Title

Inspecting fluctuation and coordination around chromophore inside green fluorescent protein from water to nonpolar solvent

Journal

Proteins Structure Function and Bioinformatics, 87(7)

ISSN

0887-3585

Authors

Dai, Liqiang
Zhang, Bo
Cui, Shuxun
et al.

Publication Date

2019-07-01

DOI

10.1002/prot.25676

Peer reviewed

Received: 23 October 2018 | Revised: 6 February 2019 | Accepted: 17 February 2019

DOI: 10.1002/prot.25676

PROTEINS WILEY

RESEARCH ARTICLE

Inspecting fluctuation and coordination around chromophore inside green fluorescent protein from water to nonpolar solvent

Liqiang Dai[1] | Bo Zhang[2] | Shuxun Cui[2] | Jin Yu[1]

[1]Complex System Research Division, Beijing Computational Science Research Center, Beijing, China

[2]Key Lab of Advanced Technologies of Materials (Ministry of Education), Southwest Jiaotong University, Chengdu, China

Correspondence
Jin Yu, Complex System Research Division, Beijing Computational Science Research Center, Beijing 100193, China.
Email: jinyu@csrc.ac.cn
and
Shuxun Cui, Key Lab of Advanced Technologies of Materials (Ministry of Education), Southwest Jiaotong University, Chengdu 610031, China.
Email: cuishuxun@swjtu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 11635002; NSAF, Grant/Award Number: U1530401; NSFC-Guangdong Joint Fund, Grant/Award Number: U1501501

Abstract

Green fluorescent protein (GFP) is a widely used biomarker that demands systematical rational approaches to its structure function redesign. In this work, we mainly utilized atomistic molecular dynamics simulations to inspect and visualize internal fluctuation and coordination around chromophore inside GFP, from water to nonpolar octane solvent. We found that GFP not only maintains its β-barrel structure well into the octane, but also sustains internal residue and water coordination to position the chromophore stably while suppress dihedral fluctuations of the chromophore, so that functional robustness of GFP is achieved. Our accompanied fluorescence microscope measurements accordingly confirmed the GFP functioning into the octane. Furthermore, we identified that crucial water sites inside GFP along with permeable pores on the β-barrel of the protein are largely preserved from the water to the octane solvent, which allows sufficiently fast exchanges of internal water with the bulk or with the water layer kept on the surface of the protein. By additionally pulling GFP from bulk water to octane, we suggest that the GFP function can be well maintained into the nonpolar solvent as long as, first, the protein does not denature in the nonpolar solvent nor across the polar-nonpolar solvent interface; second, a minimal set of water molecules are in accompany with the protein; third, the nonpolar solvent molecules may need to be large enough to be nonpermeable via the water pores on the β-barrel.

KEYWORDS

green fluorescent protein, molecular dynamics simulation, nonpolar solvent, water dynamics

1 | INTRODUCTION

Biomolecular structure is generally evolved in line with water solvent.[1–3] For example, DNA maintains a double-stranded helical structure in water but denatures quickly into nonpolar solvent.[4] Protein may become destabilized or maintain lower enzymatic activities being exposed to a hydrophobic environment, yet it may also become structurally strengthened being deprived of water.[5–7] To examine the role of water on protein structural and functional stabilities, it is of high interest to comparatively study the systems in between the water and nonpolar solvent deprived of bulk water.

In this work, we primarily used molecular dynamics (MD) simulation as a computational microscope[8] to reveal the structural stability and functional robustness of the green fluorescent protein (GFP) from water to nonpolar octane solvent (see Figure 1A). The GFP derived from jelly fish *Aequorea victoria* is widely used as a gene expression marker due to its effective bioluminescence in water-based environments.[12,13] The protein is formed by an 11-strand β-barrel, with a chromophore (S65-Y66-G67) formed inside the barrel via autocatalytic cyclization after protein folding.[14,15] The luminescent behavior of GFP relies sensitively on protein structure matrix along with water molecules.[11,16,17] How does the GFP structure along with luminescent properties react to nonpolar environments remained to be determined. Accompanied with the simulations, we also conducted an inverted fluorescent microscope measurement to verify the GFP luminescence into the octane solvent.

The GFP in the ground state maintains a neutral (A-form) or an anionic form (B-form). The fluorescence can be formed upon excitation of the ground state in the respective forms to the exited state (A*-form and B*-form), and via an intermediate state (I-form) and an excited-state proton transfer process that supports strong green fluorescence[12,18] (see Figure 1B *top*). Besides, nonfluorescent dark forms of GFP have also been found by nonradiative decay from the exited state of the

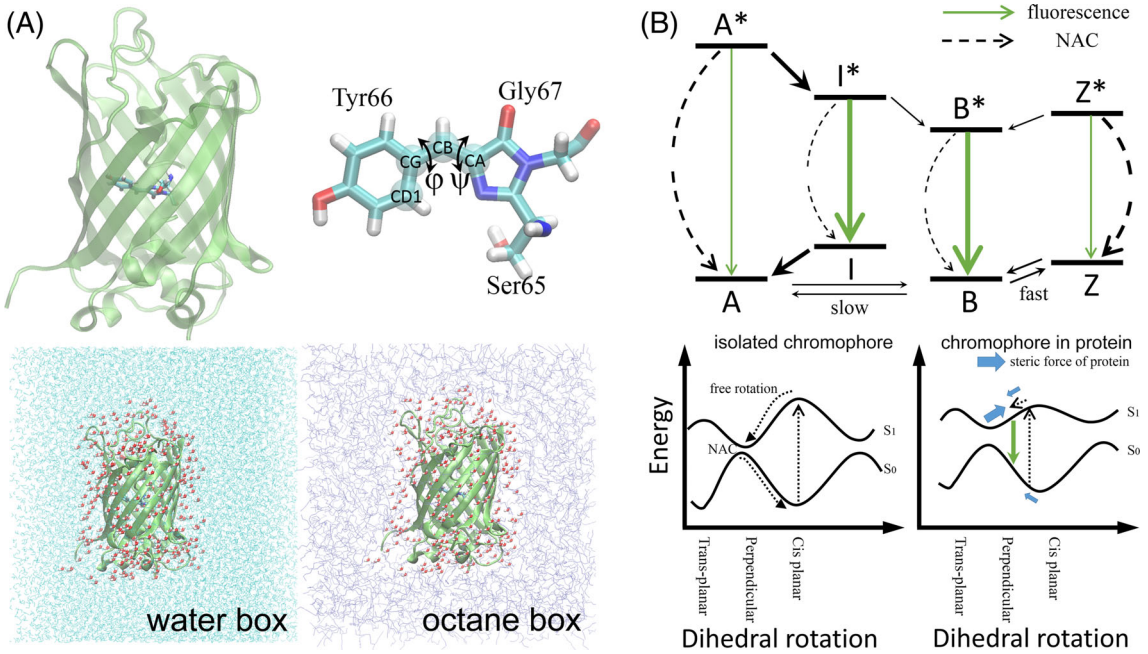

FIGURE 1 The simulation systems of green fluorescent protein (GFP) and fluorescence mechanisms. A, GFP, the chromophore, and the simulation systems. The β-barrel of GFP and the chromophore (Ser65-Tyr66-Gly67) are shown (*top*). The fluorescence requires suppression of dihedral fluctuations of the chromophore on φ (CD1-CG-CB-CA) and ψ (CG-CB-CA-N). The simulation systems included GFP (in the neutral form) in bulk water and non-polar octane (*bottom*). Crystal water molecules are explicitly shown accompanying the GFP structure (PDB: 2WUR).[9] B, The GFP fluorescence mechanisms inferred from previous quantum chemical calculations as being presented (from figure 2 of Weber et al.:[10] Copyright [1999] National Academy of Science) (*top*). A, B, I, and Z indicate the ground state neutral, anionic, intermediate, and Zwitterionic forms, respectively, with A*, B*, I*, and Z* their corresponding first exited states. The fluorescence and fluorescence-quenching nonadiabatic crossing (NAC) are indicated. The NAC usually relies on the dihedral rotation as being illustrated previously (from figure 3 with permission from Maddalo and Zimmer:[11] Copyright [2006] Wiley-Blackwell, Inc) (*bottom*). The ground state and first exited state of the chromophore in GFP are denoted by S_0 and S_1, respectively. Free dihedral rotation may lead to the NAC to quench the fluorescence, while the protein matrix may prevent the free dihedral rotation to inhibit the NAC [Color figure can be viewed at wileyonlinelibrary.com]

protein. In particular, the dark form may involve special dihedral configurations of the chromophore, which would be easily accessible if the dihedral flexibility is high (see Figure 1B *bottom*). In that case, the fluorescence-quenching nonadiabatic crossing (NAC) or nonradiative relaxation pathway between the exited state and ground state would dominate, so that the fluorescence diminishes.[10] Under the constraints of the GFP matrix, however, the chromophore dihedral would be prevented from free rotating, thus disables the fluorescence-quenching NAC.[11,17]

Accordingly, to examine whether functional robustness of GFP exists from the water to the octane condition, we conducted all-atom MD simulations, focused first on visualizing the dihedral fluctuations of the chromophore and then its coordination with surrounding amino acids and water molecules. It should be noted that for the ground state of the neutral form GFP we studied, the dihedral degree of the Y66 phenolic ring (φ in Figure 1A) is softly restrained, while the neighboring dihedral linked from Y66 toward S65 (ψ in Figure 1A) is heavily restrained and can hardly rotate,[19] hence we focused on the fluctuations of φ in this work.

We started by calibrating simulation systems on the chromophore from vacuum to bulk water and to the inside of GFP. Then we inspected GFP from water to the nonpolar octane solvent in the presence or absence of crystal water molecules captured in the high-resolution X-ray structure of GFP.[9] Discernable suppression of the chromophore dihedral fluctuations was found once the chromophore is placed into the interior of GFP. Essential water molecules seem to stabilize the chromophore inside the protein as well. Meanwhile, exchanges of waters inside and outside of the protein were frequently captured in the simulation via permeation pores on the β-barrel of GFP. When the protein is placed into the octane solvent, the protein structure seems well maintained, the chromophore stability is also sustained with the essential water molecules in accompany. In addition, the exchange of waters inside and outside on the surface of the protein was preserved in the octane solvent, while the octane molecules were nonpermeable to the inside of GFP. The fluorescence microscope measurements were then conducted to verify the luminescence of GFP into octane. Additionally, we investigated how GFP would react when it is directly pulled from water to the octane phase to assist further experimental investigations.

2 | MATERIALS AND METHODS

All the MD simulations were performed using the NAMD 2.10 software[20] with CHARMM22/27 force field.[21] The force field of chromophore was described additionally by the previous computational work.[22] The GFP structure was obtained from a high-resolution crystal structure (PDB: 2WUR).[9] For histidine protonation states, only atom ND1 was protonated except for His25, in which both ND1 and NE2 were protonated.[23]

For the GFP simulation in the water solvent, the protein structure was solvated with TIP3P water in a cubic box of 10 206 water molecules, and the minimum distance from the protein to the solvent boundary was 8 Å. We neutralized the systems with 21 Na^+ and 15 Cl^- to an ionic concentration of 0.15M. The total system consists of 35 194 atoms. The simulation was conducted under the periodic boundary condition. The van der Waals and short-range electrostatic interactions used a cutoff distance of 12 Å. The particle-mesh Ewald method was applied to deal with the long-range electrostatic interactions.[24] The solvated system was minimized with the steepest-descent algorithm, followed by 10 ps MD simulation under the canonical ensemble with a time step 1 fs. Then 5 ns equilibrium simulation was performed under the NVT ensemble with a time step of 1 fs, and position restraints on the heavy atoms of protein were imposed during the simulation. In the constant temperature simulations, the temperature was set to 300 K using the Langevin dynamics method.[25] After the constrained simulation, another constrained equilibrium simulation was performed under the NPT ensemble with a time step of 1 fs. The pressure was set at 1 bar using the Nosé-Hoover Langevin piston pressure control method.[26,27] Finally, an unconstrained equilibrium simulation or a steered molecular dynamics (SMD) simulation was carried out under the NPT ensemble with a time step of 1 fs. The equilibrium simulations of GFP in water lasted for 150 ns each. Additional 50 ns simulations were performed later to test system stabilities.

In the GFP simulation in the octane solvent, the protein structure is placed into a cubic box of octane made of 1337 octane molecules (110 Å × 81 Å × 101 Å). The total system consists of 39 302 atoms. Otherwise, the simulation setup is the same as the GFP-water simulation. The equilibrium simulations in octane also lasted for 150 ns each, with additional simulation running for the stability tests as well.

Note that for the isolated chromophore simulated in vacuum and in bulk water, the systems consist of 34 and 838 atoms, respectively. The GFP simulated in vacuum system consists of 3577 atoms. These small systems were equilibrated for 50 ns each. The equilibrations were then extended to over 1 μs in vacuum conditions to sample the *cis* to *trans* transition of the chromophore dihedral φ.

The pulling of GFP from water to octane simulation system consists of ∼90 000 atoms (88 Å × 78 Å × 150 Å), including 15 279 water molecules, 1560 octane molecules, and the GFP protein molecule, along with 27 sodium and 21 chloride ions added to make the system electrically neutral. After an energy minimization and constrained NVT and NPT simulations, a harmonic spring (with a force constant of 2 kcal/mol $Å^2$) was attached to the center of mass (COM) of the protein or to the C_α atom of a certain residue as specified below. The SMD simulations of GFP were conducted for three times under three different settings: In the first simulation, we pulled the COM of GFP from the water box toward the octane box at a speed of 1 Å/ns, for 70 ns in total; in the second SMD simulation, we dragged one end of the protein (C_α atom of ILE229) toward the octane box at a speed of 1 Å/ns, for 120 ns in total; in the last simulation, we dragged the center residue of the strand-3 (C_α atom of THR43) toward the octane box at a speed of 0.5 Å/ns, for 150 ns in total. An additional GFP pulling simulation (force implemented on C_α atom of ILE229 in the second setup) at a speed of 0.5 Å/ns was also conducted for control. During the pulling process, the COM of the octane box was constrained to avoid shifting.

3 | RESULTS

Below, we first show the dihedral degree of freedom of the chromophore when it is placed from vacuum to bulk water, and then to the inside of GFP, while GFP is placed in vacuum, in water, and in the nonpolar octane solvent, respectively. Next, we demonstrate water coordination with the chromophore along with water dynamics in and out of GFP, in the water, and the octane solvent conditions, respectively. Last, we show whether the system stabilities maintain when GFP is pulled from the water to the octane phase.

3.1 | The phenolic dihedral of the chromophore keeps restrained inside GFP from the water to the octane solvent condition

First, we placed the chromophore in vacuum, and measured its dihedral angle φ connecting the Y66 phenolic ring to the rest part. Under the current force field mimicking the ground state chromophore in the neutral form,[22] the dihedral angle φ has been softly restrained around an equilibrium *cis* planar position at φ = 0°[19] (see Supporting Information Figure S1A,B), oscillating from −43° to 44° (or around ∼0.3° ± 14.5°) within the 50-ns MD simulation (Figure 2A *top*). By extending the simulation to over 1 μs, 0° to ±180° (or *cis* to *trans*) transition of φ was identified (at ∼0.6 μs, see Supporting Information Figure S1C). φ becomes more flexible when we fully removed the soft restraint, and the φ distribution indeed spreads out within 50 ns (from *cis* to *trans*, see Figure 2A *bottom*, similar to that obtained previously[16]). The dihedral φ distribution does not show much suppression, however, around the *cis* configuration when the chromophore is placed inside GFP, still in vacuum (−0.5° ± 13.2° for Figure 2A *top*). Nevertheless, extending the simulation to 1 μs for the GFP in vacuum, φ remains fluctuating around 0° (*cis*) until the very end of the simulation (see Supporting Information Figure S1D). The flexible φ without the restrain also localizes without the *cis* to *trans* transition in the presence of GFP (see Figure 2A *bottom*). The results indicate that an enclosure of the chromophore into GFP suppresses the dihedral transition from *cis* to *trans*, consistent with the idea that the protein matrix prevents the free dihedral rotation.[11]

Slight suppression onto the dihedral fluctuations (around *cis*) is noticeable when GFP is placed in water (150 ns simulation). Note that if an isolated chromophore is simply placed into the bulk water, the φ distribution keeps almost the same as that in vacuum (−0.4° ± 14.7°). Enclosure of the chromophore into GFP solved in water, the dihedral fluctuation can be slightly quenched (7.3° ± 10.7°, see Figure 2B *top* with crystal water). That is to say, protein enclosure along with water participation together suppresses the dihedral fluctuations of the chromophore.

Then, we moved GFP into the nonpolar solvent of octane (150 ns simulation). The β-barrel structure of GFP seems maintained into the octane solvent (see Supporting Information Figure S2). The dihedral fluctuation suppression persists (−0.7° ± 10.7°, see Figure 2C *top* with crystal water). Note that on average the phenolic ring could deviate

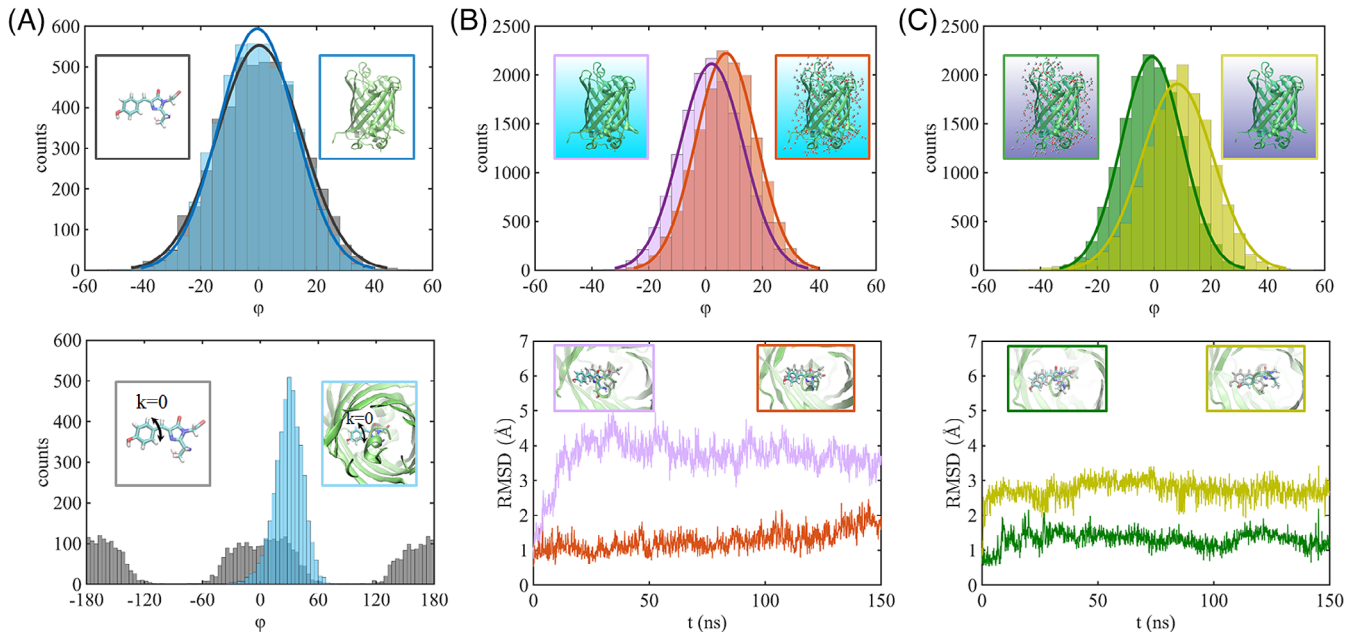

FIGURE 2 The dihedral angle φ distributions sampled from simulations of the chromophore in the vacuum (A), without (gray) and with (blue) green fluorescent protein (GFP), the *top* is for the softly restrained dihedral φ in the current force field, while the *bottom* is for the dihedral φ without the soft restraint; with GFP in the bulk water (B), in the presence (red) or absence (purple) of the crystal waters; with GFP in the octane (C), in the presence (dark green) or absence (yellow) of the crystal waters. The root-mean-square deviations (RMSDs) for the heavy atoms of the chromophore inside the aligned GFP in the water and the octane solvent are shown on the *bottom* (B) and (C), respectively, with (red/dark green) or without (purple/yellow) the crystal waters [Color figure can be viewed at wileyonlinelibrary.com]

slightly from the planar positioning, that is, positively in water, while remain close to planar in the octane.

To examine the impacts of water from bulk and from inside of the protein, we removed crystal water molecules in the control simulations of GFP (still 150 ns each). In either the water or the octane solvent condition, one would see a quick rise of the root-mean-square deviation (RMSD) of the chromophore (see Figure 2B,C *bottom*), indicating an immediate repositioning of the chromophore inside the protein. In the GFP-water simulation deprived of the crystal waters initially, however, the chromophore RMSD dropped slightly at some later stage. Indeed, it is found that bulk water molecules penetrate into GFP (see Supporting Information Figure S3), as also being suggested [23,28]. We address the water penetration later.

To test structural stabilities, we performed additional 50 ns simulations for the above systems (up to 200 ns each). During the additional simulation period, the GFP structure remained largely stable (see RMSDs in Supporting Information Figure S2), though with slightly more deviations from the crystal structure; the chromophore demonstrated more or less shifting inside GFP, yet the dihedral angle distributions remained highly stable (also see Supporting Information Figure S2). Below, we keep showing the simulation results collected from the 150 ns trajectories.

3.2 | The hydrogen bond network coordinating the chromophore inside GFP in water can be largely maintained into the GFP-octane system

Next, we inspected the residues surrounding the chromophore inside GFP and identified the hydrogen bond (HB) network that constantly dictates the chromophore positioning and suppresses the dihedral flexibilities of the chromophore, as discovered previously.[16,29] The HBs are demonstrated on the structural snapshots in Figure 3 and Supporting Information Figure S4, which were taken from four simulation systems (GFP solved in the water and octane solvent conditions, with and without crystal water molecules). The demonstrated HBs were formed for a significantly amount of time (>20% or over 30 ns in the 150-ns simulation). In the GFP-water system kept with crystal waters (see Figure 3A), one could see that the hydroxyl group on the phenolic ring of the chromophore formed a HB with a water molecule (WAT1). WAT1 bonded with two other residues, N146 and S205, simultaneously, and occasionally with H148. These residues (N146, H148, and S205) thus closely tune the dihedral flexibilities of the chromophore. Next, S205 formed a HB with E222 as in the crystal structure, which further bonded to S65 on the chromophore, providing a potential proton transfer loop.[23] Due to highly flexible side chains of S65 and E222, an additional water molecule (WAT2) is needed to maintain the interaction between them, which indeed appeared in the latter stage of the simulation. One could also see that G67 on the chromophore bonded with another water (WAT3) that simultaneously bonded with Y92 and Q94. In addition, T62 formed a HB with H181, while R96 formed a HB with Q183 as well as with Y66. There is an additional region containing several water molecules (WAT4) close around T62 and H181, which help maintain the helix structure in the central part of GFP. Removing the crystal water molecules initially in the simulation immediately disrupted a large portion of the HB network. Nevertheless, as bulk water molecules gradually penetrated into the protein and led to significant rehydration after ~80 ns (see Supporting Information Figure S3), most of the HBs recovered (see Supporting Information Figure S4A), though the HB between S205 and E222 was in a low occupancy (~10%). It suggests that the water-mediated HBs around chromophore remain robust once with sufficient

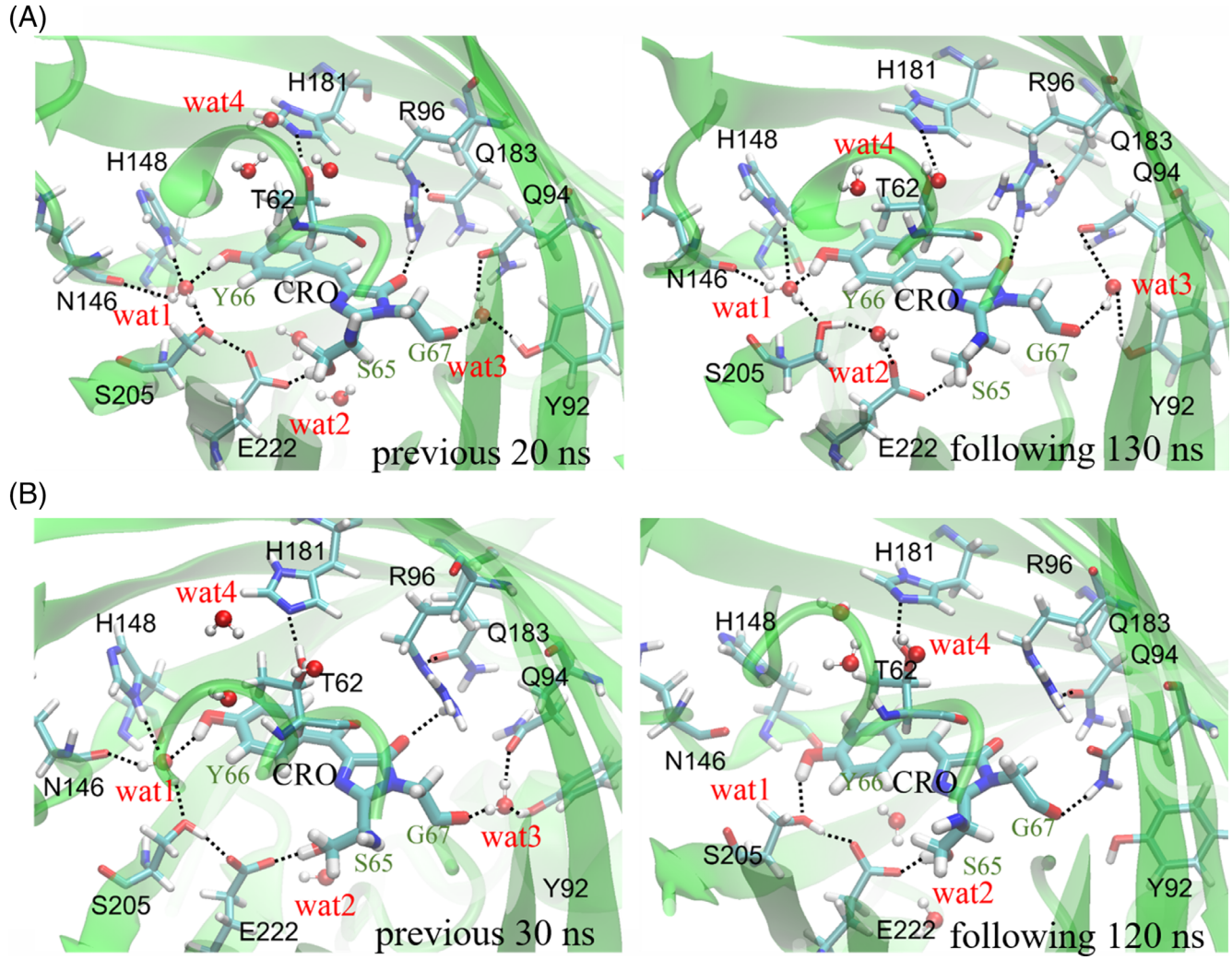

FIGURE 3 The hydrogen bond (HB) networks surrounding the chromophore in the green fluorescent protein (GFP)-water and GFP-octane simulation systems, in the presence of the crystal waters. The HBs were calculated throughout the GFP-water (A) and GFP-octane (B) simulations, respectively (only those formed >20% simulation time are depicted), in the early stage (0 to 30 ns, *left*) and later stage (30 to 150 ns, *right*). The chromophore is labeled as CRO. Four water binding sites from WAT1 to WAT4 are also labeled [Color figure can be viewed at wileyonlinelibrary.com]

hydration supplied into GFP, even though rehydration may take over hundreds of nanoseconds.

In the GFP-octane system kept with crystal waters (see Figure 3B), the HBs associated with the hydroxyl group of the Y66 phenolic ring are largely preserved at the beginning of the simulation (0 to 30 ns). Due to the lack of supply of mobile water molecules for WAT1 and WAT3, Y66 formed a HB with S205 directly, and lost connections with H148 and N146 at the later stage of the simulation (30 to 150 ns). The potential proton transfer loop Y66-S205-E222 was partial reserved, so the overall positioning and suppression of the dihedral fluctuations of the chromophore were still maintained. WAT3 was gone and Q94 connected with G67 directly, while R96 lost connection with chromophore. When the crystal water molecules were removed from the GFP-octane system, however, the chromophore rotated somehow; consequently, even some HBs were established later around chromophore, they rarely resembled those in the original HB network in the presence of the internal waters; the potential proton transfer loop from the chromophore to E222 also disappeared (see Supporting Information Figure S4B). The list of HBs identified from the above-mentioned simulation systems is provided in Supporting Information Table S1.

By comparing results from the GFP-water systems with (Figure 3A) and without crystal water (Supporting Information Figure S4A), one sees that the water-mediated HB network could gradually recover as long as bulk waters penetrate into GFP (Supporting Information Figure S3) to refill the binding sites, which may take over hundreds of nanoseconds to reach equilibrium. Essentially, according to the comparison between the GFP-water and GFP-octane systems in the presence of the crystal waters (Figure 3A,B), we show that the overall HB network is well maintained when GFP is placed into the octane solvent, as long as the essential water molecules are well kept within GFP; without these essential water molecules, the HB network around the chromophore inside GFP is mostly damaged (Supporting Information Figure S4B).

3.3 | Essential water coordination around the chromophore inside GFP can also be preserved into the octane solvent

In order to see how water molecules localize to coordinate with the chromophore inside GFP, we examined water occupancies at those four important positions (from WAT1 to WAT4) identified in the HB

network analyses above, some of which have also been reported recently.[23,28] The WAT1-site is highly crucial, as water occupying this position simultaneously forms HBs with both the hydroxyl group on the Y66 phenolic ring and amino acids N146 and S205 around, acting as a "hub" in the HB network that controls the dihedral degree of freedom of the chromophore. In the GFP-water system, a crystal water molecule could occupy the WAT1-site for over ~10 ns; when the crystal water moves away, other water molecules come and their occupancy or duration time remain significantly larger than that of a random water molecule, located at a same position relative to the chromophore in the bulk water (see Figure 4 *left* and Supporting Information Figure S5). Among them, some show comparatively long duration time as if they are bound water molecules at the WAT1-site, while the others appear transiently bound. In the octane solvent, the crystal water duration time at WAT1-site reached ~30 ns; after the release of the crystal water, other water molecules come into the site and bind rather transiently.

Similar to the WAT1-site, the WAT3-site also hosts water that bridges HBs with three residues, G67, Y92, and Q94, which appeared in both the GFP-water and the GFP-octane simulations. This site can be important for the positioning of the chromophore. Notably, the water occupancy times for this site reached up to ~50 ns and ~30 ns in the GFP-water and GFP-octane systems, respectively (see Figure 4B). Thus, the most crucial water molecules at the WAT1 and WAT3 sites that bridge for the HBs inside GFP, though occasionally transiently bound, were maintained to into the octane solvent.

Meanwhile, we checked the water occupancies at two water-clustering regions defined as the WAT2-site and the WAT4-site (see Supporting Information Figure S6 for definitions). Note that the N-C termini of GFP is placed downside. WAT2-site locates close to E222 at the downside region of GFP, or right *below* the chromophore. The water molecules located around the WAT2-site are expected to assist the proton transfer for the fluorescence.[23] Nevertheless, stable HBs were not maintained at this site, though waters forming transient HBs with E222 were identified. Indeed, waters around this site tended to form a small cluster of three to four molecules on average (see Figure 4C), likely serving as a "water pool." [23] The number of clustering waters within this "pool" could reach up to six or seven for both the GFP-water and the GFP-octane systems. For the WAT4-site located in the upside region of GTP, or *above* the chromophore, there were usually about three to four water molecules staying around T62, H181 region, while one water molecule occasionally formed HB with Q94 in both the GFP-water and the GFP-octane simulations. Indeed, the WAT2-site and the WAT4-site were connected to pores on the bottom and cap regions of the GFP β-barrel, respectively, which allow water permeations in and out of the protein.

Additionally, we counted the number of water molecules close around (within ~6 Å) the chromophore that approximates the amount of water trapped inside GFP (see Figure 4E). One can see that in the water solution, the number varies from ~10 to 20, indicating a significant amount of water fluctuation within GFP. In comparison, the number of water molecules measured inside GFP in the octane solvent remains smaller, ~10 or below, the fluctuations are also less significant than that in the water. Again, the higher number of water molecules and larger fluctuations can be attributed by the bulk water penetration into GFP. In the octane, however, only the water

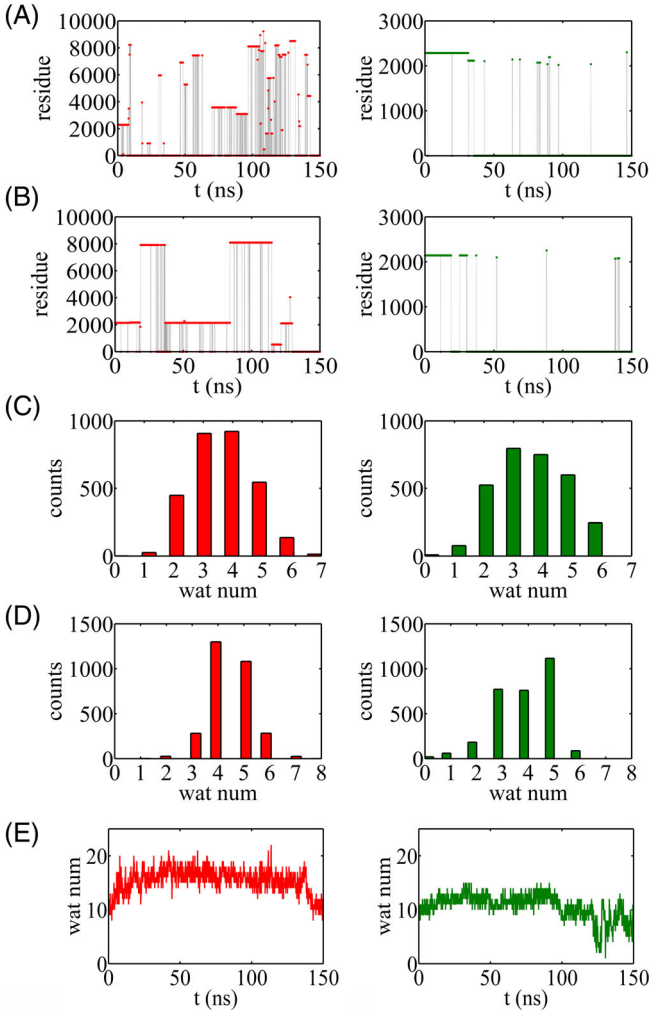

FIGURE 4 The water distributions inside green fluorescent protein (GFP) in the water and the octane solvent. The occupancy water IDs at the WAT1-site (A) and the WAT3-site (B) are shown for the GFP-water (*left*) and GFP-octane (*right*) systems. The number of water molecules clustering at the WAT2-site and the WAT4-site (with "counts" the number of timeframes sampled at every 0.05 ns) are presented in (C) and (D), respectively. The total amount of water molecules around the chromophore within GFP are shown in (E), for the GFP-water (*left*) and the GFP-octane (*right*) systems, respectively [Color figure can be viewed at wileyonlinelibrary.com]

molecules close around the protein surface remain and are able to penetrate inside. We address the water exchange behaviors next.

3.4 | Water exchange in and out GFP are present both in the water and the octane solvent

As water permeations between the inside and outside of GFP revealed, we monitored how those permeations happened in the simulation, both in the water and the octane solvent conditions (see Figure 5A,B and Supporting Information Movies S1 and S2 for the GFP-water and GFP-octane systems, respectively). First, one could see a pore-I and pore-II region quickly form on one side of GFP with accumulated water molecules from the inside and outside for exchanges. In particular, the pore-I region is close to the Y66 phenolic ring, so that it can facilitate quick water replacements at the essential WAT-1 site. After a while, as water from outside started penetrating into GFP, a pore-III (in the GFP-water

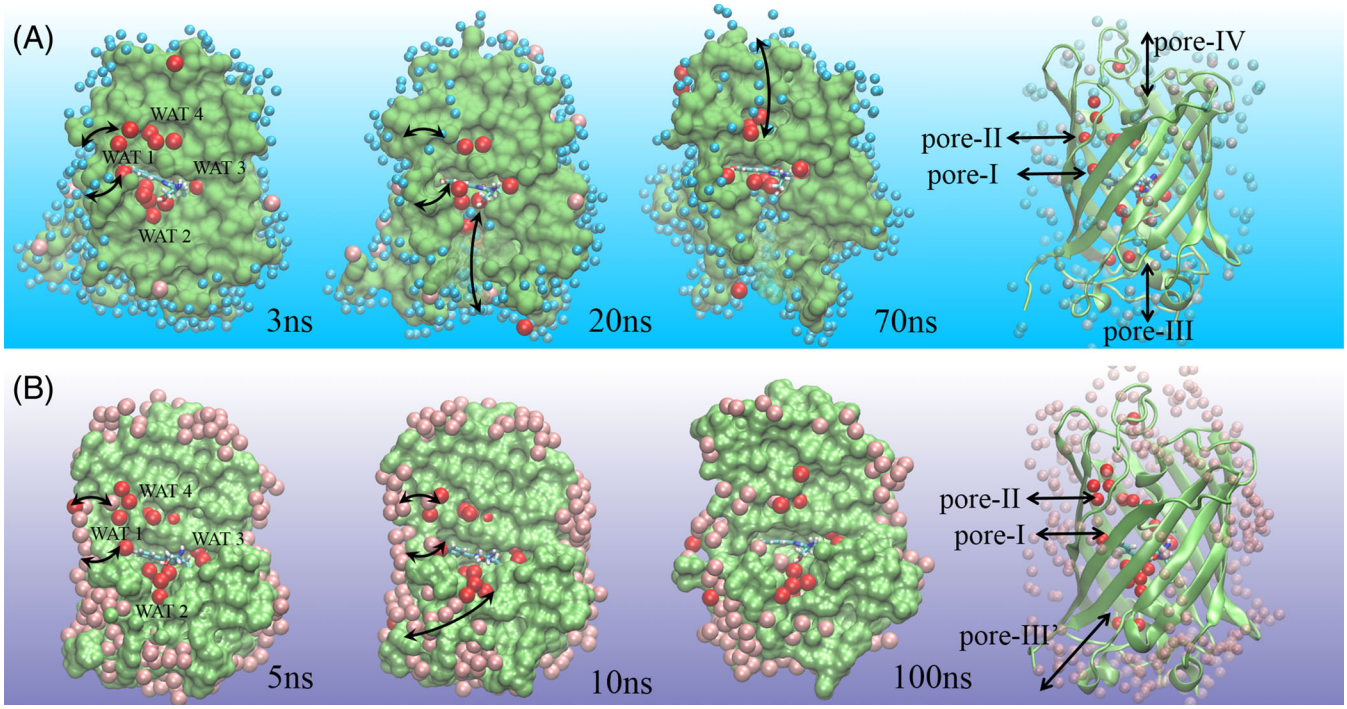

FIGURE 5 The water permeation and exchange in and out of green fluorescent protein (GFP) in the water and the octane solvent. The snapshots were taken from the respective simulation systems in water (A) and octane (B). Also see Supporting Information Movies S1 and S2 for the respective simulations. GFP is shown in green surface representation, the chromophore is shown in licorice in the center of GFP. The oxygen atoms of crystal water molecules are shown in red (inside GFP initially) and pink (on the protein surface initially) spheres, while those bulk water molecules coming close around GFP are shown in blue spheres. Arrows indicate permeation pores. The right-most column shows the β-barrel cartoon representation of GFP, with arrows again indicating the permeation pores on the side (I and II), on the bottom (III and III'), and on the cap region (IV) of GFP. Note that the N- and C-terminal are located on the bottom side of GFP in the view

system) or a pore-III' (in the GFP-octane system) formed on the N- and C-terminal bottom of GFP. The pore-III or III' then allows efficient water exchanges with the WAT2-site. Another pore-IV appeared on the cap region of GFP in the water solvent, which links the WAT4-site to the bulk water. In the GFP-octane system, however, there was no corresponding pore-IV found yet, while crystal waters kept staying around the WAT4-site throughout the simulation (see Figure 5B). These pore regions in the GFP-water system were also found in the recent study which revealed four water pathways inside GFP.[23]

By close examinations, one could see that pore-I and II, appeared on the side of GFP in both the water and octane solvent, are formed next to each other on the two sides of the strand #7 (see Supporting Information Figure S7). The strand #7 region appears highly flexible due to polar resides located on it, which actively interacted with water molecules to allow formation of the pores. Pore-III is formed in the center of the loop region on the bottom of GFP, which only appeared in water but no in the octane solvent. The slightly different Pore-III' is formed between the β-strand #3 and #11, which is tilted toward the bottom, and only opened in the octane solvent (see Supporting Information Figure S7). A close view of the pore-IV among the surrounding loops on the cap region of GFP in water is shown in Supporting Information Figure S7 as well.

3.5 | Experimental verification and dynamics on pulling GFP from water to octane

In our equilibrium simulation of GFP in the octane solvent, the protein appears structurally stable, and importantly, it maintains an internal HB network with participating water molecules to coordinate with the chromophore and suppress the chromophore dihedral fluctuations. Accordingly, the equilibrium simulation studies suggest functional robustness of GFP into the octane, which were then verified by our fluorescence microscope measurements, showing that GFP keeps fluorescent in the octane, though with lower intensity than that in water (see Supporting Information Figure S8).

Next one concerns if the GFP fluorescence can be impacted by how the protein is transferred into the octane, for example, by dynamically flowing or forcing GFP into the octane solvent. During the dynamical process, it is to be determined whether structural stability of GFP is maintained, and whether water molecules are kept along with GFP, in particular, at the water-octane interface. For that purpose, we performed additional MD simulations pulling GFP at a speed of ~0.5 to 1 Å/ns vertically moving from the water phase to the octane phase (see Figure 6). The pulling was conducted by either forcing on the center of the mass (COM) of GFP (in the first 70-ns simulation; see Figure 6A), or by forcing on two different sites of GFP (in the second and third simulations for 120 ns and 150 ns, respectively, see Figure 6B,C).

When pulling the COM of GFP from the water to the octane phase in the first simulation, one can see that GFP rotates freely across the interface between the water and octane, keeping the structure intact (also see Supporting Information Movie S3). In particular, a layer of water molecules surrounding GFP moved along with the protein across the interface, and the water layer was well maintained into the octane phase, similarly as that in the equilibrium simulation. It should be pointed out that there are ~318 crystal water molecules in

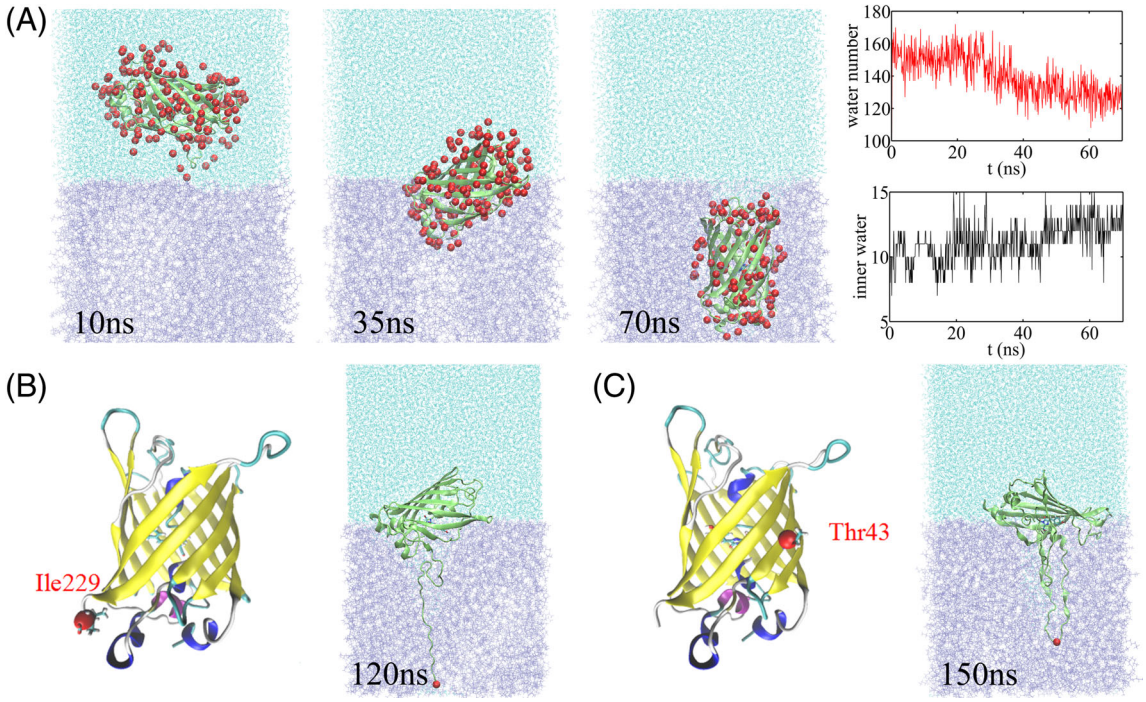

FIGURE 6 Pulling green fluorescent protein (GFP) from the water to the octane phase in the molecular dynamics simulations. A, The first pulling was conducted by forcing on the center of mass of the protein without bias (see Supporting Information Movie S3). The number of water molecules surrounding GFP (within 2.5 Å of the protein surface) and close around the chromophore (within 6 Å) inside GFP were monitored in the simulation (shown on the right). B, The second pulling was conducted by forcing on the C-terminal end of strand #11 on reside Ile229 (see Supporting Information Movie S4). Unfolding happens at the water-octane interface. C, The third pulling was conducted by forcing on the middle region of the strand #3 on residue Thr43 (see Supporting Information Movie S5). Unfolding at the interface happened as well [Color figure can be viewed at wileyonlinelibrary.com]

the currently used GFP structure. Interestingly, one found that the GFP molecule under pulling was ended up with ~120 water molecules close around the protein surface into the octane (see Figure 6A). These ~120 water molecules may constitute a minimal set of water to support the functional robustness of GFP. On the other hand, one sees approximately >10 water molecules constantly kept around the chromophore inside GFP as in the equilibrium simulation.

To probe whether the functional robustness of individual GFP molecules could be achieved, for example, subjecting to forcing in single molecule atomic force microscopy,[3,4] we conducted the second and third pulling simulations by forcing onto one terminal end residue Ile229 and one middle region residue Thr43, respectively. Unfolding of the protein however happened in both simulations at the interface between the water and the octane phase (see Supporting Information Movies S4 and S5). In addition, we performed a control simulation forcing the terminal end residue Ile229 at a lower pulling rate, in which protein unfolding also happens at the water-octane interface (see Supporting Information Movie S6). Hence, even though the β-barrel of GFP appears to be a highly stable structure, the water-octane interface may still pose a significant barrier against the protein structural stability, when the pulling force is applied in a highly biased manner onto the protein.

4 | DISCUSSION

In this work, we implemented all-atom MD simulations to inspect and visualize chromophore fluctuations and surrounding water dynamics in single molecule GFP placed comparatively in the water and the nonpolar solvent conditions. Our studies suggest the functional robustness of GFP into the nonpolar octane, as verified by accompanied fluorescence microscopic measurements.

It is known that GFP maintains bioluminescence well in the water solution, while an isolated chromophore placed in vacuum or in bulk water condition cannot sustain fluorescence.[30–32] Interestingly, it has been found that the chromophore alike molecules, after being suppressed with the dihedral flexibilities, become fluorescent in standalone conditions.[33,34] Hence, inhibition of the dihedral fluctuations of the chromophore appears to be crucial for maintaining the GFP fluorescence function. The dihedral fluctuation inhibition has been identified in our MD simulations of GFP, mainly due to the protein matrix, and additionally due to participating water molecules, consistent with previous studies.[11] The GFP enclosure significantly reduces the chance of *cis* to *trans* transition in the chromophore dihedral rotation, while water molecules trapped inside GFP (eg, over 300 crystal water molecules) help further suppression on the dihedral fluctuations. For GFP deprived of water in vacuum or into the octane, even though the chromophore dihedral fluctuations can be suppressed somehow and the overall β-barrel structure of GFP is maintained, the chromophore will be displaced. In the water condition, however, bulk water always penetrates into GFP to support the luminescent properties.

Importantly, one sees that the fluctuation and positioning of the chromophore is controlled by a HB network,[16,29] which connects key protein residues with a small number of essential water molecules. One finds that in the presence of the crystal waters, the HB network

is largely maintained no matter in the water or the octane solvent. Removing the crystal waters, the HB network is significantly impacted as the chromophore repositions immediately. In particular, there are four specific water sites embedded in the HB network. The most important hub position is the WAT1-site that allows the water to bridge the HBs among several neighboring residues (eg, N146, H148, and S205) with the hydroxyl group of the phenolic ring of the chromophore, so that the dihedral flexibility of the chromophore is controlled. The water occupancy time at the WAT1-site reaches several nanoseconds, with similar statistics in water and octane. A second WAT2-site (or a "water pool" [23]) right below the chromophore and close around E222 demonstrates clustering of two to six waters around, though no stable HBs last over nanoseconds at that site. A third WAT3-site maintains HBs between the water and neighboring residues (eg, G67, Y92, and Q94) to position the chromophore, and the water occupancy time at this site reaches up to tens of nanoseconds. A last WAT4-site locating above the chromophore hosts three to six water molecules, clustering around T62 and H181 to sustain the central helical structure above the chromophore. Hence, it is the specific HB interactions along with steric impacts from transiently bound waters and water dynamics that allow the chromophore to be maintained in a functional configuration, which remains robust from the water to the octane solvent condition. Indeed, about 10 to 20 water molecules seem to remain inside GFP to control the chromophore stabilities.

Furthermore, we examined water permeations and exchanges between the inside and outside of GFP. In particular, three regions on the β-barrel of GFP demonstrate permeable pores. In both the water and octane condition, pore I and II formed side by side along the strand #7 on the side of the β-barrel, through which the internal water molecules, in particular, the crucial one at the WAT1-site could be replenished fast. Pore-III and III' have been identified on the N- and C-terminal or bottom side of GFP, with pore-III shown in the middle of the loop region in water, while pore-III' tilted following the strand #3 and #11 when GFP is placed in octane. Both pores facilitate the constant water clustering at the WAT2-site. In addition, Pore-IV has also been found on the cap region of GFP connecting to the WAT4-site in the water condition, but not yet in the octane condition. One thus sees that the β-barrel structure does not set a closured wall for complete segregation, instead, the permeable pores allow water exchange dynamics to essentially support the bioluminescence.[23,28] It is then expected that by structurally redesigning the permeable pore regions, the luminescent properties of GFP could be modulated. On the other hand, one sees that the nonpolar solvent molecule examined in this work is unable to penetrate into the β-barrel of GFP. Indeed, an individual octane molecule (∼8 Å) appears significantly larger than the size of the permeable pore (∼4 Å) on the β-barrel. Following this line of thinking, one would expect that an alternative nonpolar solvent composed of larger solvent molecule (eg, the hexadecane) still support the GFP fluorescence, while a nonpolar solvent made of smaller solvent molecules (eg, the tetrachloroethane) may be able to penetrate into GFP to disrupt the fluorescence. Besides, cellular membrane also provides a nonpolar and nonpenetrating environment for bioluminescence implementations.

As our fluorescence microscope measurements confirmed the GFP function into the octane solvent, we further wonder whether the fluorescence properties preserve upon dynamically enforcing GFP from water to octane, particularly crossing the interface between the water and octane. For example, one may flush or flow GFP molecules from water to octane, or even individually drag GFP to cross the water-octane interface. Our simulations suggest that, however, only via an unbiased dynamical enforcing procedure, the GTP structure can be well maintained. For a biased enforcing approach, the protein unfolds crossing the interface, even though the β-barrel of GFP appears originally stable. Meanwhile, by pulling GFP on the COM in an unbiased manner from the water to the octane phase, we show that a minimal set of water molecules are required in accompany with GFP to support the fluorescence.

5 | CONCLUSIONS

We implemented all-atom MD simulations to demonstrate functional robustness of GFP from water to the nonpolar octane solvent. It is found that both the GFP protein matrix and accompanied water molecules contribute to position the chromophore and suppress the dihedral fluctuations of the chromophore that is crucial for the fluorescence. In particular, specific water occupancy sites inside GFP have been identified, which seem to be preserved from the water to the octane condition and support a substantial HB network inside GFP to allow for the bioluminescence. Water permeable pores also reveal on the β-barrel of GFP to facilitate fast water exchanges between the inside and outside of protein. The functional robustness of GFP into the nonpolar octane is thus achieved by maintaining the β-barrel structure along with a dynamical set of water molecules to coordinate with the chromophore inside GFP. Upon current verification on the GFP fluorescence into the octane, further experiments on dynamically probing the GFP function across the water-nonpolar phase can be conducted. Functional redesigns of GFP and GFP-fusion assembly in nonpolar environments can also be approached.

ACKNOWLEDGMENTS

This work is supported by NSFC Grant #11635002 and NSAF Grant U1530401. We acknowledge the computational support from the Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) under Grant No. U1501501 and from the Beijing Computational Science Research Center (CSRC).

ORCID

Liqiang Dai https://orcid.org/0000-0001-9842-2524
Shuxun Cui https://orcid.org/0000-0002-7713-7377
Jin Yu https://orcid.org/0000-0001-8224-1374

REFERENCES

1. Pace CN, Trevino S, Prabhakaran E, Scholtz JM. Protein structure, stability and solubility in water and other solvents. *Phil Trans R Soc Lond B*. 2004;359:1225-1235.
2. Prabhu N, Sharp K. Protein-solvent interactions. *Chem Rev*. 2006;106:1616-1623.

3. Cui S. The possible roles of water in the prebiotic chemical evolution of DNA. *Phys Chem Chem Phys*. 2010;12:10147-10153.
4. Cui S, Yu J, Kuhner F, Schulten K, Gaub H. Double-stranded DNA dissociates into single strands when dragged into a poor solvent. *J Am Chem Soc*. 2007;129:14710-14716.
5. Mattos C, Ringe D. Proteins in organic solvents. *Curr Opin Struct Biol*. 2001;11:761-764.
6. Klibanov AM. Improving enzymes by using them in organic solvents. *Nature*. 2001;409:241-246.
7. England JL, Haran G. Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu Rev Phys Chem*. 2011;62:257-277.
8. Lee EH, Hsin J, Sotomayor M, Comellas G, Schulten K. Discovery through the computational microscope. *Structure*. 2009;17:1295-1306.
9. Voityuk AA, Michel-Beyerle M-E, Rosch N. Structure and rotation barriers for ground and excited states of the isolated chromophore of the green fluorescent protein. *Chem Phys Lett*. 1998;296:269-276.
10. Weber W, Helms V, McCammon A, Langhoff PW. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proc Natl Acad Sci U S A*. 1999;96:6177-6182.
11. Maddalo S, Zimmer M. The role of the protein matrix in green fluorescent protein fluorescence. *Photochem Photobiol*. 2006;82:367-372.
12. Tsien RY. The green fluorescent protein. *Annu Rev Biochem*. 1998;67:509-544.
13. Peng Q, Yi Y, Shuai Z, Shao J. Toward quantitative prediction of molecular fluorescence quantum efficiency: role of duschinsky rotation. *J Am Chem Soc*. 2007;129:9333-9339.
14. Stepanenko OV, Stepaneoho OV, Kuznetsova IM, Verkhusha VV, Turoverov KK. β-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *Int Rev Cell Mol Biol*. 2013;302:221-278.
15. Grigorenko BL, Krylov AI, Nemukhin AV. Molecular modeling clarifies the mechanism of chromophore maturation in the green fluorescent protein. *J Am Chem Soc*. 2017;139:10239-10249.
16. Patnaik SS, Trohalaki S, Pachter R. Molecular modeling of green fluorescent protein: structural effects of chromophore deprotonation. *Biopolymers*. 2004;75:441-452.
17. Megley CM, Dickson LA, Maddalo SL, Chandler GJ, Zimmer M. Photophysics and dihedral freedom of the chromophore in yellow, blue, and green fluorescent protein. *J Phys Chem B*. 2009;113:302-308.
18. Shinobu A, Palm GJ, Schierbeek AJ, Agmon N. Visualizing proton antenna in a high-resolution green fluorescent protein structure. *J Am Chem Soc*. 2010;132:11093-11102.
19. Chattoraj M, King BA, Bublitz GU, Boxer SG. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc Natl Acad Sci U S A*. 1996;93:8362-8367.
20. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26:1781-1802.
21. MacKerell AD, Bashford D, Bellott M, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 1998;102:3586-3616.
22. Reuter N, Lin H, Thiel W. Green fluorescent proteins: empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulations of the wild type and S65T mutant. *J Phys Chem B*. 2002;106:6310-6321.
23. Shinobu A, Agmon N. The hole in the barrel: water exchange at the GFP chromophore. *J Phys Chem B*. 2015;119:3464-3478.
24. Dargen T, York D, Pedersen L. Particle mesh Ewald: an N·log (N) method for Ewald sums in large systems. *J Chem Phys*. 1993;98:10089-10092.
25. Paterlini MG, Ferguson DM. Constant temperature simulations using the Langevin equation with velocity Verlet integration. *Chem Phys*. 1998;236:243-252.
26. Martyna GJ, Tobias DJ, Klein ML. Constant pressure molecular dynamics algorithms. *J Chem Phys*. 1994;101:4177-4189.
27. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation: the Langevin piston method. *J Chem Phys*. 1995;103:4613-4621.
28. Li B, Shahid R, Peshkepija P, Zimmer M. Water diffusion in and out of the β-barrel of GFP and the fast maturing fluorescent protein, turboGFP. *Chem Phys*. 2012;392:143-148.
29. Helms V, Straatsma TP, McCammon JA. Internal dynamics of green fluorescent protein. *J Phys Chem B*. 1999;103:3263-3269.
30. Brejc K, Sixma TK, Kitts PA, et al. Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proc Natl Acad Sci U S A*. 1997;94:2306-2311.
31. Frankevich V, Barylyuk K, Chingin K, Nieckarz R, Zenobi R. Native biomolecules in the gas phase? The case of green fluorescent protein. *Chemphyschem*. 2013;14:929-935.
32. Brama E, Peddie C, Jones ML, et al. Standard fluorescent proteins as dual-modality probes for correlative experiments in an integrated light and electron microscope. *J Chem Biol*. 2015;8:179-188.
33. Wu L, Burgess K. Syntheses of highly fluorescent GFP-chromophore analogues. *J Am Chem Soc*. 2008;130:4089-4096.
34. Hong Y, Lam JW, Tang BZ. Aggregation-induced emission: phenomenon, mechanism and applications. *Chem Commun*. 2009;29:4332-4353.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Dai L, Zhang B, Cui S, Yu J. Inspecting fluctuation and coordination around chromophore inside green fluorescent protein from water to nonpolar solvent. *Proteins*. 2019;87:531–540. https://doi.org/10.1002/prot.25676