Electronic Supplementary Material (ESI) for Photochemical & Photobiological Sciences.

Thermostable flavin-based fluorescent protein from *Chloroflexus aggregans*: a framework for ultra-high resolution structural studies

Vera V. Nazarenko, Alina Remeeva, Anna Yudenko, Kirill Kovalev, Anton Dubenko, Ivan M. Goncharov, Pavel Kuzmichev, Andrey V. Rogachev, Pavel Buslaev, Valentin Borshchevskiy, Alexey Mishin, Gaurao V. Dhoke, Ulrich Schwaneberg, Mehdi D. Davari, Karl-Erich Jaeger, Ulrich Krauss, Valentin Gordeliy, Ivan Gushchin*

* E-mail for correspondence: ivan.gushchin@phystech.edu

Supplementary information

Supplementary Table S1

Supplementary Figures S1-S5

Table S1. Frequencies of aminoacids at the position of CagFbFP's Ile52. The most frequent occurrences are highlighted. The data are based on the dataset by Glantz *et al.*, 2016.

LOV group	**Total**	**Val**	**Ile**	**Leu**	**Thr**	**Ser**	**Cys**	**Ala**	**Others**
Archaea	156	16%	3.2%	1.9%	71.2%	4.5%	-	3.2%	-
Bacteria	1146	45.1%	22.9%	12%	12.5%	2.3%	1.5%	2.4%	1.4%
Fungi	645	50.2%	3.1%	6%	2.6%	0.3%	33.2%	3.7%	0.9%
Protists single	373	45.3%	9.7%	3.8%	16.6%	1.6%	19.6%	1.6%	1.9%
Protists tandem LOV1	169	81.7%	3.6%	1.2%	5.9%	-	2.4%	5.3%	-
Protists tandem LOV2	169	71.6%	4.1%	1.2%	1.2%	1.2%	20.7%	-	-
Land plants single	2689	92.9%	3.2%	1.2%	0.9%	0.04%	0.2%	0.9%	0.7%
Land plants tandem LOV1	1587	98.6%	0.2%	0.7%	-	-	-	0.5%	-
Land plants tandem LOV2	1587	99.9%	0.1%	-	-	-	0.1%	-	-

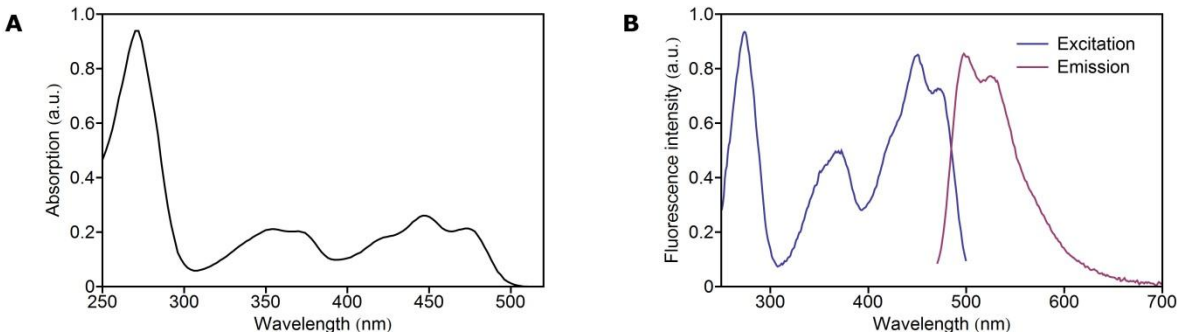

Figure S1. CagFbFP absorption (A) and fluorescence excitation and fluorescence emission (B) spectra.

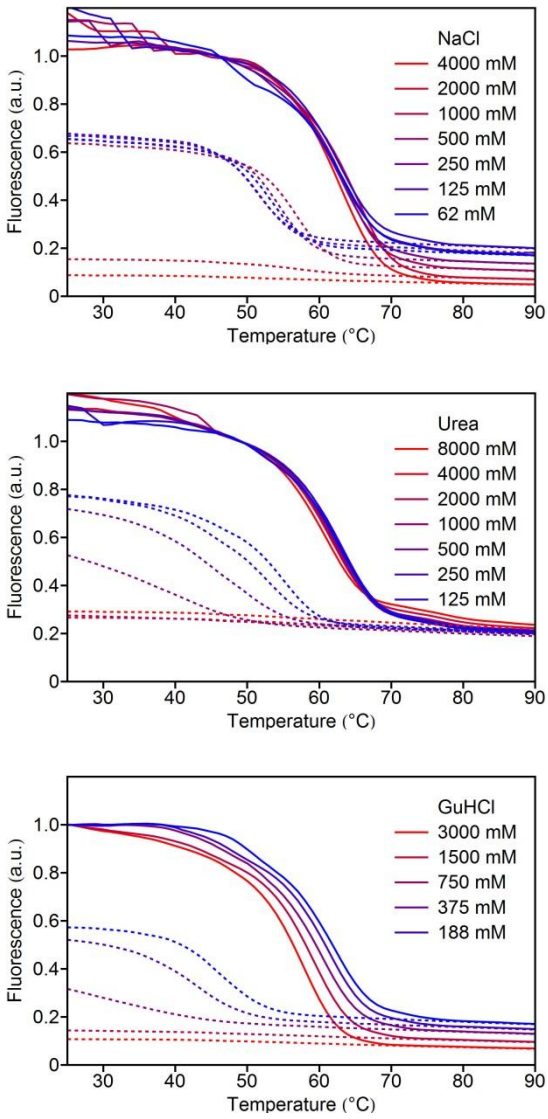

Figure S2. Temperature-induced unfolding (solid lines) and refolding (dotted lines) of CagFbFP in different solutions. Initial parts of the unfolding curves are noisy due to automatic fluorescence signal gain adjustments. To obtain the samples, purified CagFbFP was mixed with concentrated stock solutions. The measurements were performed using a Rotor-Gene Q real-time PCR cycler.

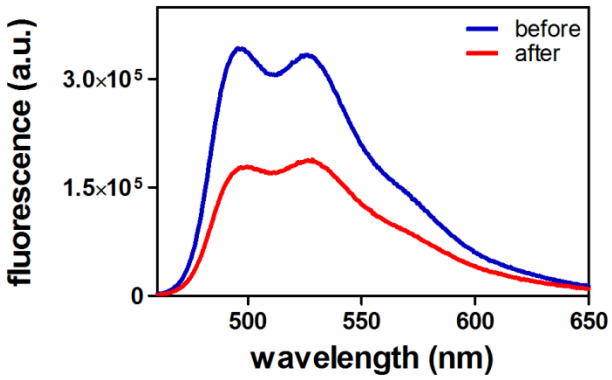

Figure S3. Exemplary fluorescence emission spectra of CagFbFP before and after melting and refolding. Samples were equilibrated for 5 minutes at 20 °C, an emission spectrum was recorded (before, blue line) and the sample was subsequently heated to 100 °C (1 °C/min ramp rate). After reaching the final temperature, the sample was cooled to 20 °C and a second emission spectrum was recorded (after, red line). In terms of fluorescence intensity at 495 nm approx. 50% of the initial value is recovered after melting. Only one data set out of three independent measurements on three independently prepared samples is shown.

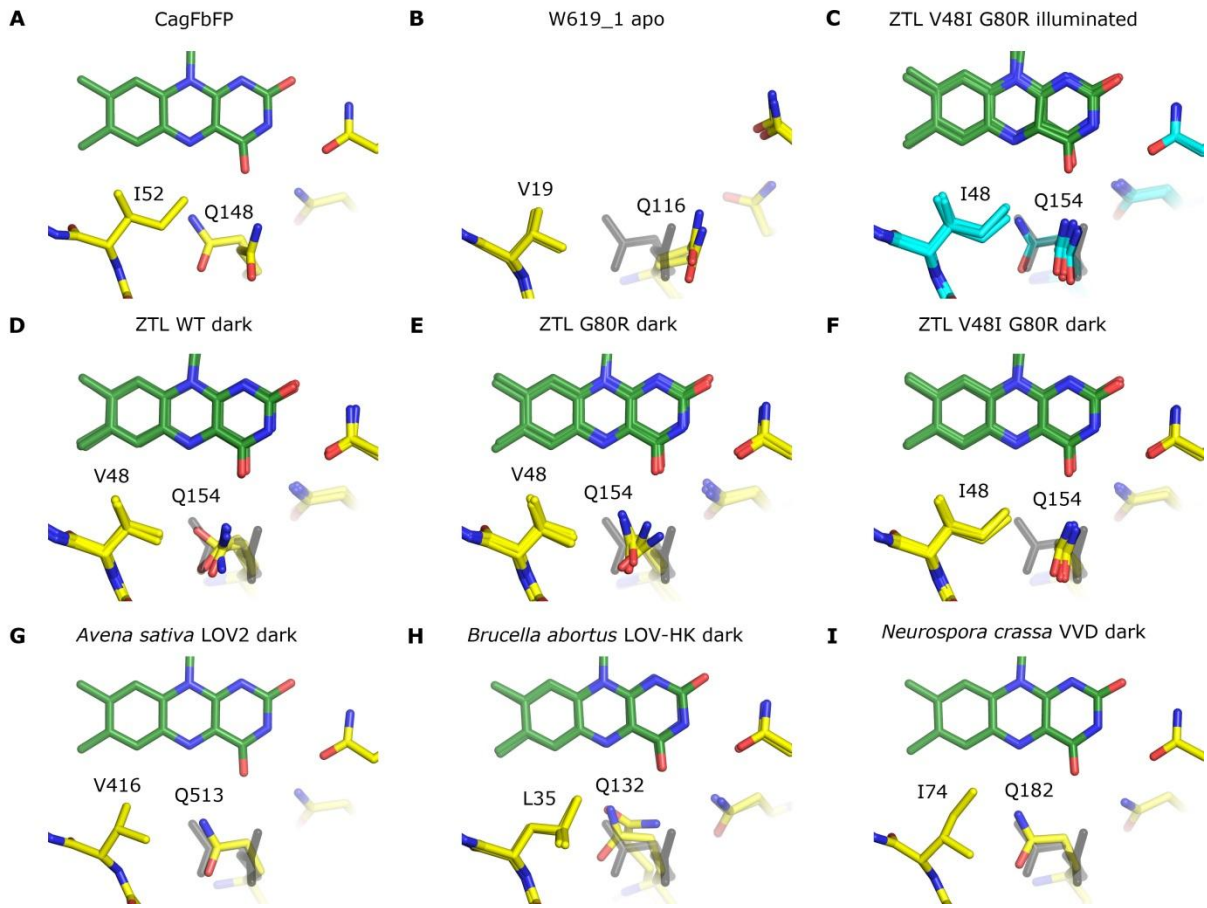

Figure S4. "Buried" and "exposed" conformations of the signaling glutamine amino acid in different crystal structures. CagFbFP Q148 positions are shown in gray for reference. (A) Structure of CagFbFP. (B) Structure of cofactor-less LOV protein W619_1 (PDB ID 5LUV). (C) Structure of V48I G80R mutant of ZTL under illumination (PDB ID 5SVW). The "buried" conformation is partially occupied in the chain C and is not observed in the chains A, B and D. (D) Structure of WT ZTL in the darkness (PDB ID 5SVG). (E) Structure of G80R mutant of ZTL in the darkness (PDB ID 5SVU). (F) Structure of V48I G80R mutant of ZTL in the darkness (PDB ID 5SVV). (G) Structure of the *Avena sativa* LOV2 in the darkness (PDB ID 2V0U). (H) Structure of the *Brucella abortus* LOV-HK in the darkness (PDB ID 3T50). (I) Structure of the *Neurospora crassa* VVD in the darkness (PDB ID 2PD7).

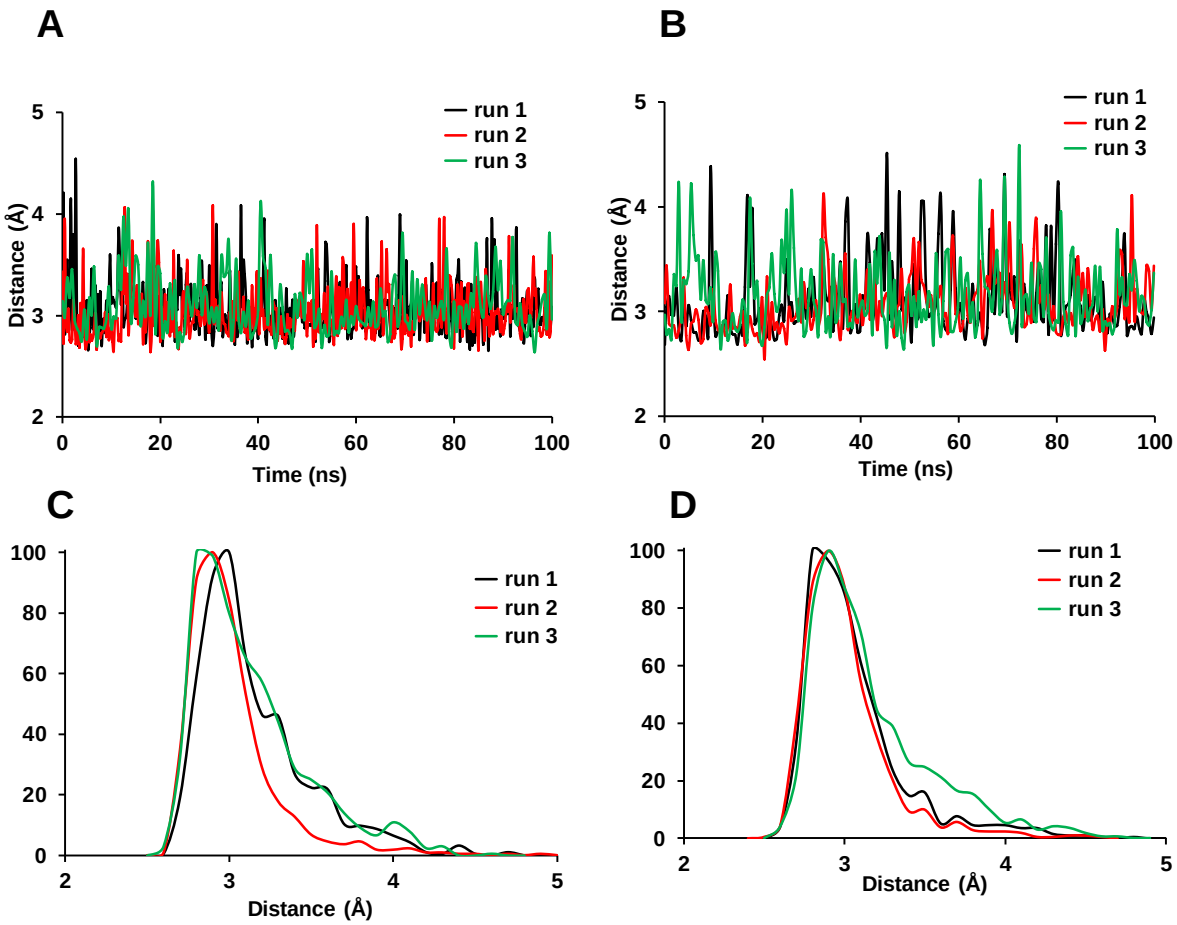

Figure S5. Distances between the Gln148 side chain and the O4 atom of FMN in molecular dynamics simulations of CagFbFP. (A) FMN-O4 and Q148-NE2 distance as a function of time starting from the "exposed" conformer. (B) FMN-O4 and Q148-NE2 distance as a function of time starting from the "buried" conformer. (C) Distance distribution curve of the interatomic distances between FMN-O4 and Q148-NE2 (in Å) calculated over the MD trajectories starting from the "exposed" conformer. (D) Distance distribution curve of the interatomic distances between FMN-O4 and Q148-NE2 (in Å) calculated over the MD trajectories starting from the "buried" conformer.