crystals

MDPI

Article

Effects of Proline Substitutions on the Thermostable LOV Domain from *Chloroflexus aggregans*

Alina Remeeva [1], Vera V. Nazarenko [1], Ivan M. Goncharov [1], Anna Yudenko [1], Anastasia Smolentseva [1], Oleg Semenov [1], Kirill Kovalev [1,2,3,4,5], Cansu Gülbahar [6], Ulrich Schwaneberg [6,7], Mehdi D. Davari [6], Valentin Gordeliy [1,2,3,4] and Ivan Gushchin [1,*]

1 Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology, 141700 Dolgoprudny, Russia
2 Institut de Biologie Structurale J.-P. Ebel, Université Grenoble Alpes-CEA-CNRS, 38000 Grenoble, France
3 Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, 52428 Jülich, Germany
4 JuStruct: Jülich Center for Structural Biology, Forschungszentrum Jülich, 52428 Jülich, Germany
5 Institute of Crystallography, RWTH Aachen University, 52062 Aachen, Germany
6 Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany
7 DWI-Leibniz Institute for Interactive Materials, 52074 Aachen, Germany
* Correspondence: ivan.gushchin@phystech.edu

Received: 7 March 2020; Accepted: 26 March 2020; Published: 28 March 2020

check for updates

Abstract: Light-oxygen-voltage (LOV) domains are ubiquitous photosensory modules found in proteins from bacteria, archaea and eukaryotes. Engineered versions of LOV domains have found widespread use in fluorescence microscopy and optogenetics, with improved versions being continuously developed. Many of the engineering efforts focused on the thermal stabilization of LOV domains. Recently, we described a naturally thermostable LOV domain from *Chloroflexus aggregans*. Here we show that the discovered protein can be further stabilized using proline substitution. We tested the effects of three mutations, and found that the melting temperature of the A95P mutant is raised by approximately 2 °C, whereas mutations A56P and A58P are neutral. To further evaluate the effects of mutations, we crystallized the variants A56P and A95P, while the variant A58P did not crystallize. The obtained crystal structures do not reveal any alterations in the proteins other than the introduced mutations. Molecular dynamics simulations showed that mutation A58P alters the structure of the respective loop (Aβ-Bβ), but does not change the general structure of the protein. We conclude that proline substitution is a viable strategy for the stabilization of the *Chloroflexus aggregans* LOV domain. Since the sequences and structures of the LOV domains are overall well-conserved, the effects of the reported mutations may be transferable to other proteins belonging to this family.

Keywords: fluorescent proteins; light-oxygen-voltage (LOV) domain; protein design; thermal stabilization; proline substitution; consensus design; X-ray crystallography

1. Introduction

Light-oxygen-voltage (LOV) domains are ubiquitous photosensory modules found in proteins from bacteria, archaea, fungi, plants and protists [1,2]. They bind different flavonoids, mainly FMN and FAD, as cofactors, and absorb blue and ultraviolet light [1]. Upon illumination, LOV domains undergo conformational changes that may result in a variety of outcomes: they may partially unfold, form homo- or heterodimers, translocate to the plasma membrane, regulate the activity of a kinase or some other effector domain [1,3–6].

Due to their versatility, LOV-based molecular tools have found widespread use in molecular biology. LOV domains lacking the conserved cysteine amino acid in the vicinity of the flavonoid cofactor have been used as fluorescent tags, which may outperform traditional fluorescent proteins under some conditions [7–10], such as the absence of molecular oxygen [11,12]. LOV domains have been used as light-induced generators of reactive oxygen species and photosensitizers [13–15]. Finally, along with microbial rhodopsins [16], they are among the most widely used optogenetic tools, employed for light-controlled peptide uncaging, association or regulation of the effector modules [5,17].

The search for LOV domains with improved or novel functions is still ongoing, with the exciting new applications being continuously discovered [2,6,18]. One of the most desired properties of many molecular tools is stability, and a significant effort has been made to develop thermo- or photostabilized LOV proteins. Recombination and directed evolution have been used to develop iLOV and phiLOV proteins, which were brighter and more photostable compared to their natural counterparts [19,20]. Rational engineering has been used to thermostabilize YtvA [21], while plasmid recombineering and directed evolution have been used to thermostabilize iLOV [22,23]. As an alternative to the engineering of mesophilic proteins, different LOV domains from thermophilic microorganisms have been screened for thermo- and photostability [24]. Recently, we identified a LOV domain from the thermophile *Chloroflexus aggregans*, and engineered its cysteine-less flavin-based fluorescent protein variant CagFbFP, which revealed remarkable stability in a variety of solvents and an ability to refold easily [25].

In this work, we inquired whether the naturally thermostable protein CagFbFP can be further stabilized using rational engineering. Out of different approaches to protein stabilization that are currently in use [26–29], we chose to combine proline substitution with consensus design. We identified three sites for proline substitution and determined the thermal stability of the resulting variants. We found that the melting temperature of the A95P mutant is raised by approximately 2 °C, whereas mutations A56P and A58P are neutral. Subsequently, we crystallized the A56P and A95P variants and modeled the effects of the mutation A58P, and determined that their structures are conserved compared to the parent protein.

2. Materials and Methods

2.1. Sequence Analysis

Light-oxygen-voltage (LOV) domain sequences presented in Figure 1 and those from Glantz et al. [2] were aligned using Clustal [30] and visualized using Jalview [31]. *Chloroflexus aurantiacus* and *Chloroflexus islandicus* LOV protein sequences were identified using the NCBI BLAST search [32]. The *Chloroflexus aurantiacus* LOV protein sequence can also be found in the dataset by Glantz et al. [2].

2.2. Cloning, Protein Expression and Purification

Genes encoding CagFbFP and its variants were cloned into pET11, as described previously [25]. The mutations (A56P, A58P and A95P) were introduced into the CagFbFP gene using polymerase chain reaction. The proteins were expressed and purified, as described previously [25]. In brief, we used the *Escherichia coli* strain C41 (DE3) and an autoinducing medium [33], and the proteins were purified using Ni-NTA resin (Qiagen, Germany) on a gravity flow column, followed by size-exclusion chromatography on a Superdex® 200 Increase 10/300 GL column (GE Healthcare Life Sciences, USA) in a buffer containing 10 mM NaCl and 10 mM sodium phosphate, pH 8.0.

2.3. Spectroscopic Characterization

Absorption and fluorescence spectra were recorded using a Synergy™ H4 Hybrid Microplate Reader (BioTek, USA). Emission spectra were measured between 470 nm and 700 nm, while the excitation wavelength was set to 450 nm. Excitation was measured between 250 nm and 500 nm, and the signal was detected at 510 nm. The thermal stability of CagFbFP and all mutated variants were determined by detecting FbFP-specific fluorescence using the Rotor-Gene Q real-time PCR cycler

(Qiagen, Germany) in heating and cooling experiments, where the temperature was changed at a rate of 1 °C per minute, changing from 25 °C to 99 °C, and then back to 25 °C. 25 µL samples with a protein concentration of 1 mg mL^{-1} were used. The excitation wavelength was set to 470 nm, and the fluorescence emission was recorded at 510 nm. Melting temperatures were determined from the first derivative of the melting curve, by fitting the smoothened data (FFT filter method, 15 points, implemented in Origin 9.0 G (OriginLab Corporation, Northampton, MA, USA)), to multiple Gaussian functions (Origin 9.0 G, Quick Peaks Gadget).

2.4. Crystallization

Protein-containing fractions were concentrated to 28 mg mL^{-1}. Crystallization trials were set up using a hanging drop vapor diffusion approach at NT8 nanovolume robotic system (Formulatrix, USA). The drops contained 100 nL concentrated protein solution and 150 nL reservoir solution. CagFbFP-A56P and CagFbFP-A95P crystallized in a broad range of crystallization conditions (pH from 5.5 to 8.5, 15–25% Polyethylene glycol (PEG) 3350–10,000, different salts), similar to ones that gave crystals of the original CagFbFP protein [25]. The crystals were grown at 22 °C and reached the size of 100–700 µm within a few days. Attempts to crystallize CagFbFP-A58P were unsuccessful. The crystallization conditions used to obtain the crystals that resulted in the presented structures were as follows: 0.1 M Ammonium acetate, 0.1 M BIS-Tris pH 5.5, 17% w/v PEG 10,000 for CagFbFP-A56P; and 0.2 M Ammonium sulfate, 0.1 M BIS-Tris pH 5.5, 25% w/v PEG 3350 for CagFbFP-A95P. All crystals were harvested using micromounts, flash-cooled and stored in liquid nitrogen.

2.5. Acquisition and Treatment of Diffraction Data

The diffraction data were collected at 100 K on the BioMAX macromolecular crystallography beamline at MAX IV Laboratory (Lund, Sweden). Diffraction images were processed using XDS [34]. POINTLESS and AIMLESS [35] were used to merge, scale and assess the quality of the data, as well as to convert intensities to structure factor amplitudes and generate Free-R labels.

2.6. Structure Determination and Refinement

The structures of CagFbFP-A56P and CagFbFP-A95P were solved using molecular replacement with MOLREP [36] and CagFbFP structure (PDB ID 6RHF) [25] as a search model. The resulting model was refined manually using Coot [37] and REFMAC5 [38].

2.7. Molecular Dynamics (MD) Simulations

The initial coordinates for the wild type CagFbFP simulations were taken from the X-ray structure (PDB ID 6RHF) [25]. The models of the CagFbFP-A58P variant were constructed using the SWAP function in YASARA Structure Version 17.4.17 [39], and optimized using the SCWRL rotamer library search [40]. The lowest energy conformers were selected for further studies. Two sets of independent simulations for each conformer of Ala58 and Pro58 ("A" and "B") were carried out. The protonation states of titratable residues at pH 7 were assigned on the basis of pKa calculations using the PROPKA 3.1 program [41] and visual inspection; all charged residues were kept at their standard protonation states. Side chains of Asn and Gln residues were checked for possible flipping. The phosphate group of the FMN cofactor was deprotonated, carrying a charge of $-2e$. Consequently, CagFbFP and CagFbFP-A58P dimers had a total charge of $-4e$. To neutralize the systems, solvent water molecules that were at least 5.5 Å away from any protein atoms were replaced by Na^+ ions. Hydrogen atoms were added, employing the *tleap* module of AmberTools14 [42]. Crystal water molecules were kept; the protein was solvated in a water box centered at the center of mass to ensure a water layer of 12 Å around the protein. The total size of the simulated systems was ~43,000 atoms, including ~13,200 TIP3P [43] water molecules. All MD simulations were carried out using the Amber14 program [42] with the Amber ff99SB [44,45] all-atom force field for proteins, the general Amber force field (GAFF) [46] for flavin mononucleotide (FMN) and the TIP3P model for water [43]. We used the atomic charges and force

field parameters for FMN moiety reported in our previous work [47]. Initially, the solvent and the ions followed by the whole system were subjected to minimization using 10,000 steps of the steepest descent, followed by 3000 steps of conjugate-gradient minimization. The system was then slowly heated from 0 to 300 K for 50 ps. In all simulations, constant pressure periodic boundary conditions using the Particle Mesh Ewald (PME) [35] method were employed. To calculate the electrostatic interactions, a cutoff of 10 Å was used. After heating, the systems were equilibrated for 1000 ps at 300 K. Finally, three independent 50 ns-long production runs were performed for each alternative conformer of WT and A58P variants. Pymol [48], VMD [49] and AmberTools14 [42] were used for molecular visualizations and analysis of MD simulations.

3. Results

3.1. Identification of Positions for Proline Substitutions

Recently, we have identified and studied a small thermostable flavin-based fluorescent protein CagFbFP, derived from a soluble light-oxygen-voltage (LOV) domain-containing histidine kinase from the thermophilic bacterium *Chloroflexus aggregans* [25]. The protein crystallized well, and an ultra-high resolution structure of CagFbFP has been determined [25]. While most of the protein is very well-ordered, backbones as well as side chains of residues Ala58 and Asp59, located in the loop between β-strands Aβ and Bβ, were observed to adopt two alternative conformations. We reasoned that targeting this region with mutations might stabilize it in a single conformation, and also stabilize the overall protein.

To understand the natural variability of the amino acids observed in these and other positions in different LOV domains, we prepared a multiple sequence alignment of different LOV-derived fluorescent proteins [9,19,25,50,51] (Figure 1). In addition to these proteins, we wanted to compare CagFbFP to its close homologs identified in the genomes of *Chloroflexus aurantiacus* [52] and *Chloroflexus islandicus* [53], with sequence identities of 72% and 87%, respectively. The sequence alignment revealed that the loop connecting Aβ and Bβ is shorter by one amino acid in the *Chloroflexi* proteins compared to others. At the same time, prolines are observed in iLOV, EcFbFP and DsFbFP at the position of CagFbFP's Ala56, as well as most of the proteins, including the ones from *Chloroflexus aurantiacus* and *Chloroflexus islandicus*, at the position of CagFbFP's Ala58 (Figure 1). Following this observation, we calculated how often prolines are observed in the sequences of all LOV domains found by Glantz et al. [2]. Prolines are observed in ~47% of proteins at the position of Ala56, and in ~77% of proteins at the position of Ala58. Since a proline's backbone is naturally more rigid than that of other amino acids, it can stabilize certain kinks in the protein, and consequently the overall protein [27,28]. Additionally, mutating a particular amino acid to the consensus one often improves the stability of the resulting protein [26]. Consequently, mutating Ala56 and Ala58 into prolines might be beneficial for the stability of CagFbFP.

Careful analysis of the sequence alignment revealed also another position, that of CagFbFP's Ala95, situated at the N-terminus of the helix Fα, which is also often occupied by prolines in other LOV proteins. Overall, 33% of the proteins from the dataset from Glantz et al. [2] contain a proline at this or the neighboring position. Consequently, we focused on probing the effects of the mutations A56P, A58P and A95P on the stability of CagFbFP.

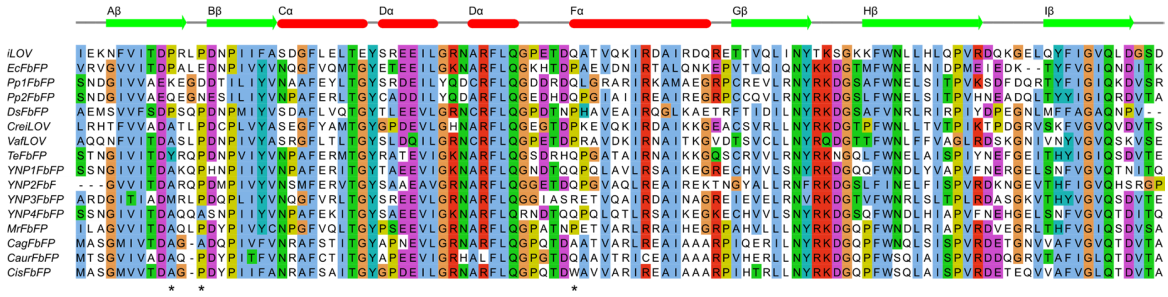

Figure 1. Multiple sequence alignment of LOV domain cores of different LOV-derived fluorescent proteins: iLOV [19], EcFbFP [9], Pp1FbFP [50], Pp2FbFP (formerly PpFbFP [9]), DsFbFP [50], CreiLOV and VafLOV [51], six different FbFPs from thermophilic microorganisms [24], CagFbFP [25] and CagFbFP homologs from *Chloroflexus aurantiacus* and *Chloroflexus islandicus*. CagFbFP alanines that were substituted with prolines in this work are marked with asterisks.

3.2. Characterization of Ala→Pro CagFbFP Mutants

Following the identification of the prospective positions for Ala→Pro substitutions, we have prepared atomistic models of the corresponding variants, CagFbFP-A56P, CagFbFP-A58P and CagFbFP-A95P, using PyMOL [48]. The models revealed that, as expected from sequence alignment (Figure 1) and structures of other LOV proteins, the replacement of Ala56, Ala58 or Ala95 with prolines should not disturb the protein backbone (Figure 2). Consequently, we produced the mutated variants and evaluated their properties in vitro. The absorption, excitation and emission spectra of CagFbFP-A56P, CagFbFP-A58P and CagFbFP-A95P are identical to those of CagFbFP, with the excitation maximum at ~449 nm and the emission maximum at ~495 nm (Supplementary Materials Figures S1–S3).

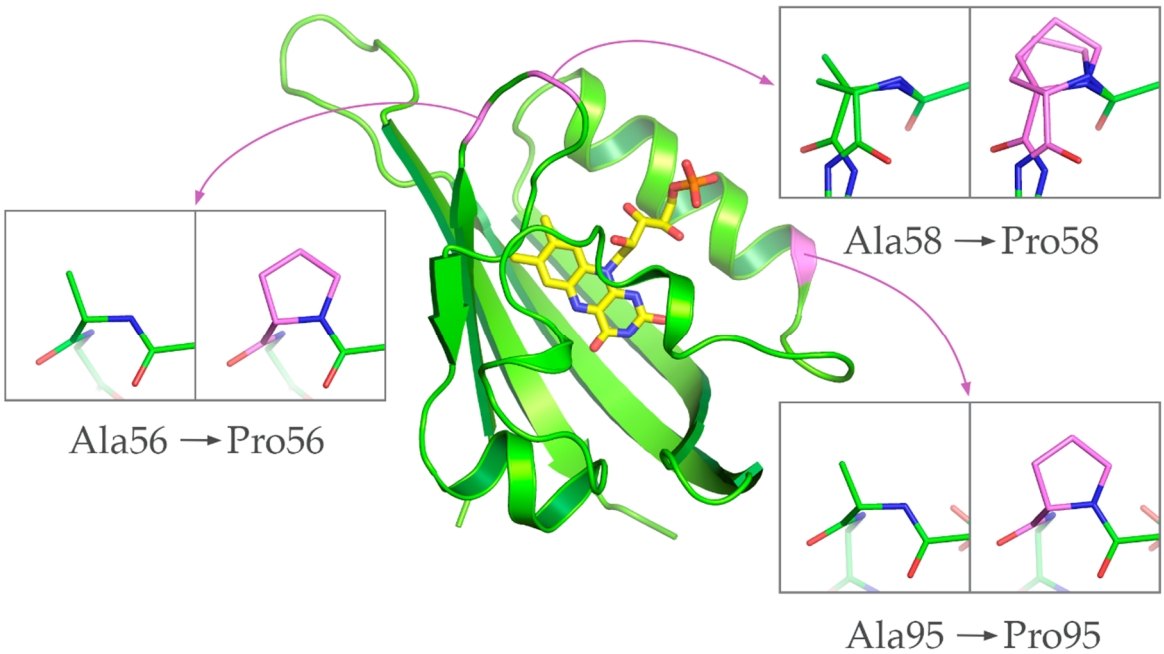

Figure 2. Location of the prospective proline substitution sites (magenta, modeled using PyMOL [48]) mapped onto the CagFbFP structure (green, PDB ID 6RHF). Ala58 (top right) adopts two alternative conformations in the original structure [25].

To evaluate the thermal stability of the mutated variants, we measured the dependence of the fluorescence intensity on temperature during heating-induced denaturation and cooling-induced refolding of the purified proteins (Figure 3). Similarly to CagFbFP, the variants CagFbFP-A56P,

CagFbFP-A58P and CagFbFP-A95P reveal two melting transitions upon denaturation, and only one upon refolding, as evidenced by derivatives of fluorescence as a function of temperature (Figure 3). The results of these experiments are summarized in Table 1. Overall, the mutation A56P slightly stabilizes the protein, but does not facilitate refolding; A58P does not significantly influence denaturation, but slightly facilitates refolding; and A95P stabilizes the protein and facilitates refolding. In all of the cases, the refolding is not complete, and refolded fraction is similar within the experimental errors. However, we should note that part of the differences in the fluorescence of the original and refolded samples is due to irradiation-induced damage of the chromophore, which is especially strong at elevated temperatures.

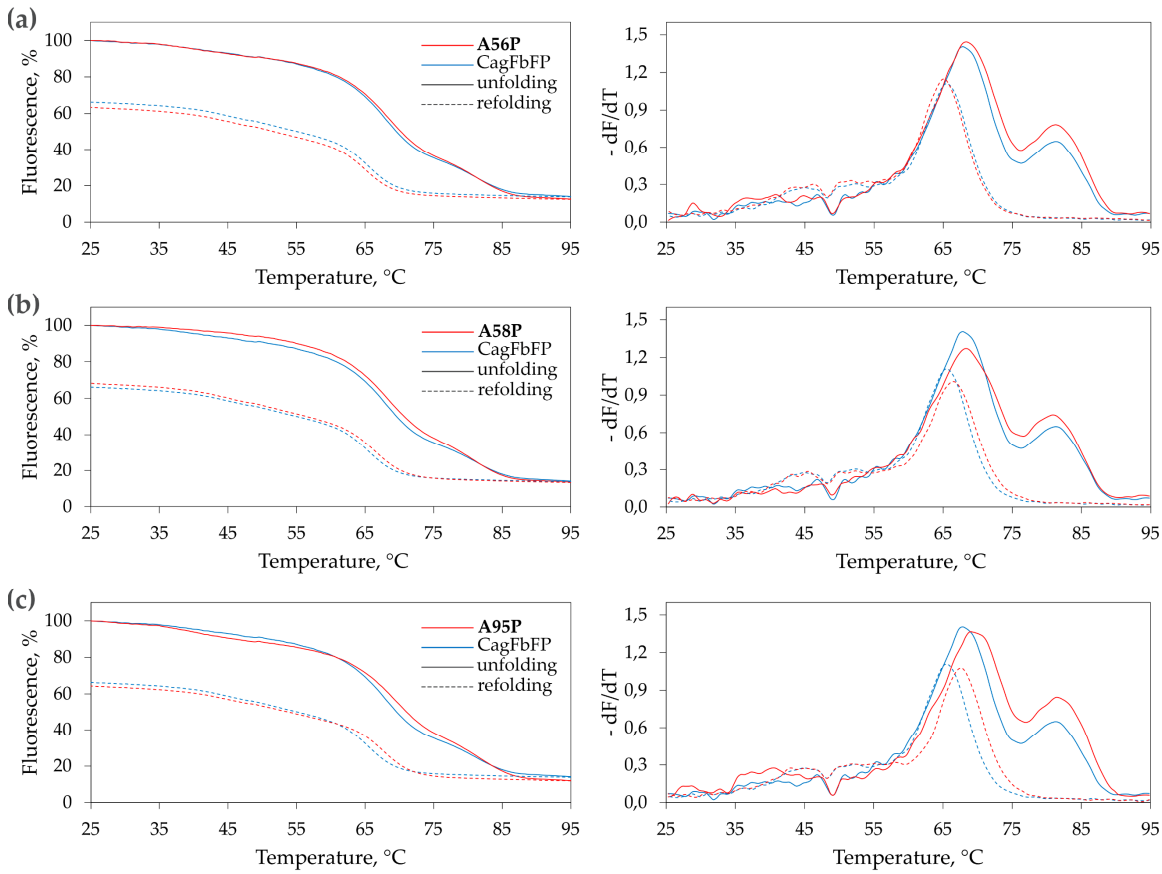

Figure 3. Thermal stability of proline mutants of CagFbFP determined based on flavin fluorescence: (**a**) CagFbFP-A56P; (**b**) CagFbFP-A58P; and (**c**) CagFbFP-A95P. The dependence of fluorescence on the temperature of the sample during heating (solid lines) and cooling (dashed lines) is shown on the left, and the derivative of the fluorescence with respect to the temperature is shown on the right. Each experiment was conducted independently four times, and the data were averaged for plotting. Characteristic unfolding and refolding temperatures are summarized in Table 1.

Table 1. Melting temperatures of CagFbFP and its proline mutants. T_{m1} and T_{m2} correspond to the two melting transitions (Figure 3), and T_r corresponds to the temperature of refolding. Errors are standard deviations of the values observed in four independent experiments.

Variant.	T_{m1}, °C	T_{m2}, °C	T_r, °C
Wild type	67.9 ± 0.3	81.3 ± 0.3	65.4 ± 0.3
A56P	68.3 ± 0.3	81.3 ± 0.4	65.2 ± 0.3
A58P	68.3 ± 0.3	80.9 ± 0.3	66.3 ± 0.3
A95P	69.2 ± 0.5	81.5 ± 0.3	67.7 ± 0.3

3.3. Crystallization of the Mutated CagFbFP Variants

To gain structural information about the effects of proline substitutions, we attempted crystallization of the mutated CagFbFP variants. CagFbFP-A56P and CagFbFP-A95P formed large crystals, reaching up to 700 µm in size, which diffracted to 1.6 Å and belonged to the same space group as CagFbFP ($P2_12_12$). The data collection statistics are reported in Table 2. On the contrary, CagFbFP-A58P did not form crystals; only amorphous aggregates were observed in some of the crystallization trials. The reason for this is not clear; possibly as Ala58 is close to Arg86 and Gln123 of the adjacent protein chains in the crystals of CagFbFP, the mutation A58P in CagFbFP-A58P resulted in steric clashes and prevented the formation of crystal contacts of the same type. While the resolution of CagFbFP-A56P and CagFbFP-A95P crystal structures is lower than that of the original CagFbFP structure [25], it could have likely been improved by the extensive screening of crystals and using advanced diffraction data collection strategies.

Table 2. Crystallographic data collection and refinement statistics. R.m.s.: root mean square.

Data Collection		
Variant	**A56P**	**A95P**
Protein Data Bank ID	6Y7R	6Y7U
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions	-	-
a, *b*, *c* (Å)	54.6, 111.3, 39.2	54.3, 111.4, 38.9
Wavelength (Å)	0.9762	0.9762
Resolution (Å)	111.29–1.60 (1.63–1.60) *	111.36–1.60 (1.63–1.60) *
R_{pim} (%)	3.3 (46.5) *	3.2 (21.5) *
$<I/\sigma I>$	14.5 (1.9) *	14.0 (3.4) *
CC1/2 (%)	99.9 (76.1) *	99.9 (92.8) *
Completeness (%)	100.0 (100.0) *	99.9 (100.0) *
Multiplicity	13.1 (13.3) *	10.5 (11.1) *
Unique reflections	32,415 (1580) *	31,994 (1582) *
Refinement		
Resolution (Å)	39.25–1.60	38.97–1.60
No. reflections	30,768	30,392
R_{work}/R_{free} (%)	18.4/21.4	17.8/22.5
No. atoms	-	-
Protein	1717	1680
FMN	62	62
Water	283	221
Average *B* factors (Å^2)	-	-
Protein	20.3	19.9
FMN	16.2	16.5
Water	33.5	30.6
R.m.s. deviations	-	-
Protein bond lengths (Å)	0.004	0.005
Protein bond angles (°)	1.3	1.3
Ramachandran analysis	-	-
Favored (%)	98.7	98.7
Outliers (%)	0	0

* The data for the highest resolution shell is shown in parenthesis.

Overall, the structures of CagFbFP-A56P and CagFbFP-A95P are very similar to the original CagFbFP structure, with the root mean square deviation of the positions of the backbone atoms of ~0.2 Å in each case. Polder OMIT maps [54] confirm the identity of the introduced mutations (Figure S4). The LOV domain fold is not changed, and the two protomers form an antiparallel dimer, with the hydrophobic surfaces of β-sheets at the dimerization interface, as observed previously for CagFbFP [25]. The protein backbone structure is slightly altered around the mutation site in

CagFbFP-A56P, and essentially unchanged in CagFbFP-A95P (Figure 4). Interestingly, in both of the mutants, Ala58 and Asp59 are still in two alternative conformations. Most likely, the differences in the free energies of the two conformations were not affected by the introduced mutations.

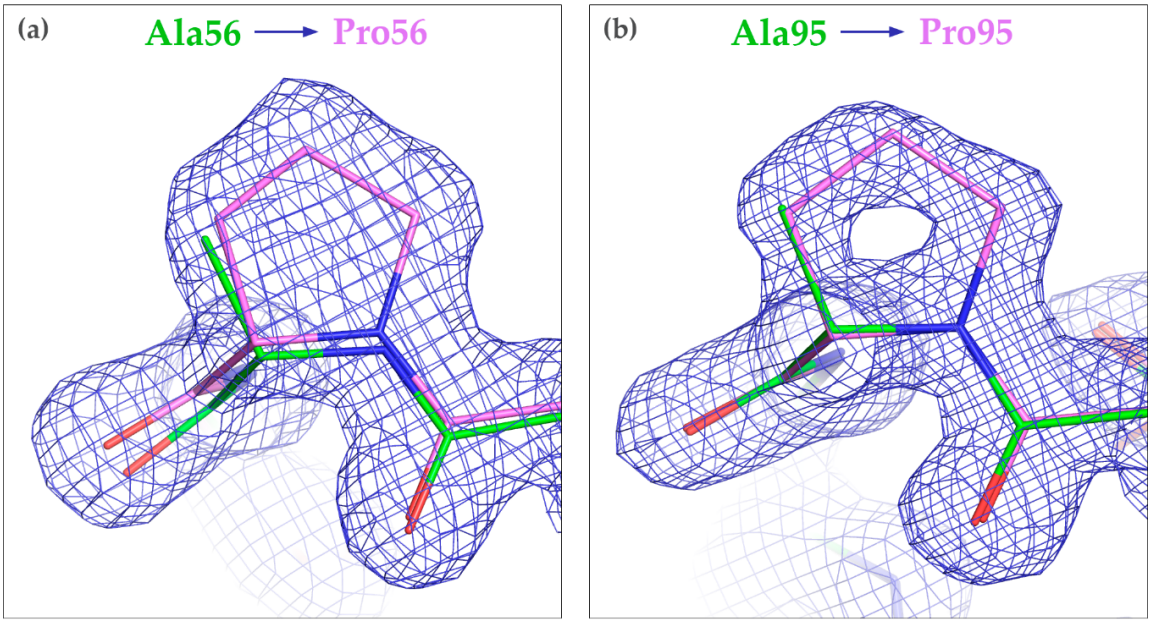

Figure 4. Effects of mutations A56P (**a**) and A95P (**b**) on the structure of CagFbFP. The backbone structure is essentially unchanged. The original structure of CagFbFP (PDB ID 6RHF) is shown in green, while the structures of the mutants are shown in magenta. $2F_o$-F_c electron densities (blue) are contoured at the level of 1.5 × r.m.s.

3.4. MD Simulations of the A58P Variant

To gain insight into the structure of the A58P variant, we conducted extensive MD simulations of the WT and mutated proteins. Both CagFbFP and CagFbFP-A58P preserve their overall structure in the simulations, as evidenced by the root mean square deviations of the atomic positions from the starting structure (Figure S5). Some of the simulations reveal relative motions of the protomers within the dimer (Figure S5), yet in each case, the structures of the individual protomers are unchanged (Figure S6). Analysis of backbone fluctuations (Figure S6) does not reveal any clear differences in the flexibility of the WT and A58P variants. However, we observed that the structure of the Aβ-Bβ loop was changed in the mutant (Figure 5). In particular, the peptide torsion angle φ is positive for Ala58 and negative for Pro58. The moderate displacement of the Aβ-Bβ loop in the A58P variant is likely the reason for its inability to form crystals, which is similar to the WT protein. Interestingly, we do not observe any interconversion between the two alternative conformations of Ala58 or Pro58. Because of this, it is not clear whether one or another conformation is preferred in the mutant.

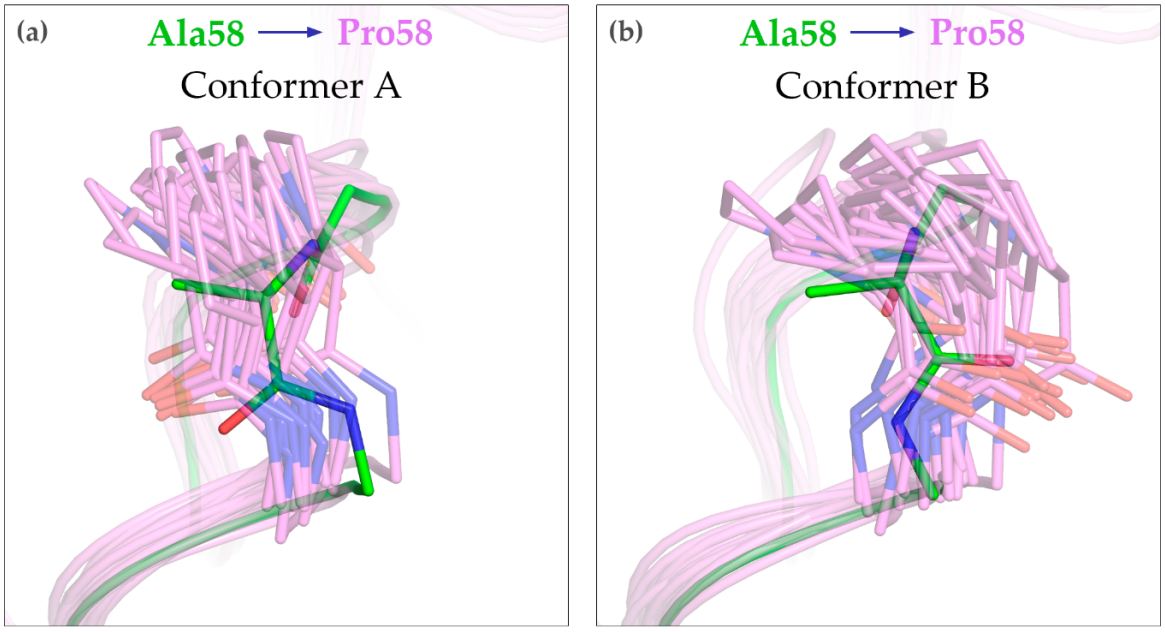

Figure 5. Structure of the conformers A (**a**) and B (**b**) of the Aβ-Bβ loop in the A58P mutant, as observed in MD simulations. The original structure of CagFbFP (PDB ID 6RHF) is shown in green, and trajectory snapshots of the A58P mutant are shown in magenta. 20 exemplary conformations for each alternative conformation of the Pro58 backbone are shown. Most notably, the position of the side chain C_β atom is different between Ala58 and Pro58.

4. Discussion

In this work, we analyzed the multiple sequence alignment of several light-oxygen-voltage (LOV) proteins and the available structural information to identify three amino acid positions in a thermostable LOV protein, CagFbFP, which could stabilize the protein when substituted with prolines. Two of the identified mutations, A56P and A58P, had a mostly neutral effect on the protein's stability, whereas the third one, A95P, moderately stabilized the protein, and also improved its refolding temperature. Complementary structural studies show that the structure of the mutated proteins remains essentially unchanged, although the structure of the Aβ-Bβ loop is slightly different in the A58P variant.

Previously, multiple studies have probed the effects of mutations on various properties of LOV domains. Earlier studies focused on the effects of mutating the conserved cysteine, which forms a covalent bond with the flavonoid cofactor during the photocycle, and some random mutations on flavin binding and photochemical reactivity [55,56]. Later studies probed the effects of mutations on other properties, particularly the absorption spectrum [47,50,57,58], photocycle lifetime [57,59], brightness of the cysteine-less variants [9,19,20,60], generation of radicals [15,61,62] and thermal stability [21–23]. Many of these mutations were rational, or could be rationalized after initial discovery, thus allowing one to apply the same principles to impart a different LOV domain with the desirable properties.

We expect that the effects of the proline substitutions may also be transferable to other LOV domains, since the rationales that we employed (consensus design [26–28], stabilization by prolines [27,28]) will hold. Interestingly, out of three LOV domain thermal stabilization studies [21–23], only one reported the proline substitutions [22], which were, however, not included in the most stable variant with multiple mutations. Thus, testing the effects of proline substitutions may be a complementary approach to recombination, and directed evolution to speed up the discovery of the most stable variants. We hope that our work will advance the development of efficient LOV-based tools for fluorescence microscopy and optogenetics.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4352/10/4/256/s1, Figure S1: Normalized absorption spectra of CagFbFP and its proline-substituted variants, Figure S2. Normalized fluorescence excitation spectra of CagFbFP and its proline-substituted variants, Figure S3. Normalized fluorescence emission spectra of CagFbFP and its proline-substituted variants, Figure S4. Omit (polder) maps for the mutants, (a) A56P and (b) A95P. The original structure of CagFbFP (PDB ID 6RHF) is shown in green, the structures of the mutants are shown in magenta. Polder electron density maps (green) are contoured at the level of 3 × r.m.s, Figure S5. Root mean square deviations of backbone atom positions as a function of time. Dimers of proteins harboring the conformers A and B of the residue 58 were simulated both for the WT and A58P variants for 3 times (runs 1–3). The values were averaged over 1 ns time intervals. Some trajectories, such as WT B run 1, display relatively high overall RMSD as a consequence of displacement of one protomer relative to another one. Structures of individual protomers are conserved well in all simulations (Figure S6), Figure S6. Root mean square fluctuations of backbone atom positions as a function of residue number during the last 40 ns of the simulations. Trajectory snapshots of each of the two protomers from each simulation were aligned and analyzed separately, and labeled chain A or chain B. No clear differences in the flexibility of different elements of the WT and A58P mutant variants are evident from the data.

Author Contributions: Conceptualization, A.R. and I.G.; investigation, A.R., V.V.N., I.M.G., A.Y., A.S., O.S., K.K., C.G., U.S., M.D.D., V.G., I.G.; writing—original draft preparation, A.R. and I.G.; writing—review and editing, A.R., M.D.D. and I.G.; visualization, V.V.N.; supervision, A.R. and I.G.; project administration, A.R. and I.G.; funding acquisition, A.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation, grant number 18-74-00092.

Acknowledgments: Atomic coordinates and structure factors for the reported crystal structures have been deposited in the Protein Data Bank under the accession codes 6Y7R (CagFbFP-A56P) and 6Y7U (CagFbFP-A95P). X-ray diffraction data were collected at the BioMAX beamline at MAX IV Laboratory (Lund, Sweden). Simulations were performed with computing resources granted by JARA-HPC from RWTH Aachen University under project JARA0065.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

FbFP	Flavin-based Fluorescent Protein
LOV	Light-Oxygen-Voltage
WT	Wild Type

References

1. Losi, A.; Gärtner, W. Solving Blue Light Riddles: New Lessons from Flavin-binding LOV Photoreceptors. *Photochem. Photobiol.* **2017**, *93*, 141–158. [CrossRef] [PubMed]
2. Glantz, S.T.; Carpenter, E.J.; Melkonian, M.; Gardner, K.H.; Boyden, E.S.; Wong, G.K.-S.; Chow, B.Y. Functional and topological diversity of LOV domain photoreceptors. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1442–E1451. [CrossRef] [PubMed]
3. Möglich, A. Signal transduction in photoreceptor histidine kinases. *Protein Sci.* **2019**, *28*, 1923–1946. [CrossRef] [PubMed]
4. Conrad, K.S.; Manahan, C.C.; Crane, B.R. Photochemistry of flavoprotein light sensors. *Nat. Chem. Biol.* **2014**, *10*, 801–809. [CrossRef]
5. Losi, A.; Gardner, K.H.; Möglich, A. Blue-Light Receptors for Optogenetics. *Chem. Rev.* **2018**, *118*, 10659–10709. [CrossRef]
6. Glantz, S.T.; Berlew, E.E.; Jaber, Z.; Schuster, B.S.; Gardner, K.H.; Chow, B.Y. Directly light-regulated binding of RGS-LOV photoreceptors to anionic membrane phospholipids. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E7720–E7727. [CrossRef] [PubMed]
7. Buckley, A.M.; Petersen, J.; Roe, A.J.; Douce, G.R.; Christie, J.M. LOV-based reporters for fluorescence imaging. *Curr. Opin. Chem. Biol.* **2015**, *27*, 39–45. [CrossRef]
8. Mukherjee, A.; Schroeder, C.M. Flavin-based fluorescent proteins: Emerging paradigms in biological imaging. *Curr. Opin. Biotechnol.* **2015**, *31*, 16–23. [CrossRef]
9. Drepper, T.; Eggert, T.; Circolone, F.; Heck, A.; Krauß, U.; Guterl, J.-K.; Wendorff, M.; Losi, A.; Gärtner, W.; Jaeger, K.-E. Reporter proteins for in vivo fluorescence without oxygen. *Nat. Biotechnol.* **2007**, *25*, 443–445. [CrossRef]

10. Kim, N.M.; Sinnott, R.W.; Sandoval, N.R. Transcription factor-based biosensors and inducible systems in non-model bacteria: Current progress and future directions. *Curr. Opin. Biotechnol.* **2020**, *64*, 39–46. [CrossRef]
11. Chia, H.E.; Marsh, E.N.G.; Biteen, J.S. Extending fluorescence microscopy into anaerobic environments. *Curr. Opin. Chem. Biol.* **2019**, *51*, 98–104. [CrossRef] [PubMed]
12. Ozbakir, H.F.; Anderson, N.T.; Fan, K.-C.; Mukherjee, A. Beyond the Green Fluorescent Protein: Biomolecular Reporters for Anaerobic and Deep-Tissue Imaging. *Bioconjug. Chem.* **2020**, *31*, 293–302. [CrossRef] [PubMed]
13. Shu, X.; Lev-Ram, V.; Deerinck, T.J.; Qi, Y.; Ramko, E.B.; Davidson, M.W.; Jin, Y.; Ellisman, M.H.; Tsien, R.Y. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol.* **2011**, *9*, e1001041. [CrossRef] [PubMed]
14. Westberg, M.; Etzerodt, M.; Ogilby, P.R. Rational design of genetically encoded singlet oxygen photosensitizing proteins. *Curr. Opin. Struct. Biol.* **2019**, *57*, 56–62. [CrossRef]
15. Endres, S.; Wingen, M.; Torra, J.; Ruiz-González, R.; Polen, T.; Bosio, G.; Bitzenhofer, N.L.; Hilgers, F.; Gensch, T.; Nonell, S.; et al. An optogenetic toolbox of LOV-based photosensitizers for light-driven killing of bacteria. *Sci. Rep.* **2018**, *8*, 1–14. [CrossRef]
16. Gushchin, I.; Gordeliy, V. Microbial Rhodopsins. *Subcell. Biochem.* **2018**, *87*, 19–56.
17. Pudasaini, A.; El-Arab, K.K.; Zoltowski, B.D. LOV-based optogenetic devices: Light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2015**, *2*, 18. [CrossRef]
18. Weber, A.M.; Kaiser, J.; Ziegler, T.; Pilsl, S.; Renzl, C.; Sixt, L.; Pietruschka, G.; Moniot, S.; Kakoti, A.; Juraschitz, M.; et al. A blue light receptor that mediates RNA binding and translational regulation. *Nat. Chem. Biol.* **2019**, *15*, 1085–1092. [CrossRef]
19. Chapman, S.; Faulkner, C.; Kaiserli, E.; Garcia-Mata, C.; Savenkov, E.I.; Roberts, A.G.; Oparka, K.J.; Christie, J.M. The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20038–20043. [CrossRef]
20. Christie, J.M.; Hitomi, K.; Arvai, A.S.; Hartfield, K.A.; Mettlen, M.; Pratt, A.J.; Tainer, J.A.; Getzoff, E.D. Structural tuning of the fluorescent protein iLOV for improved photostability. *J. Biol. Chem.* **2012**, *287*, 22295–22304. [CrossRef]
21. Song, X.; Wang, Y.; Shu, Z.; Hong, J.; Li, T.; Yao, L. Engineering a More Thermostable Blue Light Photo Receptor Bacillus subtilis YtvA LOV Domain by a Computer Aided Rational Design Method. *PLOS Comput. Biol.* **2013**, *9*, e1003129. [CrossRef] [PubMed]
22. Higgins, S.A.; Ouonkap, S.V.Y.; Savage, D.F. Rapid and Programmable Protein Mutagenesis Using Plasmid Recombineering. *ACS Synth. Biol.* **2017**, *6*, 1825–1833. [CrossRef] [PubMed]
23. Ko, S.; Hwang, B.; Na, J.-H.; Lee, J.; Jung, S.T. Engineered Arabidopsis Blue Light Receptor LOV Domain Variants with Improved Quantum Yield, Brightness, and Thermostability. *J. Agric. Food Chem.* **2019**, *67*, 12037–12043. [CrossRef] [PubMed]
24. Wingen, M.; Jaeger, K.-E.; Gensch, T.; Drepper, T. Novel Thermostable Flavin-binding Fluorescent Proteins from Thermophilic Organisms. *Photochem. Photobiol.* **2017**, *93*, 849–856. [CrossRef]
25. Nazarenko, V.V.; Remeeva, A.; Yudenko, A.; Kovalev, K.; Dubenko, A.; Goncharov, I.M.; Kuzmichev, P.; Rogachev, A.V.; Buslaev, P.; Borshchevskiy, V.; et al. A thermostable flavin-based fluorescent protein from Chloroflexus aggregans: A framework for ultra-high resolution structural studies. *Photochem. Photobiol. Sci. Off. J. Eur. Photochem. Assoc. Eur. Soc. Photobiol.* **2019**, *18*, 1793–1805. [CrossRef]
26. Porebski, B.T.; Buckle, A.M. Consensus protein design. *Protein Eng. Des. Sel.* **2016**, *29*, 245–251. [CrossRef]
27. Pezeshgi Modarres, H.; Mofrad, M.R.; Sanati-Nezhad, A. Protein thermostability engineering. *RSC Adv.* **2016**, *6*, 115252–115270. [CrossRef]
28. Kazlauskas, R. Engineering more stable proteins. *Chem. Soc. Rev.* **2018**, *47*, 9026–9045. [CrossRef]
29. Goldenzweig, A.; Fleishman, S.J. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **2018**, *87*, 105–129. [CrossRef]
30. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinforma Oxf. Engl.* **2007**, *23*, 2947–2948. [CrossRef]
31. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinforma. Oxf. Engl.* **2009**, *25*, 1189–1191. [CrossRef] [PubMed]

32. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5-9. [CrossRef] [PubMed]
33. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207–234. [CrossRef] [PubMed]
34. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef] [PubMed]
35. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *62*, 72–82. [CrossRef]
36. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *66*, 22–25. [CrossRef]
37. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [CrossRef]
38. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [CrossRef]
39. Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins Struct. Funct. Bioinforma* **2002**, *47*, 393–402. [CrossRef]
40. Krivov, G.G.; Shapovalov, M.V.; Dunbrack, R.L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins Struct. Funct. Bioinforma* **2009**, *77*, 778–795. [CrossRef]
41. Olsson, M.H.M.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537. [CrossRef] [PubMed]
42. Case, D.A.; Babin, V.; Berryman, J.; Betz, R.M.; Cai, Q.; Cerutti, D.S.; Cheatham, T.E., III; Darden, T.A.; Duke, R.E.; Gohlke, H.; et al. *Amber 14*; University of California: San Francisco, CA, USA, 2014.
43. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
44. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinforma* **2006**, *65*, 712–725. [CrossRef] [PubMed]
45. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [CrossRef]
46. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef] [PubMed]
47. Davari, M.D.; Kopka, B.; Wingen, M.; Bocola, M.; Drepper, T.; Jaeger, K.-E.; Schwaneberg, U.; Krauss, U. Photophysics of the LOV-Based Fluorescent Protein Variant iLOV-Q489K Determined by Simulation and Experiment. *J. Phys. Chem. B* **2016**, *120*, 3344–3352. [CrossRef] [PubMed]
48. DeLano, W.L. The PyMOL Molecular Graphics System. 2002. Available online: http://www.pymol.org (accessed on 28 March 2020).
49. Humphrey, W.; Dalke, A.; Schulten, K. VMD Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]
50. Wingen, M.; Potzkei, J.; Endres, S.; Casini, G.; Rupprecht, C.; Fahlke, C.; Krauss, U.; Jaeger, K.-E.; Drepper, T.; Gensch, T. The photophysics of LOV-based fluorescent proteins – new tools for cell biology. *Photochem. Photobiol. Sci.* **2014**, *13*, 875–883. [CrossRef]
51. Mukherjee, A.; Weyant, K.B.; Agrawal, U.; Walker, J.; Cann, I.K.O.; Schroeder, C.M. Engineering and characterization of new LOV-based fluorescent proteins from Chlamydomonas reinhardtii and Vaucheria frigida. *ACS Synth. Biol.* **2015**, *4*, 371–377. [CrossRef]
52. Tang, K.-H.; Barry, K.; Chertkov, O.; Dalin, E.; Han, C.S.; Hauser, L.J.; Honchak, B.M.; Karbach, L.E.; Land, M.L.; Lapidus, A.; et al. Complete genome sequence of the filamentous anoxygenic phototrophic bacterium Chloroflexus aurantiacus. *BMC Genom.* **2011**, *12*, 334. [CrossRef]
53. Gaisin, V.A.; Kalashnikov, A.M.; Grouzdev, D.S.; Sukhacheva, M.V.; Kuznetsov, B.B.; Gorlenko, V.M. Chloroflexus islandicus sp. nov., a thermophilic filamentous anoxygenic phototrophic bacterium from a geyser. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 1381–1386. [CrossRef] [PubMed]

54. Liebschner, D.; Afonine, P.V.; Moriarty, N.W.; Poon, B.K.; Sobolev, O.V.; Terwilliger, T.C.; Adams, P.D. Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. Sect. Struct. Biol.* **2017**, *73*, 148–157. [CrossRef] [PubMed]
55. Kottke, T.; Dick, B.; Fedorov, R.; Schlichting, I.; Deutzmann, R.; Hegemann, P. Irreversible photoreduction of flavin in a mutated Phot-LOV1 domain. *Biochemistry* **2003**, *42*, 9854–9862. [CrossRef] [PubMed]
56. Salomon, M.; Christie, J.M.; Knieb, E.; Lempert, U.; Briggs, W.R. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* **2000**, *39*, 9401–9410. [CrossRef] [PubMed]
57. Zayner, J.P.; Sosnick, T.R. Factors that control the chemistry of the LOV domain photocycle. *PLoS ONE* **2014**, *9*, e87074. [CrossRef]
58. Raffelberg, S.; Gutt, A.; Gärtner, W.; Mandalari, C.; Abbruzzetti, S.; Viappiani, C.; Losi, A. The amino acids surrounding the flavin 7a-methyl group determine the UVA spectral features of a LOV protein. *Biol. Chem.* **2013**, *394*, 1517–1528. [CrossRef]
59. Fettweiss, T.; Röllen, K.; Granzin, J.; Reiners, O.; Endres, S.; Drepper, T.; Willbold, D.; Jaeger, K.-E.; Batra-Safferling, R.; Krauss, U. Mechanistic Basis of the Fast Dark Recovery of the Short LOV Protein DsLOV from Dinoroseobacter shibae. *Biochemistry* **2018**, *57*, 4833–4847. [CrossRef]
60. Mukherjee, A.; Weyant, K.B.; Walker, J.; Schroeder, C.M. Directed evolution of bright mutants of an oxygen-independent flavin-binding fluorescent protein from Pseudomonas putida. *J. Biol. Eng.* **2012**, *6*, 20. [CrossRef]
61. Consiglieri, E.; Xu, Q.; Bregnhøj, M.; Westberg, M.; Ogilby, P.R.; Losi, A. Single mutation in a novel bacterial LOV protein yields a singlet oxygen generator. *Photochem. Photobiol. Sci.* **2019**, *18*, 2657–2660. [CrossRef]
62. Westberg, M.; Bregnhøj, M.; Etzerodt, M.; Ogilby, P.R. No Photon Wasted: An Efficient and Selective Singlet Oxygen Photosensitizing Protein. *J. Phys. Chem. B* **2017**, *121*, 9366–9371. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).