate analyses of the first and second halves, are given in Figs. S9–S12 and Tabs. S1 and S2.

Initially we consider briefly the differences between the dihedral distributions from the first and second halves of the molecular dynamics simulations. Overall, a perusal of Figs. S9–S12 and Tabs. S1 and S2 indicates that the distributions between the two halves are consistent, and that any trends that are present, such as for dihedral angles **1** and **2** (see below), are maintained. The principle exception to this observation is for the simulation of IFP 1.0 at low pH. At the start of the simulation, dihedral angles **1** and **3** are close to planarity, but they undergo a concerted shift to non-planarity in the second half of the simulation, adopting means of 15.9 and $-8.1°$, respectively.

In what follows we shall concentrate on the distributions from the last 5 ns of the molecular dynamics simulations, and we start with the simulations performed at neutral pH. As seen on Fig. 4a in the main text, the differences between the distributions for the two dihedral angles located between the chromophore rings A and B, labeled **1** and **2**, are quite pronounced. Ring A is adjacent to the covalent binding site of the chromophore to Cys24. In *Dr*CBD, dihedral **1** angle shows a mean value of 12.0°, indicating a noticeable distortion of planarity (see Tab. 3 for the comparison of the calculated statistical parameters of each dihedral in all proteins). In IFP1.0, which is only a weakly fluores-

cent protein, dihedral **1** is only slightly shifted towards planarity by 1.0° in comparison to the parent protein. In the most fluorescent proteins, the mean value of dihedral **1** is further reduced to 6.8° in IFP1.4 and 7.7° in IFP2.0, which indicates a dihedral about 40% more planar than in the parent protein. For the generated models of IFP1.1 and IFP1.2, the mean values were calculated to be 11.1° and 11.4°, respectively. While the mean parameter of dihedral **1** is shifted depending on the studied protein, the simulations show only small decreases in the variance parameter. For example, the flexibility of the chromophore region involving dihedral **1** was calculated to be 8.9 and 7.5 for *Dr*CBD and IFP2.0, respectively. Thus, the flexibility of the chromophore fragment involving dihedral **1** seems to be comparable in all proteins. As indicated in the main manuscript, the observed differences in the planarity can be related to the migration of water molecules in the vicinity of the chromophore.

Dihedral angle **2** displays the most significant differences in planarity, depending on the protein. The mean parameter for dihedral **2** was calculated to be 13.1° in *Dr*CBD and is already reduced to 9.6° in IFP1.0. Dihedral **2** shows a significant improvement of planarity in IFP1.4 and to a greater extent in IFP2.0. Namely, the μ parameter was calculated to be as low as 3.1° in IFP1.4 and 2.0° in IFP2.0. For IFP1.1 and IFP1.2, dihedral **2** has similar values to that of the parent protein (μ = 13.0 and 13.5°, repectively). However, the models for these proteins were generated by substitution of the relevant residues in *Dr*CBD, since the corresponding crystal structures were not available. Hence, one has to allow for a wider margin of uncertainty when considering these two proteins. Similarly to dihedral **1**, the rigidity of this dihedral is only slightly reduced during the transition from the parent protein (*Dr*CBD) to the other IFPs. The variance parameter was calculated to be σ=9.4 in *Dr*CBD and changes to only σ=8.2 in IFP2.0.

Compared to dihedrals **1** and **2**, the dynamics of the other four dihedrals show lesser dependence on the protein. Dihedral **3**, which is part of the link between rings B and C, has a mean value of 4.1° in *Dr*CBD. The value is reduced to 1.7° in IFP1.0 and 1.8° in

IFP1.4, but eventually becomes 7.1° in IFP2.0. For the models of the intermediate proteins, dihedral **3** was calculated to be even less planar than in the other proteins (μ=8.6° and μ=7.8° for IFP1.1 and IFP1.2, respectively). Dihedral **4** shows a rather irregular dynamics and its out-of-plane character generally increases on going from *Dr*CBD to IFP2.0. A similar behavior is seen in the case of dihedral **5**, where there is a considerable tilt between rings C and D. The μ parameter of dihedral **5** was calculated to be 15.2° in *Dr*CBD and 13.6° in IFP2.0, despite some improvement of planarity in the intermediate proteins. Finally, dihedral **6** shows a similar behavior to dihedrals **4** and **5** as its planarity decreases from μ=−139.2° in the parent protein to μ=−132.9° in IFP2.0, whereas its rigidity again does not change much. A somewhat similar trend in the dynamics of dihedral **6** has been reported in the earlier study on phytochromes, where a greater tilt was observed for the mutant protein.[4]

From the simulations performed at neutral pH, we note that the flexibility of all dihedrals remains similar regardless of the protein. Namely, the σ parameter indicating the flexibility of the particular fragment of the chromophore was always calculated to be within the range of σ=7.3 to σ=9.7. Fragments of the chromophore involving dihedrals **3** to **6** are located closer to the center of the protein than the fragments involving dihedrals **1** and **2**.

We next discuss the simulations performed at low pH. Protonation of different groups in the protein may disfavor the planar geometry of the chromophore, leading to the eventual loss of fluorescence. Similarly to the situation at neutral pH, dihedral **1** is significantly distorted in *Dr*CBD (μ=14.6°). On going from the parent protein to the fluorescent mutants, the improvement of planarity is not as pronounced as at neutral pH. The two most fluorescent proteins show dihedral **1** more tilted (μ=6.9° for IFP1.4 and μ=10.7° for IFP2.0) than in the intermediate proteins (for example, μ was calculated to be 5.2° for IFP1.1).

Dihedral **2** has a considerable tilt of μ=15.9° in *Dr*CBD but this reduces markedly

to μ=8.2° in IFP1.0. In contrast to the situation at neutral pH, the mean values of dihedral **2** for IFP1.1 and IFP1.2 are similar to those of IFP1.0, and were calculated to be 8.7° and 7.8°, respectively. In the most fluorescent proteins, the tilt is further slightly reduced to 6.0° for IFP1.4 and 4.3° in IFP2.0. The orientation of rings A and B of the chromophore seems more planar at neutral pH if one considers the parent protein and the most fluorescent proteins. However, the trend is reversed for the intermediate proteins, in which dihedral **2** is less tilted at low pH. For both dihedrals, there do not appear to be significant differences in flexibility between low and high pH.

At lower pH, the dynamic behavior of the remaining four dihedrals is again less systematic, reinforcing the idea that the planarity of dihedrals **1** and **2** is most critical in improving the fluorescence of the chromophore. The flexibility of dihedral **3** varies to some extent depending on the protein, but does not show the same clear trend that is visible for dihedrals **1** and **2**. Interestingly, the overall flexibility of dihedral **3** is increased in the situation at lower pH, with the exception of *Dr*CBD and IFP1.4, where the difference between neutral and lower pH is only slight. For example, the calculated variance parameter for IFP2.0 is σ=8.8 at neutral pH and σ=12.5 at low pH. The trend for dihedral **4** is somewhat mixed with it being more tilted at low pH for the parent protein, and for IFP1.1 and IFP1.4, and also more flexible in four out of six cases. Dihedral **5** shows a rather significant displacement from planarity, which is related to the orientation of ring D of the chromophore. The tilt is also greater now than in the situation at neutral pH, with the exception of IFP2.0. For *Dr*CBD, the dihedral's mean was calculated to be 15.2° at neutral pH and 18.7° at low pH. For IFP2.0, this parameter is 13.6 and 10.9 for neutral and low pH, respectively. In both acidities, there is a visible trend towards planarity for dihedral **5** on going from the parent protein towards more fluorescent proteins. Finally, the dynamics of dihedral **6** is reversed compared to the case at pH=7. With the exception of IFP1.1 and IFP1.4, dihedral **6** at lower pH was predicted to be closer to its planar conformation defined as −180°.

In summary, the performed simulations suggest that there is a connection between the introduced mutations and the dynamics of the chromophore. Dihedral angles **1** and **2** are shown to be generally the most distorted. Nevertheless, because of the additional mutations, the two dihedrals can be gradually planarized and restrained to some extent, especially in IFP1.4 and IFP2.0. Similar but less pronounced trends are visible for the other dihedrals. The mutations increasingly enforce the planarity of the chromophore and reduce motion around the conjugated bonds important for fluorescence. We note that not all mutations are beneficial for the improved planarity and rigidity of the chromophore. Some of the effects observed for IFP1.1 and IFP1.2 may result from the fact that the models representing these proteins were generated by modifying the structure of the parent protein and thus may not be fully reliable. Although the dynamics of the chromophore are somewhat altered after changing the pH from neutral to low, it is difficult to definitely pinpoint the factors responsible for the reported loss of fluorescence. Concerning the planarity of the chromophore, at low pH some of the dihedrals become more distorted compared to the situation at neutral pH, most importantly **1** and **2**. Interestingly, in four out of the six cases the flexibility of the dihedrals is increased at low pH (Tab. 3), although the observable differences in comparison to the situation at neutral pH are rather small. The overall decrease of the chromophore's rigidity at low pH may probably lead to some reduction of fluorescence, but we are not convinced that this can explain the complete loss of the fluorescence properties, as seen from the measurements performed for IFP1.4[5] and the other proteins (Fig. S3). We propose that, in addition to the increased flexibility of the chromophore, as seen in our molecular dynamics simulations, another mechanism might be involved in the reduction of fluorescence at low pH, for example quenching upon chromophore deprotonation or denaturation of the protein.

S2 Calculations of different protonation forms of biliverdin in the gas phase.

The key difficulty in the computational treatment of nonstandard titratable groups, such as the biliverdin chromophore, is that the pK_a values of these groups in aqueous solution are often not known. For a known pK_a in aqueous solution, i.e. pK_a^{aq}, one can calculate its shift while moving the protonatable group from the solvent phase to the protein. The calculation is usually done by using the Poisson-Boltzmann electrostatic model, like the one employed in the present study.[6] These pK_a^{aq} values, if not available from experiment, can be estimated computationally by using a variety of methods.[7,8] However, reliable calculations of pK_a^{aq} values represent a rather difficult task. In the present study, we took a different approach, in which we simply fixed the protonation state of the protein-bound biliverdin to its either fully protonated or B-deprotonated form during the electrostatic calculations.

The choice of the B-deprotonated form was based on the comparison of the vacuum energies of all deprotonated forms of the chromophore, calculated using the BP density functional theory method and the TZV(2d) basis sets (TZV(p) for hydrogens). The B-deprotonated form was always the lowest in energy and the A-, C- and D-deprotonated forms were higher in energy by, respectively, 1.8, 5.9 and 29.1 $^{kcal}/_{mol}$. Thus, ring B was assumed to be the most likely deprotonation site. From our vacuum calculations, we note that the conformation of the isolated chromophore is dependent on the protonation of its pyrrole rings (Fig. S13). The A-, B- and C-deprotonated forms display a relatively planar geometry compared to the D- deprotonated or fully protonated forms. Removal of a pyrrole proton from ring A, B or C reduces the mutual repulsion between the rings, which is due to the confined inner protons. It also introduces a hydrogen-bond acceptor, to which intramolecular hydrogen bonds can be formed, further stabilizing the planar structure of the chromophore. For example, removal of hydrogen B enables intramolecular hydrogen

bonds between the pyrrole nitrogen of ring B and the hydrogens from rings A and C. By contrast, the geometries of the D-deprotonated and fully protonated forms show three positively charged hydrogens on rings A, B and C in rather close contact. Thus, in the optimized geometry of the fully protonated chromophore, the closest distance, of 2.2 Å is between the A and B pyrrole hydrogens, where the distances between hydrogens A-C and B-C are 2.5 Å and 2.3 Å, respectively. Since fluorescence is due to the conjugated system of π-electrons, the chromophore should ideally have a planar and rigid structure to maximize its fluorescent properties. Therefore, it is conceivable that the B-deprotonated form of the chromophore should be more fluorescent than the fully protonated form. Indeed, the fluorescence of IFPs diminishes at low pH (Fig. S3), at which we presume the chromophore is fully protonated. Note that we excluded from our considerations the doubly, triply and fully deprotonated forms of the chromophore since they are energetically unfavorable and therefore unlikely to occur under biological conditions.

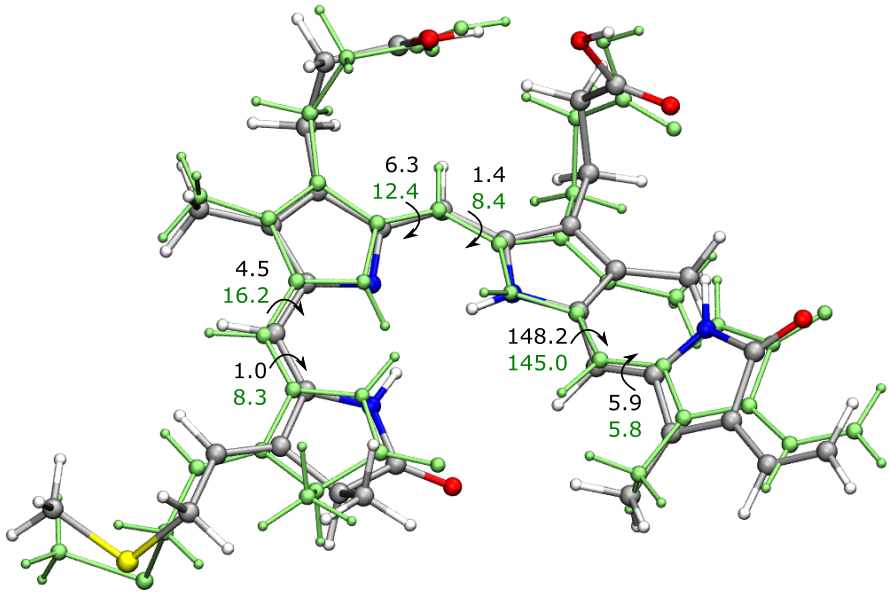


Figure S13: Superposition of the geometries of the fully protonated (shown in green) and B-deprotonated (shown in colors) forms of the free chromophore optimized at the BP/TZV(p)/TZV(2d) level. Cysteine linker was modeled during the optimizations as a thiomethyl group. Carboxylic groups at rings B and C were neutralized by adding protons. Arrows indicate central bonds of the six important dihedral angles that define the planarity of the chromophore. Black and green labels indicate the values of the dihedrals for the B-deprotonated and fully protonated forms, respectively. The pyrrole rings A-B-C in the fully protonated form show an out-of-plane geometry because of the repulsion of the inner protons. The inclusion of the pyrrole water in the vacuum models does not change the overall picture.

S3 Derivation of parameters for the deprotonated chromophore.

We refer to the biliverdin chromophore deprotonated at the B pyrrole ring simply as the B-deprotonated chromophore. The missing molecular mechanics parameters for the B-deprotonated chromophore were derived starting from the available parameters of the fully protonated chromophore.[9,10] All quantum chemical calculations were done in ORCA[11] using the BP density functional theory method and the TZV(p) and TZV(2d) basis sets for hydrogens and heavy atoms, respectively. In the first step, we prepared geometries of the isolated chromophore in its five protonation forms, namely the fully protonated form and the A-, B-, C-, and D-deprotonated form. The fully protonated form had a total charge of +1 and the other forms were neutral. The geometries of all forms were subsequently optimized in vacuum. Frequency analyses were performed on top of the optimized geometries of each form to ensure that the geometries were in their energy minima. In further calculations, only the fully protonated and the B-deprotonated forms were considered. The atomic charges for the isolated chromophore, both in its fully protonated and B-deprotonated form, were calculated using the CHELPG method.[12] Differences in the CHELPG charges were calculated between the corresponding atoms of the two forms. These differences were subsequently added to the existing MM-charges of the fully protonated form to obtain the MM-charges charges for the B-deprotonated form. The missing force field parameters for the four dihedral angles between rings A, B and C were derived following a two-step procedure. First, the complete chromophore was truncated to two separate molecular fragments, one comprising the rings A and B, and the other one comprising the rings B and C. The carboxylic groups at rings B and C were removed from the fragments to prevent possible sterical clashes during the consecutive scans. Points of truncation were saturated with hydrogen atoms. A relaxed potential energy surface scan was performed in steps of 5° along the central bond of each of the four dihedrals.

The scans were done both for the fully protonated form and the B-deprotonated form of the chromophore, using both the QC- and MM-potentials. The QC- and MM-scans were done in ORCA and pDynamo, respectively. Second, the dihedral parameters for the B-deprotonated form were manually tuned and the MM-scans were redone as long as a satisfactory fitting of the MM- and QC-generated energy profiles was achieved. The obtained fits for the B-deprotonated form show good agreement between the QC-profiles and MM-profiles (see Fig. S14).

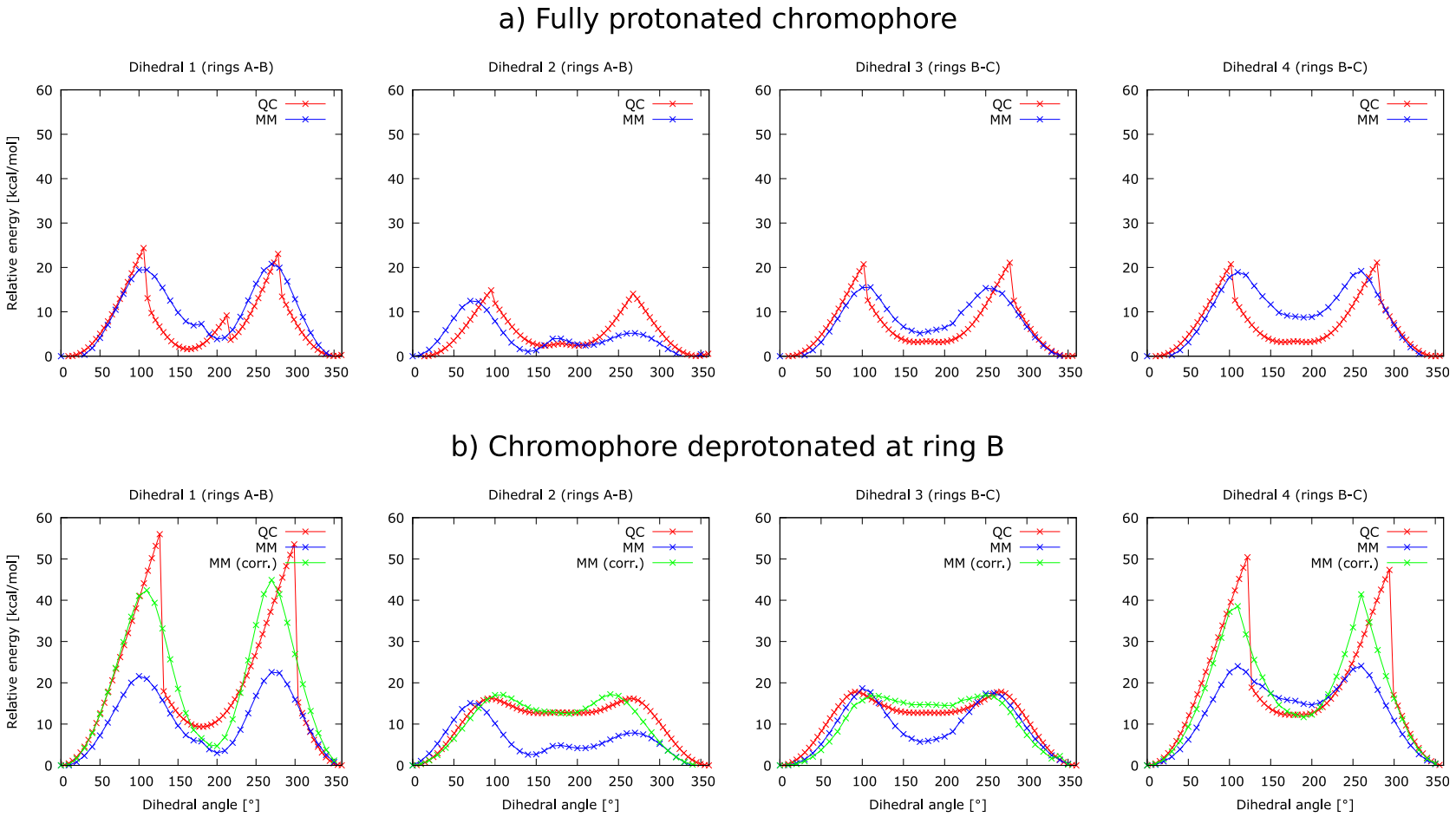


Figure S14: Energy profiles calculated for rotations around the central bond of each of the four dihedral angles of the chromophore. Each dihedral involves the B pyrrole ring, which is the most likely deprotonation site of the chromophore. The rotations were done for the fully protonated chromophore (a), and for the chromophore deprotonated at ring B (b). Red and blue lines indicate rotations performed with QC- and MM-potentials, respectively. As the QC-potential, the BP density functional theory method and the TZV(p) basis set (TZV(2d) for hydrogens) were used. As the MM-potential, the CHARMM27 force field was used together with the parameters for the fully protonated chromophore taken from the literature.[9,10] Green lines indicate MM-energy profiles calculated using MM-parameters (charges, dihedral parameters) corrected for the B-deprotonated chromophore. The new MM-parameters were obtained by fitting the MM-profiles to the corresponding QC-profiles.

Figure S15: Topology file for the CHARMM force field to describe the biliverdin chromophore deprotonated at pyrrole ring B, based on the previous topology of the fully protonated chromophore.[9,10]

```
* B-deprotonated chromophore
*
  31    1

MASS    201     CPY1      12.011    C
MASS    202     CPY2      12.011    C
MASS    203     CPY3      12.011    C
MASS    204     CPY4      12.011    C
MASS    205     CPY5      12.011    C
MASS    206     CPY6      12.011    C
MASS    207     SE        32.060    S

AUTO  ANGLE DIHE

!=======================================
RESI BLB               -2.00
GROUP       ! methine bridge
ATOM    CAC     CPM    -0.034
ATOM    HAC     HA      0.110
GROUP       ! RING A
ATOM    C1C     C       0.267
ATOM    H2C     HA      0.045
ATOM    N_C     NR1    -0.381
ATOM    H_C     H       0.266
ATOM    C4C     CA      0.166
ATOM    C3C     CPY1   -0.140
ATOM    C2C     CT1     0.333
ATOM    O_C     O      -0.505
GROUP
ATOM    CHD     CPY3   -0.441
ATOM    HHD     HA      0.185
GROUP       ! RING B
ATOM    C1D     CPA     0.448
ATOM    N_D     NR1    -0.705
!
ATOM    C4D     CPA     0.465
ATOM    C3D     CPB    -0.207
ATOM    C2D     CPB    -0.002
GROUP
ATOM    CHA     CPM    -0.086
```

```
ATOM    HHA     HA      0.304
GROUP       ! RING C
ATOM    C1A     CPA     0.021
ATOM    N_A     NR1    -0.499
ATOM    H_A     H       0.335
ATOM    C2A     CPY4   -0.038
ATOM    C3A     CPB    -0.202
ATOM    C4A     CPA     0.432
GROUP       ! RING D
ATOM    C4B     C       0.467
ATOM    N_B     NR1    -0.579
ATOM    H_B     H       0.429
ATOM    C1B     CA      0.475
ATOM    C2B     CPY5   -0.132
ATOM    C3B     CPY6   -0.112
ATOM    O_B     O      -0.462
GROUP
ATOM    CHB     CPY2   -0.491
ATOM    HHB     HA      0.268
GROUP       ! VINYL RING D
ATOM    CAB     CE1    -0.299
ATOM    HAB     HE1     0.292
ATOM    CBB     CE2    -0.348
ATOM    HV1     HE2     0.179
ATOM    HV2     HE2     0.162
GROUP       ! CH3 RING A (Cys binding point)
ATOM    CBC     CT2    -0.107
ATOM    HL1     HA      0.087
ATOM    HL2     HA      0.051
GROUP       ! CH3 RING A (close to Cys)
ATOM    CMC     CT3    -0.173
ATOM    HE1     HA      0.040
ATOM    HE2     HA      0.042
ATOM    HE3     HA      0.032
GROUP       ! CH3 RING B
ATOM    CMD     CT3    -0.052
ATOM    HD1     HA     -0.001
ATOM    HD2     HA      0.011
ATOM    HD3     HA      0.012
GROUP       ! CH3 RING C
ATOM    CMA     CT3    -0.049
ATOM    HA1     HA      0.028
ATOM    HA2     HA      0.030
ATOM    HA3     HA      0.038
GROUP       ! CH3 RING D
ATOM    CMB     CT3    -0.039
```

```
ATOM    HB1     HA      0.045
ATOM    HB2     HA      0.069
ATOM    HB3     HA      0.036
GROUP       ! CH2_COO- RING B   (Heme)
ATOM    CBD     CT2    -0.28
ATOM    HO3     HA      0.09
ATOM    HO4     HA      0.09
ATOM    CGD     CC      0.62
ATOM    O2D     OC     -0.76
ATOM    O1D     OC     -0.76
GROUP       ! CH2 RING B        (Heme)
ATOM    CAD     CT2    -0.159
ATOM    HO7     HA      0.056
ATOM    HO8     HA      0.045
GROUP       ! CH2_COO- RING C   (Heme)
ATOM    CBA     CT2    -0.28
ATOM    HO1     HA      0.09
ATOM    HO2     HA      0.09
ATOM    CGA     CC      0.62
ATOM    O2A     OC     -0.76
ATOM    O1A     OC     -0.76
GROUP       ! CH2 RING C        (Heme)
ATOM    CAA     CT2    -0.137
ATOM    HO5     HA      0.052
ATOM    HO6     HA      0.057
!
BOND  CHA  HHA    CHA  C1A    CHA  C4D    CGA  O1A
BOND  N_A  H_A    N_A  C1A    N_A  C4A    N_B  C4B
BOND  C1A  C2A    C2A  C3A    C2A  CAA    C3B  C4B
BOND  C3A  C4A    C3A  CMA    C4A  CHB    CMB  HB3
BOND  CMA  HA1    CMA  HA2    CMA  HA3    CBB  HV2
BOND  CAA  HO5    CAA  HO6    CAA  CBA    CGA  O2A
BOND  CBA  HO1    CBA  HO2    CBA  CGA    CMC  HE3
BOND  CHB  C1B    N_B  H_B    N_B  C1B    CMC  HE2
BOND  C1B  C2B    C2B  C3B    C2B  CMB    CAC  CBC
BOND  C4B  O_B    CMB  HB1    CMB  HB2    C3B  CAB
BOND  CAB  HAB    CAB  CBB    CBB  HV1    CMD  HD3
BOND  N_C  H_C    N_C  C1C    N_C  C4C    CGD  O1D
BOND  C1C  C2C    C1C  O_C    C2C  H2C    CGD  O2D
BOND  C2C  C3C    C2C  CMC    CMC  HE1
BOND  C3C  C4C    C3C  CAC    CAC  HAC
BOND  CBC  HL1    CBC  HL2    CHB  HHB
BOND  C4C  CHD    CHD  HHD    CHD  C1D
BOND  N_D  C1D    N_D  C4D
BOND  C1D  C2D    C2D  C3D    C3D  C4D
BOND  C2D  CMD    CMD  HD1    CMD  HD2
```

```
BOND  C3D  CAD    CAD  HO7    CAD  HO8
BOND  CAD  CBD    CBD  HO3    CBD  HO4
BOND  CBD  CGD
!
IMPH  HHA  C4D  C1A  CHA    C1A  CHA  N_A  C2A    C2A  C1A  C3A  CAA
IMPH  C3A  C2A  C4A  CMA    C4A  C3A  N_A  CHB    C2B  N_B  C1B  CHB
IMPH  CGA  O1A  O2A  CBA    CHB  C4A  C1B  HHB    C4B  N_B  C3B  O_B
IMPH  C2B  C1B  C3B  CMB    C3B  CAB  C4B  C2B    C3C  N_C  C4C  CHD
IMPH  CAB  CBB  C3B  HAB    CBB  CAB  HV1  HV2    C3D  C2D  C4D  CAD
IMPH  C1C  N_C  C2C  O_C    CHD  C4C  C1D  HHD
IMPH  C3C  C2C  C4C  CAC    C3C  CBC  CAC  HAC
IMPH  C1D  C2D  N_D  CHD    C2D  C1D  C3D  CMD
IMPH  C4D  C3D  N_D  CHA    CGD  O2D  O1D  CBD
!
DONOR     H_C  N_C  H_A  N_A  H_B  N_B
ACCEPTOR  O_C  C1C  O1D  CGD  O2D  CGD  O1A  CGA  O2A  CGA  O_B  C4B

!=======================================
! Patch to link BV to a cysteine residue
PRES  ABV              0.00
GROUP
ATOM   CB    CT2     -0.11
ATOM   HB1   HA       0.09
ATOM   HB2   HA       0.09
ATOM   SG    SE      -0.07
GROUP
ATOM   C     C        0.51
ATOM   O     O       -0.51
DELETE  ATOM  1HG1
!
BOND    2CBC  1SG
ANGLE   1CB   1SG   2CBC  1SG   2CBC  2CAC
!
ANGLE   2HL1  2CBC  1SG   2HL2  2CBC  1SG
!
DIHE    1CB   1SG   2CBC  2CAC  1HB1  1CB   1SG   2CBC  1HB2  1CB  1SG  2CBC
DIHE    1SG   2CBC  2CAC  2C3C  1SG   2CBC  2CAC  2HAC
DIHE    1CB   1SG   2CBC  2HL1  1CB   1SG   2CBC  2HL2
END
```

Figure S16: Parameter file for the CHARMM force field describing the biliverdin chromophore deprotonated at pyrrole ring B, based on the previous parameters for the fully protonated chromophore.[9,10]

```
* B-deprotonated chromophore
*

BONDS
NR1     CPA      377.200    1.3817
NR1     C        260.000    1.4190
CT2     CPM      230.000    1.4830
NR1     CA       377.200    1.4117
SE      CT2      198.000    1.8180
CPY1    CA       305.000    1.4772
CT1     CPY1     230.000    1.4907
CPM     CPY1     360.000    1.3266
CPY2    CA       360.000    1.3800
HA      CPY2     367.600    1.0902
CPY2    CPA      360.000    1.4455
CPY3    CA       360.000    1.3676
HA      CPY3     367.600    1.0900
CPY3    CPA      360.000    1.4225
CPY4    CPA      299.800    1.4052
CPY4    CPB      340.700    1.4018
CT2     CPY4     230.000    1.4946
CPY5    CA       305.000    1.4620
CT3     CPY5     230.000    1.4850
CPY6    C        250.000    1.4711
CPY6    CE1      450.000    1.4290
CPY6    CPY5     305.000    1.3470

!=======================================
ANGLES
NR1   CPA   CPB      122.00       110.000
CPA   NR1   CPA      139.30       116.300
H     NR1   CPA       30.000      125.50      20.00     2.15000
CPM   CPA   NR1       88.000      131.800
NR1   C     O         80.000      118.000
C     NR1   H         34.000      123.000
NR1   C     CT1       20.000      112.500
HA    CT2   CPM       49.300      107.500
HA    CPM   CT2       34.500      110.10      22.53     2.17900
C     NR1   CA        50.000      133.50
H     NR1   CA        30.000      125.50      20.00     2.15000
```

```
CT2   SE    CT2       34.000       95.000
SE    CT2   CPM       58.000      112.500
SE    CT2   CT1       58.000      112.500
SE    CT2   HA        38.000      111.000
HA    CPM   CPY1      12.700      117.440
CT2   CPM   CPY1      45.800      117.490
NR1   CA    CPY1     122.00       115.000
CT1   CPY1  CPM       45.80       116.600
CT1   CPY1  CA        45.80       128.000
CPY1  CT1   C         52.00       113.700
HA    CT1   CPY1      49.30       107.500
CT3   CT1   CPY1      51.80       107.500
CPM   CPY1  CA        61.60       124.100
CPY2  CPA   NR1       88.00       112.390
CPY2  CPA   CPB       61.60       124.070
CPY2  CA    NR1       88.00       129.000
CA    CPY2  CPA       94.20       127.000
HA    CPY2  CPA       12.70       117.440
HA    CPY2  CA        12.70       117.440
CPY3  CA    NR1       88.00       124.390
CPY1  CA    CPY3      61.60       127.570
CA    CPY3  CPA       94.20       122.800
HA    CPY3  CPA       12.70       117.440
CPY3  CPA   NR1       88.00       124.390
CPY3  CPA   CPB       61.60       125.070
HA    CPY3  CA        12.70       117.440
CPM   CPA   CPY4      61.60       132.500
NR1   CPA   CPY4     122.00       111.540
CPB   CPY4  CPA       30.80       136.010
CT2   CPY4  CPB       65.00       126.750
CPY4  CPB   CPA       30.80       145.010
CT3   CPB   CPY4      65.00       126.750
CT2   CT2   CPY4      70.00       114.700
HA    CT2   CPY4      50.00       109.500
CT2   CPY4  CPA       65.00       126.740
NR1   CA    CPY5     122.00       112.400
CPY5  CA    CPY2      61.60       124.970
CT3   CPY5  CA        45.80       115.900
HA    CT3   CPY5      49.30       107.500
NR1   C     CPY6      20.00       109.500
CPY6  C     O         80.00       118.000
CT3   CPY5  CPY6      45.80       122.300
C     CPY6  CPY5      52.00       116.000
C     CPY6  CE1       70.00       114.490
CPY5  CPY6  CE1       70.00       126.740
HE1   CE1   CPY6      50.00       120.000
```

```
CE2   CE1   CPY6      40.00       117.600
CA    CPY5  CPY6      40.00       116.500

!========================================
DIHEDRALS
H       NR1     CPA     CPB        1.0000   2   180.00
H       NR1     CPA     CPM        1.0000   2   180.00
CPM     CPA     NR1     CPA       14.0000   2   180.00
CPA     NR1     CPA     CPB       14.0000   2   180.00
NR1     C       CT1     HA         0.1900   3     0.00
O       C       CT1     HA         0.0000   3   180.00
NR1     C       CT1     CT3        0.0000   1     0.00
H       NR1     C       CT1        2.5000   2   180.00
HA      CT2     CPM     HA         0.1600   3   100.00
H       NR1     C       O          2.5000   2   180.00
CA      NR1     C       O          2.7500   2   180.00
CA      NR1     C       CT1        2.7500   2   180.00
HA      CT3     CT1     C          0.0000   6     0.00
CPA     CPB     CPB     CT2       14.0000   2   180.00
CPA     CPB     CPB     CT3       14.0000   2   180.00
SE      CT2     CPM     CA         0.0000   6     0.00
CT1     CT2     SE      CT2        0.2400   1   180.00
HA      CT2     SE      CT2        0.2800   3     0.00
SE      CT2     CPM     HA         0.0100   3     0.00
HS      SE      CT2     CPM        0.2700   3     0.00
NR1     CA      CPY1    CPM        3.0000   2   180.00
CPM     CPY1    CT1     C          3.0000   3     0.00
HA      CT1     CPY1    CPM        0.0400   3     0.00
CT3     CT1     CPY1    CPM        3.1000   2   180.00
HA      CPM     CPY1    CT1        1.2000   2   180.00
C       NR1     CA      CPY1       0.0000   2     0.00
CA      CPY1    CT1     C          0.0400   3     0.00
HA      CT1     CPY1    CA         0.0400   3     0.00
NR1     C       CT1     CPY1       0.0000   1     0.00
NR1     CA      CPY1    CT1        0.1900   3     0.00
H       NR1     CA      CPY1       1.0000   2   180.00
CT3     CT1     CPY1    CA         3.1000   2   180.00
HA      CT2     CPM     CPY1       0.0000   3     0.00
O       C       CT1     CPY1       1.4000   1   180.00
CT2     CPM     CPY1    CT1        3.1000   2   180.00
CT1     CPY1    CA      CPM        3.1000   2   180.00
HA      CPM     CPY1    CA         0.0000   2     0.00
HA      CPM     CA      CPY1       0.0000   2     0.00
CPY2    CPA     NR1     CPA       14.0000   2   180.00
NR1     CPA     CPY2    HA         1.9000   2   180.00
H       NR1     CPA     CPY2       0.0000   2   180.00
```

CPB	CPA	CPY2	HA	0.0000	2	180.00
CPA	CPY2	CA	NR1	3.0000	2	180.00
C	NR1	CA	CPY2	0.0000	2	0.00
NR1	CA	CPY2	HA	0.0000	2	0.00
H	NR1	CA	CPY2	1.0000	2	180.00
NR1	CPA	CPY2	CA	0.0000	2	180.00
CPY3	CPA	NR1	CPA	14.0000	2	180.00
NR1	CPA	CPY3	HA	0.1900	3	0.00
H	NR1	CPA	CPY3	1.0000	2	180.00
CPB	CPA	CPY3	HA	0.0000	2	0.00
CPA	CPY3	CA	NR1	3.0000	2	180.00
C	NR1	CA	CPY3	0.0000	2	0.00
NR1	CA	CPY3	HA	0.0000	2	0.00
H	NR1	CA	CPY3	1.0000	2	180.00
NR1	CPA	CPY3	CA	3.0000	2	180.00
CPM	CPY1	CA	CPY3	3.1000	2	180.00
CPY1	CA	CPY3	HA	3.1000	2	180.00
CT1	CPY1	CA	CPY3	3.1000	2	180.00
CPM	CPA	CPY4	CPB	0.0000	2	0.00
CPM	CPA	CPY4	CT2	0.0000	2	0.00
CPA	CPY4	CPB	CT3	0.0000	2	180.00
CPA	CPY4	CT2	CT2	0.0000	6	0.00
CPA	CPY4	CT2	HA	0.0000	6	0.00
NR1	CPA	CPY4	CPB	0.0000	2	0.00
NR1	CPA	CPY4	CT2	0.0000	2	0.00
H	NR1	CPA	CPY4	0.0000	2	0.00
CPY4	CPA	NR1	CPA	14.0000	2	180.00
CPB	CPY4	CT2	CT2	0.0000	6	0.00
CPB	CPY4	CT2	HA	0.0000	6	0.00
CPA	CPY4	CPB	CPA	14.0000	2	180.00
CPA	CPB	CPY4	CT2	1.0000	2	180.00
CT3	CPB	CPY4	CT2	0.0000	2	180.00
C	NR1	CA	CPY5	0.0000	2	0.00
CA	CPY5	CT3	HA	0.0000	3	0.00
CPY5	CA	CPY2	HA	0.0000	2	0.00
CPY2	CA	CPY5	CT3	3.1000	2	180.00
NR1	CA	CPY5	CT3	3.0000	2	180.00
H	NR1	CA	CPY5	1.0000	2	180.00
CE1	CA	CPY5	CA	3.1000	2	180.00
CA	CPY5	CPY6	C	0.0400	3	0.00
CT3	CPY5	CPY6	C	14.0000	3	0.00
C	CPY6	CE1	HE1	0.0000	2	180.00
C	CPY6	CE1	CE2	1.0000	2	180.00
NR1	C	CPY6	CPY5	0.0000	1	0.00
NR1	C	CPY6	CE1	0.0000	1	0.00
NR1	CA	CPY5	CPY6	3.0000	2	180.00

```
H      NR1     C       CPY6       2.5000   2   180.00
CA     NR1     C       CPY6      14.0000   2   180.00
CE1    CPY6    CPY5    CA         3.1000   2   180.00
O      C       CPY6    CPY5       0.4000   1   180.00
CPY5   CPY6    CE1     HE1        1.0000   2   180.00
CPY6   CPY5    CA      CPY2       1.1000   2   180.00
CPY6   CPY5    CT3     HA         0.0000   3     0.00
CPY6   CE1     CE2     HE2        5.2000   2   180.00
O      C       CPY6    CE1        1.4000   1     0.00
CE1    CPY6    CPY5    CT3        3.1000   2   180.00
!
! multiple terms for dihedral optimization
! VINYL
CPY5   CPY6    CE1     CE2        0.90     1     0.00
CPY5   CPY6    CE1     CE2        2.35     2   180.00
CPY5   CPY6    CE1     CE2        0.50     3     0.00
!
! DB pyrrol-pyrrolinone ring A-B
CPY1   CA      CPY3    CPA        4.8      1      0.0
CPY1   CA      CPY3    CPA       16.0      2    178.0
CPY1   CA      CPY3    CPA        4.5      3    160.0
!
! EB pyrrol-pyrrolinone ring A-B
CA     CPY3    CPA     CPB      2.0        3    180.0
CA     CPY3    CPA     CPB      4.0        2    180.0
CA     CPY3    CPA     CPB      6.0        1      0.0
!
! pyrrol-pyrrol 1
CPY4   CPA     CPM     CPA       18.0      2    183.0
CPY4   CPA     CPM     CPA        4.5      1      2.0
CPY4   CPA     CPM     CPA        5.0      3    190.0
!
! pyrrol-pyrrol 2
CPB    CPA     CPM     CPA       1.5       3    180.0
CPB    CPA     CPM     CPA       5.0       2    180.0
CPB    CPA     CPM     CPA       5.5       1      0.0
!
! DB pyrrol-pyrrolinone ring C-D
CPA    CPY2    CA      CPY5      11.8      1      0.0
CPA    CPY2    CA      CPY5       3.0      2    190.0
!
! EB pyrrol-pyrrolinone ring C-D
CPB    CPA     CPY2    CA         2.0      3      0.0
CPB    CPA     CPY2    CA         4.4      2    138.0
CPB    CPA     CPY2    CA         2.1      1    243.0
!
```

```
! linking fragment
CT2    SE      CT2     CPM        0.7      3     10.00
CT2    SE      CT2     CPM        1.1      1     88.00
SE     CT2     CPM     CPY1       1.4      2     45.00
CT2    CPM     CPY1    CA         2.4      2    180.00
CT2    CPM     CPY1    CA        16.4      1      4.00

!=======================================
IMPROPER
CPA    CPB     NR1     CPM      200.0000   0      0.0000
HE2    HE2     CE1     CE2        3.0000   0      0.0000
CD     OB      OH1     CT2       96.0000   0      0.0000
CPB    CPB     CPA     CT2       90.0000   0      0.0000
CPB    CPA     CPB     CT2       90.0000   0      0.0000
CPY1   CT1     CA      CPM        0.0000   0      0.0000
CPY1   CT2     CPM     HA        29.4000   0    180.000
CPA    CPB     NR1     CPY2     140.0000   0      0.0000
CPY2   CPA     CA      HA        29.4000   0      0.0000
CPA    CPB     NR1     CPY3     145.000    0      0.0000
HA     CPA     CA      CPY3      29.4000   0      0.0000
CPY1   NR1     CA      CPY3     140.0000   0    180.00
CPA    CPM     NR1     CPY4      61.0000   0      0.0000
CPY4   CPA     CPB     CT2       64.0000   0      0.0000
CPB    CPY4    CPA     CT3       90.0000   0      0.0000
CPY5   NR1     CA      CPY2       7.0000   0    180.000
CPY5   C       CA      CE1       90.0000   0    180.000
CPY5   X       X       CT3        0.0000   0      0.0000
CPY6   CE1     C       CPY5     180.0000   0      0.0000
HE1    X       X       CE1        3.0000   0      0.0000

!=======================================
NONBONDED nbxmod 5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
CPY1     0.000000     -0.070000      1.992400
CPY2     0.000000     -0.090000      1.800000
CPY3     0.000000     -0.090000      1.800000
CPY4     0.000000     -0.090000      1.800000
CPY5     0.000000     -0.070000      1.992400
CPY6     0.000000     -0.070000      1.992400
SE       0.000000     -0.470000      2.200000

END
```

S4 Chromophore dynamics and migration of water molecules in the situation at neutral pH and with the B-deprotonated chromophore.

In this section, we examine a physiologically less likely situation, when the chromophore is deprotonated at the B pyrrole ring at neutral pH (see Fig. 3 in the main text for the ring definitions). Although the chromophore is likely fully protonated in the whole spectrum of pH, as suggested by earlier experiments, the analysis of the B-deprotonated chromophore may still provide some insights into the dynamics of IFPs. As discussed in Section S2 (also see Fig. S13), the unbound biliverdin adopts a more planar geometry in its B-deprotonated form, which may favor the enhancement of fluorescence.

Similarly to the situation with the fully protonated chromophore, the most significant differences between the proteins are visible for dihedrals **1** and **2** (see Fig. S17). The planarity of dihedral angle **1** is calculated to be 12.6° in *Dr*CBD and decreases to 4.1° in IFP2.0 (see Tab. S3). The rigidity of dihedral **2** does not show a clear trend, which is reminiscent of the simulation of the fully protonated chromophore. Interestingly, dihedral **1** is now consistently more planar and rigid in comparison to the case with the fully protonated chromophore (see Tab. 3 in the paper). Dihedral **2** shows a significant improvement of planarity on going from the parent protein (μ=17.9°) to IFP2.0 (μ=−0.9°), which is the most dramatic change in all simulations that we have performed. The fragments of the chromophore described by dihedral **2** display again a rather irregular rigidity depending on the protein. Nevertheless, these fragments appear most flexible in *Dr*CBD (σ=12.4) and most rigid in IFP2.0 (σ=8.4). Compared to the previous situation, dihedral **2** in the B-deprotonated chromophore is initially more distorted (μ=13.1° vs μ=17.9°), but shows a regular trend towards planarity with the increasing fluorescence. The rigidity of dihedral **2** appears comparable for both protonation states of biliverdin. The inner pyrrole

rings described by dihedral angles **3** and **4** show similar dynamics as before. Dihedral **3** is predicted to be slightly more planar and rigid with the B-deprotonated chromophore. Dihedral **5** connecting rings C and D is now more displaced out of planarity, in particular for the intermediate IFPs. For example, the mean value of dihedral **5** in IFP1.4 is calculated to be 11.0° and 16.1° for the fully protonated and B-deprotonated chromophore, respectively. The flexibility of dihedral **5** does not change much between different protonation forms of biliverdin. Lastly, dihedral **6** is slightly more displaced out of planarity with the B-deprotonated chromophore, whereas its flexibility remains similar.

We have also analyzed the mobility of water molecules in the neighborhood of the B-deprotonated chromophore (see Fig. S18). In *Dr*CBD, the pyrrole water is promptly displaced from its crystallographic position and shows reasonable mobility inside the cavity. The water forms transient hydrogen bonds to the nitrogen atom of the deprotonated ring B of the chromophore, or to the carbonyl carbon atom at ring A. At the same time, it maintains contact with other water molecules from the bulk solvent through a chain of hydrogen bonds. The closest of these waters can push the pyrrole water further towards the region between the rings, thus causing ring A to move out of planarity. After about 3 ns of simulation, the pyrrole water is displaced from its position and replaced by another water. Water replacements in the cavity recur during the rest of the simulation.

Unlike in *Dr*CBD, the pyrrole water in IFP1.0 is isolated from the solvent because His207 hinders access of solvent to the chromophore. His207 is ϵ-protonated and forms a hydrogen bond to the carbonyl oxygen of the A pyrrole ring. This hydrogen bond remains in place throughout the simulation, although the His207 side-chain displays some flexibility. We do not observe any water exchange between the cavity of the chromophore and the outer solvent; the pyrrole water is permanently locked between the rings. After the first 0.5 ns of simulation, the pyrrole water establishes two hydrogen bonds, one to the δ-nitrogen of His260 and another to the carbonyl oxygen at ring A. These bonds remain intact until the end of the simulation.

In IFP1.4, the orientation of the His207 side-chain is different than in IFP1.0. Namely, the δ-hydrogen now forms a hydrogen bond to the carbonyl oxygen of the protein backbone. The carbonyl oxygen at ring A is coordinated both by the pyrrole water and by solvent waters from the outer side. Some limited contact is possible between these solvent waters and the pyrrole water. The contact is improved after 8.5 ns of simulation when the His207 side chain rotates outwards and more solvent waters approach the cavity occupied by the pyrrole water. However, the exchange of water molecules, which was seen in *Dr*CBD, does not generally occur here. The transient opening of the cavity is followed by a prompt return of His207 to its previous conformation.

IFP2.0 is exceptional in the sense that His260 is calculated to be doubly protonated with the B-deprotonated chromophore at neutral pH (Fig. S8). The pyrrole water is initially shifted towards the solvent. The shift is likely triggered by the electrostatic repulsion from the additional δ-proton of His260 or for sterical reasons. No direct contacts are visible between the pyrrole water and solvent waters, which may be due to the exposure of the hydrophobic part of the Thr207 side-chain towards the solvent and a different network of hydrogen bonds. However, the pyrrole water is promptly exchanged during the simulation by another water coming from the vicinity of ring A. This new water molecule remains locked between His260 and the pyrrole rings until the end of the simulation. Except for the initial replacement of the pyrrole water, no further exchange of water molecules is observed throughout the simulation. Overall, the mobility of water in IFP2.0 is predicted to be somewhat higher than in IFP1.0 and IFP1.4, but nevertheless noticeably lower than in *Dr*CBD (see Fig. S18).

Comparison to the simulations with the fully protonated chromophore. We note that the overall picture of the simulations generally holds up regardless of the protonation state of the chromophore, with the exception of IFP2.0. The most intensive water exchange is observed in the case of *Dr*CBD for both protonation states of biliverdin. In IFP1.0 and IFP1.4, the inner pyrrole water is always locked in the cavity of the chro-

mophore and its contact with the solvent is hindered by His207. In IFP2.0, we observe a rather intensive water migration when the chromophore is fully protonated (see the discussion in the paper). However, this migration does not seem to displace the chromophore out of planarity, because Thr207 interacts only weakly with the chromophore. When the chromophore is in its B-deprotonated form, His260 is calculated to be fully protonated at neutral pH (see Fig. S8). In this situation, the dynamics of IFP2.0 resembles that of the other IFPs and the pyrrole water remains isolated from the solvent.

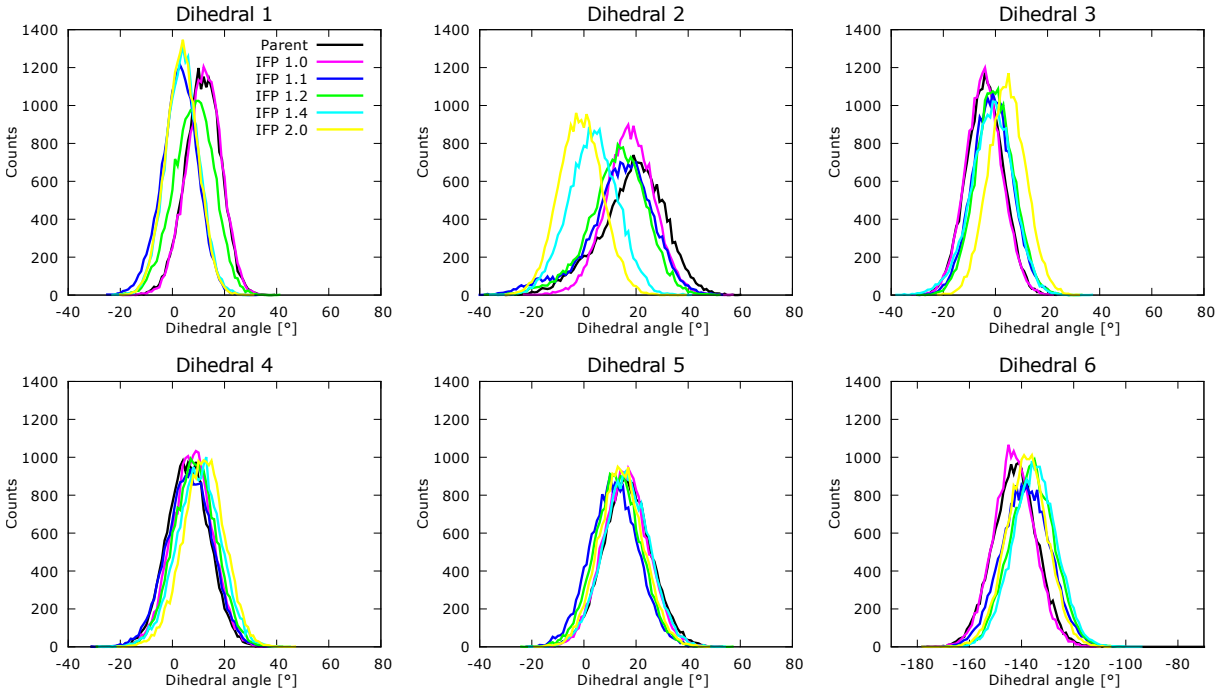


Figure S17: Statistical distributions of the six dihedral angles of the chromophore for all proteins. The simulations were performed at neutral pH and with the chromophore in its B-deprotonated form. See Fig. 3 in the paper for the definitions of the dihedrals.

Table S3: Distributions of the six dihedral angles of the chromophore when the chromophore is deprotonated at the B pyrrole ring at neutral pH. See Tab. 3 in the main text for a detailed description and Fig. 3 for the numbering of the dihedrals.

Dihedral→	**1**		**2**		**3**		**4**		**5**		**6**	
Protein↓	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
						Neutral pH						
*Dr*CBD	12.6	6.9	17.9	12.4	−3.9	7.2	6.1	8.0	16.4	9.3	−142.8	8.7
IFP1.0	11.9	6.8	17.8	9.4	−4.0	7.0	7.7	7.9	15.9	8.9	−142.7	8.1
IFP1.1	3.5	7.0	13.2	13.3	−1.0	7.6	6.7	8.6	12.1	9.3	−137.8	9.3
IFP1.2	8.4	7.7	12.9	11.4	−0.4	7.4	8.6	8.2	13.7	8.9	−136.1	8.6
IFP1.4	4.0	6.4	3.5	9.5	−1.4	8.3	10.4	8.5	16.1	9.0	−134.6	8.6
IFP2.0	4.1	6.2	−0.9	8.4	5.0	7.2	12.1	8.1	14.0	8.6	−137.6	8.1

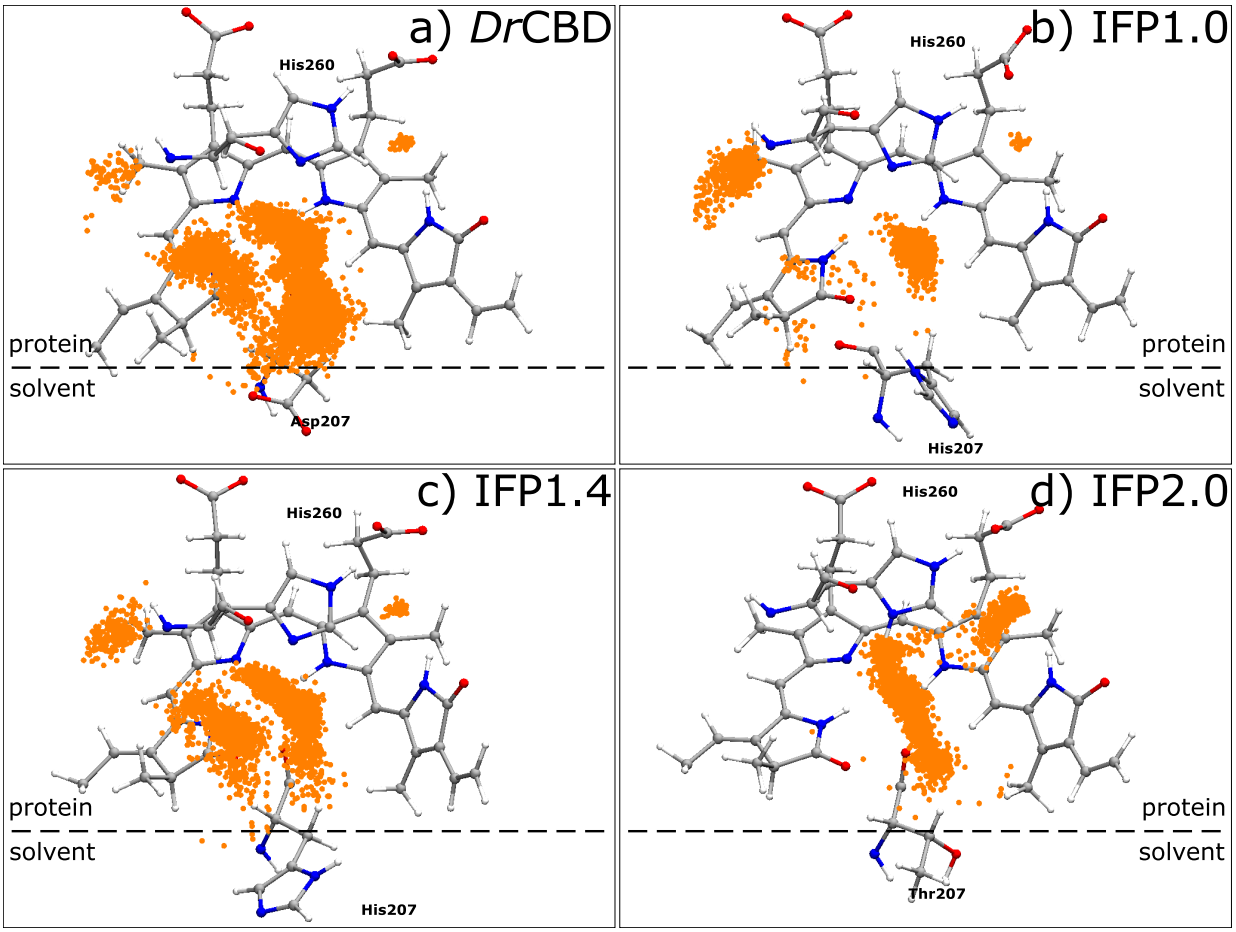


Figure S18: Mobility of water molecules inside the chromophore binding pocket in different proteins at pH=7 and with the B-deprotonated chromophore. See Fig. 6 in the main text for details.

Table S4: A list of animations showing the observed key events during the molecular dynamics simulations. In the animations, the chromophore and water molecules are shown in a ball-and-stick representation, while the surrounding residues, His260, Tyr263, Asp/His/Thr207 are shown in a licorice representation. His260 and Tyr260 are colored in lime and brown, respectively. For clarity, only water molecules within a radius of 7.0 Å from the center of the chromophore are shown. To obtain a smooth animation and eliminate minor thermal motion, the coordinates of the system for each simulation frame were averaged over 5 consecutive frames. The animations were generated for the simulations at neutral pH and with the fully protonated chromophore. Since IFP1.0 and IFP1.4 have a similar dynamic behavior, only an animation of the latter protein is included.

Animation file	Description
parent_wat_mobil.avi	Mobility of water in *Dr*CBD
ifp14_lock_pyr_wat.avi	Locked pyrrole water, rotation of His207 in IFP1.4
ifp20_wat_rel.avi	Release of the pyrrole water in IFP2.0

References

[1] Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G., Clustal W and Clustal X Version 2.0. *Bioinformatics*, **2007**, *23*, 2947–2948.

[2] Robert, X.; Gouet, P., Deciphering Key Features in Protein Structures with the New Endscript Server. *Nucleic Acids Res.*, **2014**, *42*, 320–324.

[3] The PyMOL Molecular Graphics System, Version 1.3r1, Schrodinger, LLC. 2010.

[4] Velazquez Escobar, F.; Hildebrandt, T.; Utesch, T.; Schmitt, F. J.; Seuffert, I.; Michael, N.; Schulz, C.; Mroginski, M. A.; Friedrich, T.; Hildebrandt, P., Structural Parameters Controlling the Fluorescence Properties of Phytochromes. *Biochemistry*, **2014**, *53*, 20–29.

[5] Shu, X.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y., Mammalian Expression of Infrared Fluorescent Proteins Engineered From a Bacterial Phytochrome. *Science*, **2009**, *324*, 804–807.

[6] Ullmann, G. M.; Knapp, E. W., Electrostatic Models for Computing Protonation and Redox Equilibria in Proteins. *Eur. Biophys. J.*, **1999**, *28*, 533–551.

[7] Liptak, M. D.; Shields, G. C., Accurate pK_a Calculations for Carboxylic Acids Using Complete Basis Set and Gaussian-N Models Combined with CPCM Continuum Solvation Methods. *J. Am. Chem. Soc.*, **2001**, *123*, 7314–7319.

[8] Muckerman, J. T.; Skone, J. H.; Ning, M.; Wasada-Tsutsui, Y., Toward the Accurate Calculation of pK_a Values in Water and Acetonitrile. *Biochim. Biophys. Acta, Bioenerg.*, **2013**, *1827*, 882–891.

[9] Kaminski, S.; Daminelli, G.; Mroginski, M. A., Molecular Dynamics Simulations of the Chromophore Binding Site of Deinococcus Radiodurans Bacteriophytochrome Using New Force Field Parameters for the Phytochromobilin Chromophore. *J. Phys. Chem. B*, **2009**, *113*, 945–958.

[10] Kaminski, S.; Mroginski, M., Molecular Dynamics of Phycocyanobilin Binding Bacteriophytochromes: A Detailed Study of Structural and Dynamic Properties. *J. Phys. Chem. B*, **2010**, *114*, 16677–16686.

[11] Neese, F., The ORCA Program System. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.*, **2012**, *2*, 73–78.

[12] Breneman, C. M.; Wiberg, K. B., Determining Atom-Centered Monopoles From Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.*, **1990**, *11*, 361–373.