Title: Insights into the mechanisms of LOV domain color tuning from a set of high-resolution X-ray structures

Running title: Insights into LOV domain color tuning

Alina Remeeva¹, Vera V. Nazarenko¹, Kirill Kovalev^{1,2,3,4,5}, Ivan Goncharov¹, Anna Yudenko¹, Roman Astashkin^{1,2}, Valentin Gordeliy^{1,2,3,4}, 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

* Correspondence: ivan.gushchin@phystech.edu.

Acknowledgements: The diffraction data were collected on the beamlines ID23-1, ID29 and ID30B at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to local contacts at ESRF for providing assistance in using these beamlines, to Daria Melikhova for help with expression of the Q148R variant, and to Ulrich Krauss for fruitful discussions. This research was funded by the Russian Science Foundation, grant number 18-74-00092.

Abstract

Light-oxygen-voltage (LOV) domains are widespread photosensory modules that can be used in fluorescence microscopy, optogenetics and controlled production of reactive oxygen species. All of the currently known LOV domains have absorption maxima in the range of ~440 to ~450 nm, and it is not clear whether they can be shifted significantly using mutations. Here, we have generated a panel of LOV domain variants by mutating the key chromophore-proximal glutamine amino acid of a thermostable flavin based fluorescent protein CagFbFP (Gln148) to asparagine, aspartate, glutamate, histidine, lysine and arginine. Absorption spectra of all of the mutants are blue-shifted, with the maximal shift of 8 nm observed for the Q148H variant. While CagFbFP and its Q148N/D/E variants are not sensitive to pH, Q148H/K/R reveal a moderate red shift induced by acidic pH. To gain further insight, we determined high resolution crystal structures of all of the mutants studied at the resolutions from 1.07 Å for Q148D to 1.63 Å for Q148R. Whereas in some of the variants, the amino acid 148 remains in the vicinity of the flavin, in Q148K, Q148R and partially Q148D, the C-terminus of the protein unlatches and the side chain of the residue 148 is reoriented away from the chromophore. Our results explain the absence of color shifts from replacing Gln148 with charged amino acids and pave the way for rational design of color-shifted flavin based fluorescent proteins.

Keywords: LOV domain, color tuning, X-ray crystallography, mutagenesis, protein design

Introduction

Light-oxygen-voltage (LOV) domains are widespread photosensory modules found in proteins from organisms belonging to all three kingdoms of life^{1,2}. They bind different flavins such as flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) and respond to illumination with blue light^{1,3,4}. At the moment, more than 6000 LOV sequences may be found in genomic databases², and many of them have been characterized in detail. They generate considerable interest since they regulate a wide variety of processes such as phototropism in plants, circadian expression in fungi and lifestyle and virulence in bacteria^{1,5}. Also, LOV domains have been engineered for applications in fluorescence microscopy⁶⁻¹¹, optogenetics^{12,13} and controlled production of reactive oxygen species^{14,15}. Among the major advantages of using LOV domains are their small size (around 110 amino acids) and independence of maturation on molecular oxygen or any other exogenous molecules.

The optical properties of the LOV domains are dictated by the flavin chromophores. The proteins have characteristic absorption spectra with the maxima at 370 and 450 nm, respectively.

Absorption of a photon by the flavin launches a cascade of photochemical reactions called photocycle, where two widely conserved nearby amino acids, cysteine and glutamine, play a major role^{1,3,4}. Cysteine's sulfur forms a covalent bond with flavin's C4a atom in the early steps of the photocycle, which leads to reorganization of the flavin's hydrogen bonding pattern and, in particular, to change in the conformation of the glutamine^{1,3,4}. Reorientation of the glutamine side chain further relays the photosignal^{1,3,4}. Eventually, the cysteine-flavin bond breaks, and the protein returns thermally to the initial state^{1,3,4}.

The role of different amino acids surrounding flavins in the photocycle of LOV domains has been extensively studied. Replacing the conserved cysteine with another sulfur-containing amino acid, methionine, leads to irreversible light-induced formation of the methionine-FMN adduct¹⁶. Replacing the cysteine with non-reactive amino acids abolishes the adduct formation^{17,18} and enhances the fluorescent properties of the protein^{6,7}, because the adducted form is non-fluorescent. Yet, some LOV domains lacking the conserved cysteine are still able to generate a response to illumination¹⁹.

While the cysteine-FMN interaction is probably the most important one for the photocycle in most of the LOV domains, other amino acids regulate the duration of the photocycle to a high degree. The photocycles can last from seconds to days, and can be effectively regulated by mutations^{20–23}.

Early mechanistic studies have also highlighted the role of a conserved glutamine, situated closely to FMN and the cysteine, as another key amino acid for signal transduction^{24–27}. Replacement of the glutamine with a leucine has been used to remove the hydrogen bond with FMN, and impairs the signal transduction^{22,27–30}. On the contrary, replacement with asparagine locks the LOV domain in an illuminated-like state and has been used as its mimic^{30–33}. Finally, replacements of the glutamine with an aspartate or histidine have profound effects on the LOV domain photochemistry^{22,34}.

Yet, none of the tested mutations had a profound effect on the absorption spectra of the LOV domains. While the absorption in the UVA region (320–400 nm) is affected by the amino acids surrounding the flavin's 7a-methyl group³⁵, the visible light spectra of the mutants are usually only slightly blue-shifted, with the maximum shift of 10 nm for Gln→Leu replacements^{22,29,30,36}.

This highlights an intriguing question: can the visible light absorption of LOV domains be color-tuned at all? The native LOV domain-containing proteins seem all to have very similar spectra, with the absorption maxima of different proteins all being within several nanometers of each other (depending slightly on the equipment used in experiments), whereas other photoproteins display a remarkable variety of colors. For example, animal and microbial rhodopsins may have absorption maxima in essentially the whole range of the visible light from at least 400 to 700 nm, without any covalent modification of the retinal chromophore^{37,38}. Beta-barrel fluorescent proteins such as the green fluorescent protein (GFP), or related blue, cyan, yellow and red FPs, have absorption maxima from ~350 to ~650 nm, although in this case the chromophores do differ^{39–41}.

The common rationales in the studies of color tuning are that the spectra may be affected either by steric interactions (changing the geometry of the chromophore, which is not always possible), or electronic interactions (such as π -stacking, hydrogen bonding, or electrostatic)⁴². Since the isoalloxazine moiety in FMN and FAD is formed by three conjugated unsaturated heterocycles, substantially changing its geometry does not seem to be a viable idea. On the other hand, affecting its spectral properties by changing the charge distributions of the surrounding amino acids seems possible. In 2015, Khrenova *et al.* suggested that replacing the conserved glutamine 489 with a lysine in a widely used LOV-based fluorescent protein iLOV may result in a ~50 nm red shift⁴³. However, the experiments and calculations by Davari *et al.* showed that the Q489K mutant of iLOV is unexpectedly blue-shifted, and that the lysine's ϵ -amino group is unlikely to remain in the close vicinity of the FMN's isoalloxazine ring⁴⁴. Consequently, Khrenova *et al.* suggested, based on simulations, that introducing a second substitution may fix the lysine side chain and result in a red shift⁴⁵. Overall, electrostatic spectral tuning maps can be used to

estimate the effects of the redistribution of charges around the chromophore on its spectra, in particularly for FMN-binding proteins⁴².

An alternative route to introducing mutations is modification of the chromophore. Replacement of FMN with roseoflavin strongly alters the LOV domain spectra and shifts the absorption maximum to 485 nm⁴⁶. Replacement of FMN with 7-methyl-8-chloro-riboflavin has a similar effect⁴⁷, whereas replacement with 7-bromo-8-methyl-riboflavin does not have a substantial effect, and replacement with lumichrome causes a blue shift of the absorption maximum to 421 nm⁴⁷. Replacement of FMN with 5-deazaflavin mononucleotide similarly results in blue shifts to 423 nm⁴⁸ or even ~400 nm⁴⁹, whereas 1-deazaflavin mononucleotide produces a remarkable red shift to ~530 nm⁴⁹. Finally, the absorption maxima of a LOV domain 8-bromo- and 8-trifluoromethyl-substituted flavins are relatively not shifted⁵⁰. Although not yet confirmed experimentally, computational studies indicate that thioflavins or fluorinated flavins may also be used^{51,52}, and that red shifts of up to 100 nm may also be attainable by combining modified chromophores with protein mutations⁵³. Still, we must note that using the modified chromophores for applications of LOV domains does not seem to be particularly practical. While the natural chromophores such as FMN or FAD are ubiquitous in most living cells, the modified chromophores will have to be synthesized and supplemented exogenously for *in vivo* applications; they may be toxic for the cells, or their binding may be weaker compared to omnipresent natural flavins.

Therefore, a search for mutations that may shift the spectra of LOV domains, both to understand the color tuning mechanisms of flavins, and for applications, is fully justified. Among the amino acids surrounding the chromophore in LOVs, most are tightly constrained by their neighbors and/or the flavin, so that drastic mutations are very likely to result in loss of binding or

misfolding. The only residue shown to tolerate some mutations to polar and charged amino acids is the glutamine in the vicinity of flavin atoms N4 and O5, which is involved in signaling. Previously, mutations of this glutamine were suggested to confer large spectral shifts^{43,45}. Consequently, in this work, we have generated a panel of variants of the thermostable LOV protein CagFbFP, where the conservative glutamine is substituted with different polar and charged amino acids (aspartate, glutamate, lysine, arginine, asparagine and histidine) that might strongly affect the electrostatic environment of the flavin and thus change the absorption spectrum significantly. We showed that the mutations are tolerated and the protein remains folded and binds flavin, measured the absorption and fluorescence spectra of the generated variants, and determined their high-resolution crystallographic structures that allow us to rationalize the effects of mutations.

Materials and Methods

Cloning, protein expression and purification

The nucleotide sequence encoding the original variant CagFbFP (amino acid sequence derived from the gene Cagg_3753, UniProt accession code B8GAY9, *Chloroflexus aggregans* strain MD-66 / DSM 9485, with the mutation C85A and C-terminal 6×His tag) was obtained as a synthetic gene from Evrogen (Russia) and introduced into the pET11 expression vector (Novagen, Merck, Germany) via NdeI and BamHI restriction sites as described previously⁵⁴. Mutations Q148N, Q148D, Q148E, Q148H, Q148K and Q148R were introduced into the CagFbFP gene using polymerase chain reaction.

CagFbFP and its mutants were expressed in *Escherichia coli* strain C41 (DE3) as described previously⁵⁴. All cultures were grown in ZYP-5052 autoinducing medium⁵⁵ supplemented with 100 µg mL⁻¹ ampicillin at 37 °C. 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. Protein-containing fractions were concentrated to 30 mg mL⁻¹ for crystallization.

Spectroscopic methods

Thermal stability of CagFbFP and all mutated variants were measured using the Rotor-Gene Q real-time PCR cycler (Qiagen, Germany) with fluorescence excitation at 470 nm and fluorescence detection at 510 nm. Absorption and fluorescence spectra were recorded using Synergy™ H4 Hybrid Microplate Reader (BioTek, USA). Emission was measured between 470

nm and 700 nm in increments of 1 nm, with excitation fixed at 450 nm. Excitation was measured between 250 nm and 500 nm in increments of 1 nm, the signal was detected at 510 nm.

Crystallization

All of the crystals were obtained using the sitting drop vapor diffusion approach. Crystallization trials were set up using the NT8 robotic system (Formulatrix, USA). The crystals were grown at 22 °C. The drops contained 150 nL concentrated protein solution and 100 nL reservoir solution.

CagFbFP-Q148N was crystallized with 0.06 M Magnesium chloride hexahydrate, 0.06 M Calcium chloride dihydrate, 0.1M MES monohydrate, 0.1M Imidazole, pH 6.5, 20% v/v PEG 500 MME, 10% w/v PEG 20000.

Crystals of CagFbFP-Q148D and CagFbFP-Q148E were obtained with the following precipitant: 0.1 M DL-Glutamic acid monohydrate, 0.1 M DL-Alanine, 0.1 M Glycine, 0.1 M DL-Lysine monohydrochloride, 0.1 M DL-Serine, 0.1M MES monohydrate, 0.1M Imidazole, pH 6.5, 20% v/v PEG 500 MME, 10% w/v PEG 20000.

Crystals of CagFbFP-Q148H were obtained using 0.5 M Lithium sulfate monohydrate, 15 % w/v PEG 8000 as precipitant.

As described in the accompanying article⁵⁶, CagFbFP-Q148K was crystallized using the precipitant solution containing 0.1 M MES monohydrate pH 6.5 and 12% w/v PEG 20000.

Morpholine-free structures were obtained with 0.5 M Ammonium Sulfate, 1 M Lithium sulfate monohydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6.

Crystals of CagFbFP-Q148R were obtained with the following precipitant: 0.12 M 1,6-Hexanediol, 0.12 M 1-Butanol, 0.12 M 1,2-Propanediol, 0.12 M 2-Propanol, 0.12 M 1,4-Butanediol, 0.12 M 1,3-Propanediol, 0.1M MES monohydrate, 0.1M Imidazole, pH 6.5, 20% v/v PEG 500 MME, 10% w/v PEG 20000 as precipitant.

To prevent drying and for cryoprotection, precipitant solutions with 20% glycerol were added to the crystallization drops during crystal harvesting. The crystals were harvested using micromounts and then flash-cooled and stored in liquid nitrogen.

Acquisition and treatment of diffraction data

The diffraction data were collected at 100 K on the beamlines ID23-1, ID29 and ID30B at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Diffraction images were processed using XDS⁵⁷. POINTLESS and AIMLESS⁵⁸ were used to merge, scale and assess the quality of the data, as well as to convert intensities to structure factor amplitudes and to generate Free-R labels. As described in the accompanying article⁵⁶, datasets 6YX6 and 6YXC revealed strong anisotropy and were processed using STARANISO⁵⁹ using the following criterion for determining the diffraction-limit surface: $\langle I_{\text{mean}}/\sigma I_{\text{mean}} \rangle$ of more than 2.

Structure determination and refinement

All structures were solved using molecular replacement with MOLREP⁶⁰. CagFbFP structure⁵⁴ (PDB ID 6RHF) was used as a search model. The resulting solution was refined manually using Coot⁶¹ and REFMAC5⁶².

Results

General properties of the Q148X variants

Introduction of strongly charged amino acids at the chromophore-binding site may affect its binding, and also destabilize the protein overall. Therefore, we speculated that starting with an initially stable protein, which can also produce well-diffracting crystals, is desirable for this study, and chosen the recently identified thermostable LOV domain from *Chloroflexus aggregans*⁵⁴ as a framework. We used the C85A mutant, dubbed CagFbFP (*Chloroflexus aggregans* flavin based fluorescent protein), which is fluorescent and does not undergo a photocycle under illumination. Below, the designation WT is used for the original CagFbFP variant.

All of the tested variants (Q148N, Q148D, Q148E, Q148H, Q148K and Q148R) were expressed in *E. coli* in the flavin-bound form and could be purified and crystallized. We tested the effects of the substitutions on thermal stability of CagFbFP by measuring the dependence of fluorescence on temperature (Figures 1 and S1, and Table 1) and found that all of the substitutions destabilized the protein. Similarly to CagFbFP, the variants had complex melting curves, often with two different melting transitions. Among the variants tested, Q148K is the most detrimental mutation, with the melting temperature reduced by ~20 °C.

Spectroscopic characterization of the Q148X mutants

Having generated the Q148X mutants, we measured their absorption and fluorescence spectra (Figure 2 and Table 1). We note that positions of absorption and fluorescence excitation maxima are slightly different, with the former being blue-shifted by 2-3 nm compared to the latter.

Interestingly, none of the variants revealed a red shift in the absorption or fluorescence excitation spectrum; substitutions of the glutamine 148 with potentially positively charged amino acids resulted in the largest blue shifts, for as much as 8 nm for Q148H. The blue shift of the Q148K variant is smaller than that of the homologous iLOV Q489K mutant⁴⁴. While the fluorescence emission maxima of the Q148D and Q148E variants are formally red-shifted, the spectra are overall widened and cannot be considered red-shifted.

Structural analysis of the Q148X mutants

Following the spectroscopic characterization, we attempted crystallization of all of the generated variants. The original variant, CagFbFP, has been crystallized previously, and its structure has been determined at the resolution of 1.07 Å⁵⁴. The new variants are also readily crystallizable, and X-ray structures were solved for all of them at the resolutions ranging from 1.07 Å for Q148D to 1.63 Å for Q148R. All of the crystals belonged to the same space group as CagFbFP crystals, P2₁2₁2. For Q148H and Q148K, crystals belonging to C2 and P2₁, respectively, were also obtained. In comparison to the P2₁2₁2 CagFbFP crystals, P2₁2₁2 Q148K and Q148R crystals had slightly extended unit cell dimension a (~57.5 Å instead of ~53.5 Å). The structures have been deposited to the Protein Data Bank (PDB), and respective PDB identifiers are used for convenience below. Crystallographic data collection and refinement statistics are presented in Supplementary Table S1. Crystals of the Q148K variant in the space group P2₁ contained two dimers in the asymmetric unit; all other crystals contained a single dimer in the asymmetric unit. In many cases this was helpful as more data could be obtained from a single crystal.

Since aspartate, glutamate, histidine, lysine and arginine side changes can be protonated or deprotonated depending on pH and their environment, we tested the effects of pH on the

fluorescence spectra (Figures 3 and S3-S7). The parent protein, CagFbFP, didn't reveal any differences in the spectra collected at different pH values, similarly to EcFbFP, which retained its fluorescence in a wide range of pH values⁶³. Fluorescence spectra of the variants Q148N,D,E were also essentially identical at pH 3 and pH 8 (Figures S3-5). On the contrary, Q148H,K,R fluorescence spectra were slightly red-shifted at pH 3 compared to pH 8 (Figures 3 and S6-7). A possible explanation for this will be given below in the context of the structural data.

Structure of CagFbFP

To put the new findings in context, we describe first the structure of the original variant, CagFbFP⁵⁴ (Figure 4A). The chromophore FMN is coordinated by, among others, Asn117 and Asn127 side chains, and Gln148 occupies two alternative conformations, in both cases forming a hydrogen bond to the FMN's atom O4. Whereas one of the conformations is similar to that observed in most other LOV domains, the second one is unusual and may be a consequence of steric interactions with the nearby Ile52 side chain. The asymmetric unit contains two CagFbFP molecules, which are almost identical in structure. Yet, there is a small difference: in the chain A, the C-terminus occupies a single (latched) conformation, and separates the Gln148-harboring pocket from the solvent. However, in the chain B, two alternative conformations of the C-terminus are observed. In the second one, Thr149 is displaced, and Asp150 is not very well resolved in the electron density maps.

Structure of the Q148N variant

Similarly to glutamine, asparagine has a carboxamide side chain, and we begin our comparisons with the structure of the Q148N variant. In both chains of the dimer, Asn148 side chain occupies the same position (Figures 4B,C). Its atom N_{δ2} is within hydrogen bonding distance of the

FMN's O4. Whereas the protein's C-terminus is in a single latched conformation in the chain A (Figure 4B), it is clearly in two conformations in the chain B: one is like that in the chain A, and another is different (Figure 4C). Overall, the structure is identical to that of the original variant CagFbFP in all regards except for the position of the residue 148 side chain. Consequently, it is not clear how the mutation may convert the protein into illuminated-like state, as it does in *Avena sativa* phototropin 1 LOV2³⁰; perhaps, it changes the dynamics of the proteins in solution.

Structure of the Q148E variant

As opposed to glutamine, glutamate has a carboxylic acid side chain that can become negatively charged if deprotonated. Surprisingly, we observe that Glu148 side chain occupies almost the same position in the Q148E variant as Gln148 does in the WT CagFbFP, except that it loses the hydrogen bond to FMN (Figure 4D). The distances from the O_{ε1} and O_{ε2} carboxyl atoms to FMN's O4 are 4.3 and 5.1 Å, respectively, and there are no evident hydrogen bonding partners in the vicinity of the side chain. In both chains, the C-termini are clearly in the same single latched conformations. In the chain A, residues 149-155, including His154 and His155 of 6×His purification tag, are well resolved.

Structure of the Q148D variant

Aspartate has a shorter side chain compared to glutamate, and this might be the reason for the observed differences in the structures (Figure 4E,F). Asp148 and C-terminus occupy different positions in the two protein chains in the asymmetric unit. In the chain A (Figure 4E), the C-terminus is latched, and positions of residues 149-155, including His154 and His155 of 6×His purification tag are resolved similarly to those in the Q148E variant. Asp148 occupies two alternative conformations: one within hydrogen bonding distance of the FMN's O4 and

Asn127's N_{δ2}, and another one rotated and without any evident hydrogen bonding partners in the vicinity. The hydrogen-bonded conformation is similar to that observed by Kopka *et al.* in a simulation of the Q489D variant of iLOV with protonated D489³⁴. In the chain B, surprisingly, the C-terminus is fully unlatched, and Asp148 side chain is oriented away from FMN (Figure 4F), as observed in a simulation of the Q489D variant of iLOV with deprotonated D489³⁴.

Structure of the Q148K variant

As opposed to the previously discussed amino acids, lysine has an amino group in its side chain. It is positively charged and protonated when exposed to the solvent at physiological conditions, but its pK_a may be as low as 5.3⁶⁵ when it is confined within a protein (as observed, for example, in some microbial rhodopsins⁶⁶). Previously, both Khrenova *et al.*⁴³ and Davari *et al.*⁴⁴ proposed that the homologous Lys489 in the Q489K mutant of iLOV is protonated, and its ε-amino group remains within 6 Å of the FMN's O4.

As described in the accompanying article⁵⁶, we were able to obtain three types of Q148K variant crystals, all using the precipitant solutions with pH of 6.5. Unexpectedly, and contrary to the results of previous simulations^{43,44}, in all of the obtained models the C-terminus of the protein is clearly unlatched, and the side chain of Lys148 is extended away from FMN (Figure 5A,B,C).

Surprisingly, the highest-resolution structure (PDB ID 6YX4, determined at the resolution of 1.36 Å), revealed a six-membered ring in a chair conformation close to FMN (Figure 5A). No such molecule has been added to the sample during purification or crystallization. Yet, we noted that the precipitant solution contained 2-(N-morpholino)ethanesulfonic acid (MES), and morpholine moiety fitted perfectly in the density near FMN with the occupancies of 0.7 and 0.5 for the protein chains A and B, respectively. Thus, we assume that MES was degraded to

morpholine during storage of the precipitant solution or during crystallization, and the morpholine bound to the pocket in CagFbFP. Similar presumable degradation of a related buffer component 3-(N-morpholino)propanesulfonic acid (MOPS) and binding of morpholine to the protein was observed by Sooriyaarachchi *et al.* during crystallization of *Plasmodium falciparum* 1-deoxy-d-xylulose-5-phosphate deductoisomerase⁶⁷ (PDB ID 5JO0).

Although we did not observe any influence of morpholine on the Q148K spectra or thermal stability, it could not be excluded that it affected the conformation of the protein. Consequently, we crystallized the protein using another precipitant solution in two crystal forms⁶⁸. P₂₁2₁2 crystals (PDB ID 6YX6) diffracted anisotropically, and we applied ellipsoidal truncation using the resolution cut-offs of 1.5 and 2.05 Å. Chain A revealed some residual density in the vicinity of FMN observed at the level of the polder maps of $3.5 \times \text{r.m.s.}$ (Figure 5B, green) that couldn't be ascribed to any specific molecule. Chain B revealed only a very weak density at the same level (Figure 5B, red). P₂₁ crystals (PDB ID 6YXB) produced almost isotropic diffraction patterns and contained two protein dimers within the asymmetric unit. Chains A, B and D contained very weak densities at the same position (Figure 5C, green, red and blue), whereas chain C didn't have any. The densities are likely due to disordered water molecules present in the pockets of some of the proteins in the crystal. Therefore, we conclude that irrespective of the buffer, the C-terminus of the Q148K variant is unlatched and Lys148 occupies the outermost position. These findings are corroborated by the structure of the Q489K variant of iLOV described in the accompanying article⁵⁶.

Structure of the Q148R variant

Arginine has a guanidinium group in its side chain that is also normally protonated and has an even higher pK_a in solution (~ 12.5) compared to lysine (~ 10.5). Similarly to Lys148 in the Q148K variant, Arg148 in Q148R occupies the outermost position (Figure 5D). There are two density blobs in the vicinity of FMN, a smaller one within hydrogen-bonding distance of N5 and O4 atoms, and a bigger one at the position of morpholine and other densities in the Q148K variant. The smaller density likely reflects partial occupancy of a water molecule, whereas the bigger one cannot be ascribed to any particular molecules in the precipitant solution. The C-terminus is in an intermediate conformation similar to that of the morpholine-bound Q148K variant. Interestingly, Asp150 side chain is nearby and may interact with the Arg148 side chain (Figure 5D).

Structure of the Q148H variant

Histidine is the most difficult side chain for analysis and modeling since it has several protonation states, tautomeric forms, and in solution its pK_a is around 6.5. We obtained two crystal forms of Q148H, belonging to space groups $P2_12_12$ (PDB ID 6YWQ) and $C2$ (PDB ID 6YWR). In the first structure (PDB ID 6YWQ), in both chains, the C-terminus is in the latched conformation, and His148 side chain does not contact FMN (Figure 5E). Unlike in any of the previously discussed structures, Asn127 occupies two alternative conformations; in the new one it is hydrogen-bonded to His148 (Figure 5E). In the second structure (PDB ID 6YWR), the chain A adopts the same conformation, with possibly a weakly occupied second conformation of His148 attained via rotation around the C_β - C_γ torsion. However, chain B has a number of unique features: besides the two alternative conformations for Asn127, two conformations are also

observed for His148 (the new one is hydrogen-bonded to FMN), and Asp150 is reoriented towards His148 (Figure 5F).

Discussion

Herein, we have generated and characterized a panel of 6 variants of a thermostable LOV domain CagFbFP, where the conservative glutamine amino acid has been replaced by asparagine, glutamate, aspartate, histidine, lysine or arginine. All of the tested mutations were destabilizing and affected the melting and refolding temperatures strongly (Table 1), which highlights the importance of starting with an initially stable protein⁶⁹. Surprisingly, none of the mutations produced a red shift in the absorbance or fluorescence excitation spectra, and the observed blue shifts were smaller than those resulting from Gln→Leu replacement in other proteins³⁶. Thus, significant color tuning would probably require two or more simultaneous mutations^{45,56}.

We also determined 9 high-resolution crystal structures of the mutants (Table S1). The structures highlight the overall robustness of the LOV domain fold, and the conformations of various amino acids at the position 148. In the original version CagFbFP, residue 148 adopts two conformations: one is classical and observed in many other LOV domain structures, and the other is perpendicular to it⁵⁴. This second conformation was implicated in the signaling mechanism in ZEITLUPE⁷⁰ (dubbed the “exposed” conformation) and in *Bacillus subtilis* YtvA⁷¹. Interestingly, in the mutated variants, the side chains are in most cases observed in a single conformation (Figures 4 and 5). In particular, the structures reveal that the potentially strongly charged amino acids (Lys, Arg, and partially Asp), which might have affected the spectroscopic properties of the flavin chromophore, are in fact expelled from the protein core

(Figures 4 and 5). This explains the seeming contradiction between the notion that the strong positive charges in the region of the flavin's N5 and O4 atoms should result in a strong red shift⁴²⁻⁴⁴, and the experimental data revealing modest blue shifts (reported by Davari *et al.*⁴⁴, here, and also in the accompanying article⁵⁶). Potentially negatively charged amino acids also do not produce a substantial blue shift, contrary to what might have been expected from the electrostatic tuning maps^{42,72}. Evidently, Glu148 and Asp148 either do not interact with FMN, or interact with it in the protonated form (hydrogen-bonded conformation in Figure 4E). These findings underscore the importance of using experimental structural data for prediction of properties of photoactive proteins⁷³.

Our data also shows that Q148H, K and R mutants are mildly pH-sensitive, revealing red shifts at lower pH (Figures 3 and S6-7). This sensitivity comes from the mutations, because CagFbFP and the spectra of the variants Q148N, D and E do not depend on pH (Figures S3-5). At the same time, Lys148 and Arg148 in Q148K and Q148R are highly likely to be protonated and charged already at physiological pH and thus should not be affected by lower pH values. We believe that lower pH values may result in neutralization of Asp150, which might serve as a counterion for Lys148 or Arg148 (Figure 5A-D), and consequently in a higher probability of lysine/arginine side chain coming in the vicinity of FMN and affecting its spectra.

As can be seen from the structures, the reorientation of the residue 148 side chain away from the chromophore in Q148K, Q148R and partially Q148D variants is possible due to unlatching of the C-terminus of the LOV domain. Thus, it might be possible to keep the residue 148 side chain nearby the chromophore either by creating favorable interactions in the pocket, as suggested by Khrenova *et al.*⁴⁵, or by stabilizing the C-terminus in the latched conformation. The former is demonstrated in the accompanying article⁵⁶. On the other hand, unlatching and opening of FMN-

proximal pocket presents an interesting possibility of influencing the LOV domain spectra by small molecules, or even designing LOV-based fluorescent reporters that change their properties upon binding of a particular molecule.

Finally, our study highlights CagFbFP as a useful framework for engineering of new LOV domain variants. It clearly tolerates drastic alterations in the FMN-binding pocket, and even then, it still remains relatively thermostable. CagFbFP can be split for detection of protein-protein interactions⁷⁴. It can be further thermostabilized by replacing Ala95 with a proline⁷⁵ for testing even more destabilizing alterations. Finally, CagFbFP can be crystallized and produce high-resolution structures on par with the best resolution structures available for other LOV proteins (1.17 Å for miniSOG¹⁵ and 1.2 Å for iLOV⁷⁶).

Supporting Information. Supporting information (Supporting Table S1 and Supporting Figures S1-S8) are available.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

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

Abbreviations

FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; FP, fluorescent protein; LOV domain, light oxygen voltage domain; PDB, Protein Data Bank; WT, wild type.

References

- (1) Losi, A.; Gärtner, W. Solving Blue Light Riddles: New Lessons from Flavin-Binding LOV Photoreceptors. *Photochemistry and Photobiology* **2017**, *93* (1), 141–158. <https://doi.org/10.1111/php.12674>.
- (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. *PNAS* **2016**, *113* (11), E1442–E1451. <https://doi.org/10.1073/pnas.1509428113>.
- (3) Möglich, A. Signal Transduction in Photoreceptor Histidine Kinases. *Protein Science* **2019**, *28* (11), 1923–1946. <https://doi.org/10.1002/pro.3705>.
- (4) Conrad, K. S.; Manahan, C. C.; Crane, B. R. Photochemistry of Flavoprotein Light Sensors. *Nat. Chem. Biol.* **2014**, *10* (10), 801–809. <https://doi.org/10.1038/nchembio.1633>.
- (5) Herrou, J.; Crosson, S. Function, Structure and Mechanism of Bacterial Photosensory LOV Proteins. *Nature Reviews Microbiology* **2011**, *9* (10), 713–723. <https://doi.org/10.1038/nrmicro2622>.
- (6) 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 Biotech* **2007**, *25* (4), 443–445. <https://doi.org/10.1038/nbt1293>.
- (7) 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. *PNAS* **2008**, *105* (50), 20038–20043. <https://doi.org/10.1073/pnas.0807551105>.
- (8) Mukherjee, A.; Schroeder, C. M. Flavin-Based Fluorescent Proteins: Emerging Paradigms in Biological Imaging. *Curr. Opin. Biotechnol.* **2015**, *31*, 16–23. <https://doi.org/10.1016/j.copbio.2014.07.010>.
- (9) Chia, H. E.; Marsh, E. N. G.; Biteen, J. S. Extending Fluorescence Microscopy into Anaerobic Environments. *Current Opinion in Chemical Biology* **2019**, *51*, 98–104. <https://doi.org/10.1016/j.cbpa.2019.05.008>.
- (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. *Current Opinion in Biotechnology* **2020**, *64*, 39–46. <https://doi.org/10.1016/j.copbio.2019.09.009>.
- (11) 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. <https://doi.org/10.1016/j.cbpa.2015.05.011>.
- (12) 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. <https://doi.org/10.3389/fmolb.2015.00018>.
- (13) Losi, A.; Gardner, K. H.; Möglich, A. Blue-Light Receptors for Optogenetics. *Chem. Rev.* **2018**, *118* (21), 10659–10709. <https://doi.org/10.1021/acs.chemrev.8b00163>.
- (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. <https://doi.org/10.1016/j.sbi.2019.01.025>.
- (15) Endres, S.; Wingen, M.; Torra, J.; Ruiz-González, R.; Polen, T.; Bosio, G.; Bitzenhofer, N. L.; Hilgers, F.; Gensch, T.; Nonell, S.; Jaeger, K.-E.; Drepper, T. An Optogenetic Toolbox of LOV-Based Photosensitizers for Light-Driven Killing of Bacteria. *Scientific Reports* **2018**, *8* (1), 1–14. <https://doi.org/10.1038/s41598-018-33291-4>.

- (16) 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* (33), 9854–9862. <https://doi.org/10.1021/bi034863r>.
- (17) 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* (31), 9401–9410. <https://doi.org/10.1021/bi000585+>.
- (18) Swartz, T. E.; Corchnoy, S. B.; Christie, J. M.; Lewis, J. W.; Szundi, I.; Briggs, W. R.; Bogomolni, R. A. The Photocycle of a Flavin-Binding Domain of the Blue Light Photoreceptor Phototropin. *J. Biol. Chem.* **2001**, *276* (39), 36493–36500. <https://doi.org/10.1074/jbc.M103114200>.
- (19) Yee, E. F.; Diensthuber, R. P.; Vaidya, A. T.; Borbat, P. P.; Engelhard, C.; Freed, J. H.; Bittl, R.; Möglich, A.; Crane, B. R. Signal Transduction in Light–Oxygen–Voltage Receptors Lacking the Adduct-Forming Cysteine Residue. *Nature Communications* **2015**, *6* (1), 1–10. <https://doi.org/10.1038/ncomms10079>.
- (20) Jentzsch, K.; Wirtz, A.; Circolone, F.; Drepper, T.; Losi, A.; Gärtner, W.; Jaeger, K.-E.; Krauss, U. Mutual Exchange of Kinetic Properties by Extended Mutagenesis in Two Short LOV Domain Proteins from *Pseudomonas Putida*. *Biochemistry* **2009**, *48* (43), 10321–10333. <https://doi.org/10.1021/bi901115z>.
- (21) Zoltowski, B. D.; Vaccaro, B.; Crane, B. R. Mechanism-Based Tuning of a LOV Domain Photoreceptor. *Nature Chemical Biology* **2009**, *5* (11), 827–834. <https://doi.org/10.1038/nchembio.210>.
- (22) Zayner, J. P.; Sosnick, T. R. Factors That Control the Chemistry of the LOV Domain Photocycle. *PLoS ONE* **2014**, *9* (1), e87074. <https://doi.org/10.1371/journal.pone.0087074>.
- (23) Fettweiss, T.; Röllén, 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 Shiba*. *Biochemistry* **2018**, *57* (32), 4833–4847. <https://doi.org/10.1021/acs.biochem.8b00645>.
- (24) Crosson, S.; Moffat, K. Photoexcited Structure of a Plant Photoreceptor Domain Reveals a Light-Driven Molecular Switch. *The Plant Cell* **2002**, *14* (5), 1067–1075. <https://doi.org/10.1105/tpc.010475>.
- (25) Fedorov, R.; Schlichting, I.; Hartmann, E.; Domratcheva, T.; Fuhrmann, M.; Hegemann, P. Crystal Structures and Molecular Mechanism of a Light-Induced Signaling Switch: The Phot-LOV1 Domain from *Chlamydomonas Reinhardtii*. *Biophysical Journal* **2003**, *84* (4), 2474–2482. [https://doi.org/10.1016/S0006-3495\(03\)75052-8](https://doi.org/10.1016/S0006-3495(03)75052-8).
- (26) Halavaty, A. S.; Moffat, K. N- and C-Terminal Flanking Regions Modulate Light-Induced Signal Transduction in the LOV2 Domain of the Blue Light Sensor Phototropin 1 from *Avena Sativa*. *Biochemistry* **2007**, *46* (49), 14001–14009. <https://doi.org/10.1021/bi701543e>.
- (27) Zoltowski, B. D.; Schwerdtfeger, C.; Widom, J.; Loros, J. J.; Bilwes, A. M.; Dunlap, J. C.; Crane, B. R. Conformational Switching in the Fungal Light Sensor Vivid. *Science* **2007**, *316* (5827), 1054–1057. <https://doi.org/10.1126/science.1137128>.
- (28) Jones, M. A.; Feeney, K. A.; Kelly, S. M.; Christie, J. M. Mutational Analysis of Phototropin 1 Provides Insights into the Mechanism Underlying LOV2 Signal Transmission. *J. Biol. Chem.* **2007**, *282* (9), 6405–6414. <https://doi.org/10.1074/jbc.M605969200>.

- (29) Nozaki, D.; Iwata, T.; Ishikawa, T.; Todo, T.; Tokutomi, S.; Kandori, H. Role of Gln1029 in the Photoactivation Processes of the LOV2 Domain in Adiantum Phytochrome3. *Biochemistry* **2004**, *43* (26), 8373–8379. <https://doi.org/10.1021/bi0494727>.
- (30) Nash, A. I.; Ko, W.-H.; Harper, S. M.; Gardner, K. H. A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration. *Biochemistry* **2008**, *47* (52), 13842–13849. <https://doi.org/10.1021/bi801430e>.
- (31) Raffelberg, S.; Mansurova, M.; Gärtner, W.; Losi, A. Modulation of the Photocycle of a LOV Domain Photoreceptor by the Hydrogen-Bonding Network. *J. Am. Chem. Soc.* **2011**, *133* (14), 5346–5356. <https://doi.org/10.1021/ja1097379>.
- (32) Diensthuber, R. P.; Engelhard, C.; Lemke, N.; Gleichmann, T.; Ohlendorf, R.; Bittl, R.; Möglich, A. Biophysical, Mutational, and Functional Investigation of the Chromophore-Binding Pocket of Light-Oxygen-Voltage Photoreceptors. *ACS Synth. Biol.* **2014**, *3* (11), 811–819. <https://doi.org/10.1021/sb400205x>.
- (33) 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. U.S.A.* **2018**, *115* (33), E7720–E7727. <https://doi.org/10.1073/pnas.1802832115>.
- (34) Kopka, B.; Magerl, K.; Savitsky, A.; Davari, M. D.; Röllén, K.; Bocola, M.; Dick, B.; Schwaneberg, U.; Jaeger, K.-E.; Krauss, U. Electron Transfer Pathways in a Light, Oxygen, Voltage (LOV) Protein Devoid of the Photoactive Cysteine. *Sci Rep* **2017**, *7* (1), 13346. <https://doi.org/10.1038/s41598-017-13420-1>.
- (35) 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* (11), 1517–1528. <https://doi.org/10.1515/hsz-2013-0163>.
- (36) 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. *Photochemical & Photobiological Sciences* **2014**, *13* (6), 875–883. <https://doi.org/10.1039/C3PP50414J>.
- (37) Ernst, O. P.; Lodowski, D. T.; Elstner, M.; Hegemann, P.; Brown, L. S.; Kandori, H. Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chem. Rev.* **2014**, *114* (1), 126–163. <https://doi.org/10.1021/cr4003769>.
- (38) Gushchin, I.; Gordeliy, V. Microbial Rhodopsins. *Subcell. Biochem.* **2018**, *87*, 19–56. https://doi.org/10.1007/978-981-10-7757-9_2.
- (39) Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.* **2010**, *90* (3), 1103–1163. <https://doi.org/10.1152/physrev.00038.2009>.
- (40) Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends in Biochemical Sciences* **2017**, *42* (2), 111–129. <https://doi.org/10.1016/j.tibs.2016.09.010>.
- (41) Lambert, T. J. FPbase: A Community-Editable Fluorescent Protein Database. *Nature Methods* **2019**, *16* (4), 277–278. <https://doi.org/10.1038/s41592-019-0352-8>.
- (42) Orozco-Gonzalez, Y.; Kabir, M. P.; Gozem, S. Electrostatic Spectral Tuning Maps for Biological Chromophores. *J. Phys. Chem. B* **2019**, *123* (23), 4813–4824. <https://doi.org/10.1021/acs.jpcc.9b00489>.

- (43) Khrenova, M. G.; Nemukhin, A. V.; Domratcheva, T. Theoretical Characterization of the Flavin-Based Fluorescent Protein ILOV and Its Q489K Mutant. *J. Phys. Chem. B* **2015**, *119* (16), 5176–5183. <https://doi.org/10.1021/acs.jpcc.5b01299>.
- (44) 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* (13), 3344–3352. <https://doi.org/10.1021/acs.jpcc.6b01512>.
- (45) Khrenova, M. G.; Meteleshko, Y. I.; Nemukhin, A. V. Mutants of the Flavoprotein ILOV as Prospective Red-Shifted Fluorescent Markers. *J. Phys. Chem. B* **2017**, *121* (43), 10018–10025. <https://doi.org/10.1021/acs.jpcc.7b07533>.
- (46) Mathes, T.; Vogl, C.; Stolz, J.; Hegemann, P. In Vivo Generation of Flavoproteins with Modified Cofactors. *Journal of Molecular Biology* **2009**, *385* (5), 1511–1518. <https://doi.org/10.1016/j.jmb.2008.11.001>.
- (47) Arinkin, V.; Granzin, J.; Röllén, K.; Krauss, U.; Jaeger, K.-E.; Willbold, D.; Batra-Safferling, R. Structure of a LOV Protein in Apo-State and Implications for Construction of LOV-Based Optical Tools. *Scientific Reports* **2017**, *7* (1), 1–10. <https://doi.org/10.1038/srep42971>.
- (48) Kalvaitis, M. E.; Johnson, L. A.; Mart, R. J.; Rizkallah, P.; Allemann, R. K. A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation. *Biochemistry* **2019**, *58* (22), 2608–2616. <https://doi.org/10.1021/acs.biochem.9b00255>.
- (49) Silva-Junior, M. R.; Mansurova, M.; Gärtner, W.; Thiel, W. Photophysics of Structurally Modified Flavin Derivatives in the Blue-Light Photoreceptor YtvA: A Combined Experimental and Theoretical Study. *ChemBioChem* **2013**, *14* (13), 1648–1661. <https://doi.org/10.1002/cbic.201300217>.
- (50) Mansurova, M.; Simon, J.; Salzmann, S.; Marian, C. M.; Gärtner, W. Spectroscopic and Theoretical Study on Electronically Modified Chromophores in LOV Domains: 8-Bromo- and 8-Trifluoromethyl-Substituted Flavins. *ChemBioChem* **2013**, *14* (5), 645–654. <https://doi.org/10.1002/cbic.201200670>.
- (51) Bracker, M.; Dinkelbach, F.; Weingart, O.; Kleinschmidt, M. Impact of Fluorination on the Photophysics of the Flavin Chromophore: A Quantum Chemical Perspective. *Phys. Chem. Chem. Phys.* **2019**, *21* (19), 9912–9923. <https://doi.org/10.1039/C9CP00805E>.
- (52) Marian, C. M.; Nakagawa, S.; Rai-Constapel, V.; Karasulu, B.; Thiel, W. Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region: Thioflavins As Potential Cofactors of Photoswitches. *J. Phys. Chem. B* **2014**, *118* (7), 1743–1753. <https://doi.org/10.1021/jp4098233>.
- (53) I. Meteleshko, Y.; V. Nemukhin, A.; G. Khrenova, M. Novel Flavin-Based Fluorescent Proteins with Red-Shifted Emission Bands: A Computational Study. *Photochemical & Photobiological Sciences* **2019**, *18* (1), 177–189. <https://doi.org/10.1039/C8PP00361K>.
- (54) Nazarenko, V. V.; Remeeva, A.; Yudenko, A.; Kovalev, K.; Dubenko, A.; Goncharov, I. M.; Kuzmichev, P.; Rogachev, A. V.; Buslaev, P.; Borshchevskiy, V.; Mishin, A.; Dhoke, G. V.; Schwaneberg, U.; Davari, M. D.; Jaeger, K.-E.; Krauss, U.; Gordeliy, V.; Gushchin, I. A Thermostable Flavin-Based Fluorescent Protein from *Chloroflexus Aggregans*: A Framework for Ultra-High Resolution Structural Studies. *Photochem. Photobiol. Sci.* **2019**, *18* (7), 1793–1805. <https://doi.org/10.1039/c9pp00067d>.

- (55) Studier, F. W. Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expression and Purification* **2005**, *41* (1), 207–234. <https://doi.org/10.1016/j.pep.2005.01.016>.
- (56) Röllen, K.; Granzin, J.; Remeeva, A.; Davari, M. D.; Gensch, T.; Nazarenko, V. V.; Kovalev, K.; Bogorodskiy, A.; Borshchevskiy, V.; Hemmer, S.; Schwaneberg, U.; Gordeliy, V.; Jaeger, K.-E.; Batra-Safferling, R.; Gushchin, I.; Krauss, U. Structural and Mechanistic Insight into Spectral Tuning in Flavin-Binding Fluorescent Proteins. *bioRxiv* **2021**, 2021.01.08.425906. <https://doi.org/10.1101/2021.01.08.425906>.
- (57) Kabsch, W. XDS. *Acta Crystallographica Section D Biological Crystallography* **2010**, *66* (2), 125–132. <https://doi.org/10.1107/S0907444909047337>.
- (58) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallographica Section D Biological Crystallography* **2005**, *62* (1), 72–82. <https://doi.org/10.1107/S0907444905036693>.
- (59) Tickle, I. J.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. *STARANISO*; Global Phasing Ltd.: Cambridge, United Kingdom.
- (60) Vagin, A.; Teplyakov, A. Molecular Replacement with MOLREP. *Acta Crystallographica Section D Biological Crystallography* **2009**, *66* (1), 22–25. <https://doi.org/10.1107/S0907444909042589>.
- (61) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Cryst D* **2010**, *66* (4), 486–501. <https://doi.org/10.1107/S0907444910007493>.
- (62) 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 Crystallographica Section D Biological Crystallography* **2011**, *67* (4), 355–367. <https://doi.org/10.1107/S0907444911001314>.
- (63) Rupprecht, C.; Wingen, M.; Potzkei, J.; Gensch, T.; Jaeger, K.-E.; Drepper, T. A Novel FbFP-Based Biosensor Toolbox for Sensitive in Vivo Determination of Intracellular PH. *Journal of Biotechnology* **2017**, *258*, 25–32. <https://doi.org/10.1016/j.jbiotec.2017.05.006>.
- (64) 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 D Struct Biol* **2017**, *73* (Pt 2), 148–157. <https://doi.org/10.1107/S2059798316018210>.
- (65) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; E, B. G.-M. Large Shifts in PKa Values of Lysine Residues Buried inside a Protein. *PNAS* **2011**, *108* (13), 5260–5265. <https://doi.org/10.1073/pnas.1010750108>.
- (66) Petrovskaya, L. E.; Balashov, S. P.; Lukashev, E. P.; Imasheva, E. S.; Gushchin, I. Y.; Dioumaev, A. K.; Rubin, A. B.; Dolgikh, D. A.; Gordeliy, V. I.; Lanyi, J. K.; others. ESR—A Retinal Protein with Unusual Properties from *Exiguobacterium Sibiricum*. *Biochemistry (Moscow)* **2015**, *80* (6), 688–700.
- (67) Sooriyaarachchi, S.; Chofor, R.; Risseuw, M. D. P.; Bergfors, T.; Pouyez, J.; Dowd, C. S.; Maes, L.; Wouters, J.; Jones, T. A.; Van Calenbergh, S.; Mowbray, S. L. Targeting an Aromatic Hotspot in *Plasmodium Falciparum* 1-Deoxy-d-Xylulose-5-Phosphate Reductoisomerase with β -Arylpropyl Analogues of Fosmidomycin. *ChemMedChem* **2016**, *11* (18), 2024–2036. <https://doi.org/10.1002/cmdc.201600249>.
- (68) Röllen, K.; Granzin, J.; Remeeva, A.; Davari, M. D.; Gensch, T.; Nazarenko, V. V.; Kovalev, K.; Hemmer, S.; Schwaneberg, U.; Gordeliy, V.; Jaeger, K.-E.; Batra-Safferling, R.; Gushchin, I.; Krauss, U. Rational Design and Structure of Blue- and Red-Shifted Flavin-

Binding Fluorescent Proteins. *Submitted, available in “Supporting Information for Review Only.”*

- (69) Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein Stability Promotes Evolvability. *PNAS* **2006**, *103* (15), 5869–5874. <https://doi.org/10.1073/pnas.0510098103>.
- (70) Pudasaini, A.; Shim, J. S.; Song, Y. H.; Shi, H.; Kiba, T.; Somers, D. E.; Imaizumi, T.; Zoltowski, B. D. Kinetics of the LOV Domain of ZEITLUPE Determine Its Circadian Function in Arabidopsis. *eLife* **2017**, *6*, e21646. <https://doi.org/10.7554/eLife.21646>.
- (71) Polverini, E.; Schackert, F. K.; Losi, A. Interplay among the “Flipping” Glutamine, a Conserved Phenylalanine, Water and Hydrogen Bonds within a Blue-Light Sensing LOV Domain. *Photochem. Photobiol. Sci.* **2020**, *19* (7), 892–904. <https://doi.org/10.1039/D0PP00082E>.
- (72) Kabir, M. P.; Orozco-Gonzalez, Y.; Gozem, S. Electronic Spectra of Flavin in Different Redox and Protonation States: A Computational Perspective on the Effect of the Electrostatic Environment. *Phys. Chem. Chem. Phys.* **2019**, *21* (30), 16526–16537. <https://doi.org/10.1039/C9CP02230A>.
- (73) Nemukhin, A. V.; Grigorenko, B. L.; Khrenova, M. G.; Krylov, A. I. Computational Challenges in Modeling of Representative Bioimaging Proteins: GFP-Like Proteins, Flavoproteins, and Phytochromes. *J. Phys. Chem. B* **2019**, *123* (29), 6133–6149. <https://doi.org/10.1021/acs.jpcc.9b00591>.
- (74) Yudenko, A.; Smolentseva, A.; Maslov, I.; Semenov, O.; Goncharov, I. M.; Nazarenko, V. V.; Maliar, N. L.; Borshchevskiy, V.; Gordeliy, V.; Remeeva, A.; Gushchin, I. Rational Design of a Split Flavin-Based Fluorescent Reporter. *ACS Synth. Biol.* **2021**, *10* (1), 72–83. <https://doi.org/10.1021/acssynbio.0c00454>.
- (75) Remeeva, A.; Nazarenko, V. V.; Goncharov, I. M.; Yudenko, A.; Smolentseva, A.; Semenov, O.; Kovalev, K.; Gülbahar, C.; Schwaneberg, U.; Davari, M. D.; Gordeliy, V.; Gushchin, I. Effects of Proline Substitutions on the Thermostable LOV Domain from *Chloroflexus Aggregans*. *Crystals* **2020**, *10* (4), 256. <https://doi.org/10.3390/cryst10040256>.
- (76) 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* (26), 22295–22304. <https://doi.org/10.1074/jbc.M111.318881>.

Table 1. Properties of glutamine-substituted CagFbFP variants.

Variant	T _{m1} , °C	T _{m2} , °C	T _r , °C	λ _{ex} , nm	λ _{em} , nm
WT	67.9 ± 0.3	81.3 ± 0.3	65.4 ± 0.3	450	498
Q148N	62.2 ± 1.0	73.6 ± 0.3	61.7 ± 0.3	447	496
Q148D	55.7 ± 0.3	69.0 ± 0.7	45.3 ± 0.7***	449	501
Q148E	58.3 ± 0.6	72.4 ± 0.9	49.2 ± 1.3	449	505
Q148H	64.3 ± 0.3	73.5 ± 0.2	65.4 ± 0.3	444	492
Q148K	56.0 ± 0.5	n.d.**	55.6 ± 1.9***	446	492
Q148R	61.6 ± 0.4	n.d.**	48.7 ± 2.1***	447	499

* The errors are standard deviations of the peak positions in five melting experiments.

** For these variants, presence and position of the second melting transition is not clear.

*** Only a small fraction is refolded, the temperature is not reliable.

Figure 1. Dependence of CagFbFP variants' fluorescence on temperature during heating.

A) Temperature-induced decay of fluorescence. The values are normalized to 100% at 25 °C, the temperature in the beginning of the experiment. B) Derivatives of the fluorescence traces. Each experiment was conducted independently for five times, and the data were averaged for plotting. The troughs observed at 48 °C are due to the instrument and are observed in all samples. Characteristic unfolding temperatures are summarized in Table 1. Refolding data is shown in Figure S1.

Figure 2. Spectroscopic properties of CagFbFP variants at pH 8. A) Absorption spectra. B) Fluorescence excitation spectra. C) Fluorescence emission spectra. The data are summarized in Table 1. Absorption and excitation data for the wavelength range 250-500 nm are presented in Supporting Figure S2.

Figure 3. Dependence of CagFbFP Q148H fluorescence excitation and emission spectra on pH.

All of the structures reveal the same fold of the proteins. Residues 48-148 are well resolved in all of the structures, and their backbones occupy essentially identical positions. The C-terminus, starting from the residue 149, has sometimes different conformations, as do the amino acids Asn127 and Asn/Asp/Glu/His/Lys/Arg148. The obtained structures are depicted in Figures 4 and 5, and exemplary $2F_o - F_c$ and omit (polder^{64}) maps may be found in Figure S8.

Figure 4. Structure of the flavin-binding pocket in CagFbFP and its Q148N, Q148E and Q148D mutants. Positions of CagFbFP's Gln148 and C-terminus are shown in all panels in grey for reference. Putative hydrogen bonding distances are indicated in ångströms. A) Structure of CagFbFP that was determined previously⁵⁴. Note that Gln148 was observed to occupy two alternative positions. B) Structure of the Q148N mutant, chain A. Asn148 forms a hydrogen bond with FMN. C-terminus occupies the same position as in the original variant. C) Structure of the Q148N mutant, chain A. Two alternative conformations of the C-terminus are observed: unlatched (shown) and latched, as in panel B (not shown). D) Structure of the Q148E mutant. No evident hydrogen bonding partners are observed for the Glu148 side chain. E) Structure of the Q148D mutant, chain A. Two alternative conformations of the Asp148 side chain are observed: in the first one, it forms a hydrogen bond with FMN, similarly to Asn148 in the Q148N mutant;

in the second one, no evident hydrogen bonding partners are observed. F) Structure of the Q148D mutant, chain B. Asp148 side chain is reoriented away from FMN; C-terminus in unlatched.

Figure 5. Structure of the flavin-binding pocket in Q148K, Q148R and Q148H mutants of CagFbFP. Positions of CagFbFP's Gln148 and latched and unlatched C-terminus are shown in grey for reference. A) Structure of the Q148K mutant with a bound morpholine molecule (blue). Polder omit map⁶⁴ is contoured at the level of $4 \times \text{r.m.s.}$ (green). Lys148 side chain is reoriented away from FMN; C-terminus in unlatched, but differently from what is observed for Q148N and Q148D mutants. Position of Asp150 side chain cannot be determined from the available electron densities. B) Structure of the Q148K mutant without morpholine (space group $P2_12_12$). Lys148 side chain is reoriented away from FMN; C-terminus in unlatched. Residual polder omit densities (green for chain A and red for chain B, contoured at the level of $3.5 \times \text{r.m.s.}$) are observed near FMN. C) Structure of the Q148K mutant without morpholine (space group $P2_12_12$). Residual polder omit densities (green for chain A, red for chain B, blue for chain D,

contoured at the level of $3.5 \times \text{r.m.s.}$) are observed near FMN. Lys148 side chain is reoriented away from FMN; C-terminus in unlatched. D) Structure of the Q148R mutant. Polder omit map⁶⁴ is contoured at the level of $3 \times \text{r.m.s.}$ (green). Arg148 side chain is reoriented away from FMN; C-terminus in unlatched, but differently from what is observed for Q148N and Q148D mutants. E) Structure of the Q148H mutant (space group C2). Two alternative conformations for Asn127 are observed: similar to the one observed in other variants, and a new one hydrogen-bonded to the N_{δ1} atom of His148. F) Structure of the Q148H mutant (space group P2₁2₁2). Two alternative conformations for Asn127 are observed. Two alternative conformations for His148 are also observed: similar to the one observed in the C2 crystals, and one where His148 is hydrogen-bonded to FMN. Asp150 is reoriented towards His148.