Photochemical & Photobiological Sciences

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pps

Journal Name

ARTICLE

Near infrared fluorescent biliproteins generated from bacteriophytochrome AphB of *Nostoc* sp. PCC 7120

Received 00th January 20xx,
Accepted 00th January 20xx

Che Yuan[a], Hui-Zhen Li[a], Kun Tang[a,b], Wolfgang Gärtner[b], Hugo Scheer[c], Ming Zhou*[a] and Kai-Hong Zhao*[a]

DOI: 10.1039/x0xx00000x

www.rsc.org/

The genome of the cyanobacterium *Nostoc* sp. PCC 7120 encodes a large number of putative bacteriophytochrome and cyanobacteriochrome photoreceptors that, due to their long-wavelength absorption and fluorescence emission, might serve as fluorescent tags in intracellular investigations. We show that the PAS-GAF domain of the bacteriophytochrome, AphB, binds biliverdin covalently and exhibits, besides its reversible photochemistry, a moderate fluorescence in the near infrared spectral region (NIR). It was selected for further increasing the brightness while retaining the NIR fluorescence. In a first step, amino acids assumed to improve fluorescence were selectively mutated. The resulting variants were then subjected to several rounds of random mutagenesis and screened for enhanced fluorescence in the NIR. The brightness of optimized PAS-GAF variants increased more than threefold compared to that of wt AphB(1-321), with only insignificant spectral shifts (A_{max} around 695 nm, and F_{max} around 720 nm). In general, the brightness increases with decreasing wavelengths, which allows for a selection of the fluorophore depending on the optical properties of the tissue. A spectral heterogeneity was observed when residue His260, located in close proximity to the chromophore, was mutated to Tyr, emphasizing the strong effects of the environment on the electronic properties of the bound biliverdin chromophore.

Introduction

Bacteriophytochromes (Bphs) are phylogenetically related to the canonical, red/far-red responsive phytochromes from plants and cyanobacteria. As in these proteins, the photosensory domain is composed of a PAS, a GAF, and a PHY domain. Bphs bind their biliverdin (BV) chromophore, however, in the PAS (SMART acc. no. SM00091) domain, while the canonical phytochromes bind their phycocyanobilin (PCB) chromophore into the GAF (SMART acc. no. SM00065) domain. Chromophore binding is autocatalytic, and retained a truncated PAS-GAF construct of Bphs [1]. However, many of these C-terminally truncated Bphs are impaired in their photochemistry inasmuch the Pfr state is not formed or shows only moderate thermal stability [2]. As in canonical phytochromes, the photoconversion of Bphs consists of a 15Z/E isomerization of BV [3]; its thermostable ('dark') state carries the 15Z-isomer of the BV chromophore (λ_{Amax} = 700 nm), and the metastable photoproduct state contains the 15E-isomer of BV (λ_{Amax} = 750 nm) [4, 5].

BV is the farthest red absorbing natural bilin chromophore and available in aerobic cells from heme-degradation. The free chromophore is non-fluorescent, and fluorescence is generally only marginally enhanced when bound to Bphs in both the thermostable and metastable states: Bphs are optimized for photochemistry. Yet, several studies on PAS-GAF constructs demonstrated that the fluorescence of the (parental) thermostable state can be significantly increased after several rounds of (random) mutagenesis [6-9]. These modified PAS-GAFs express well in various cells and fluoresce in the near infrared (NIR) region after BV incorporation. The increase in fluorescence quantum yield obtained so far, the bathochromic absorbance and fluorescence and the general intracellular availability of biliverdin from oxygenic degradation of heme make these BV-PAS-GAFs promising candidates as NIR fluorescent biomarkers [6, 10-13].

Absorbance and fluorescence of Bphs depend on the bound BV chromophores and its non-covalent interactions with the apoprotein [9, 14, 15]. Whereas the BV chromophore is the best choice for NIR fluorescence, one might intend to modify the apo-protein to further improve the properties of NIR fluorescent biliproteins [9]. *Nostoc* sp. PCC 7120 contains a canonical phytochrome, AphA [16], a bacteriophytochrome AphB [17], and a remarkably large amount of cyanobacteriochromes (CBCRs) [18] that act as sensory photoreceptors [19, 20]. We here present modifications of the PAS-GAF domain of AphB by, firstly, site-directed mutagenesis of amino acids (aas) that were reported in recent studies on fluorescent Bphs to enhance the NIR fluorescence of BV-PAS-GAF [6, 9-12]. These site-specific mutated proteins were then subjected to random

[a] State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, P.R. China.
[b] Max-Planck-Institute for Chemical Energy Conversion, Stiftstrasse 34-36, D-45470 Mülheim, Germany.
[c] Department Biologie I, Universität München, D-80638 München, Germany.

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

mutagenesis. After several iterative cycles, automated screening of AphB-expressing *E. coli* cells detected increased fluorescence for several of these randomly mutated chromoproteins. These fluorescence-improved proteins were then analyzed individually in greater detail by steady state absorption and fluorescence spectroscopy as well as by time-resolved fluorescence spectroscopy.

Materials and methods

Cloning and expression

All genetic manipulations were carried out following standard protocols [21]. The DNA fragment of the PAS-GAF domain of AphB was PCR-amplified from the genomic DNA of *Nostoc* sp. PCC 7120 using the primers in Table S1 and *Taq* DNA polymerase (MBI Fermentas). Amino acids that are considered relevant to increase the fluorescence yield were mutated (for primers see Table S1) employing the mutation kit (TaKaRa) according to the manufacturer's instructions. PCR products were ligated into the cloning vector pBluescript (Stratagene), verified by sequence analysis, and subcloned into pET30 (Novagen).

Random mutagenesis was performed using error-prone PCR [22] or the PCR random mutagenesis kit (TaKaRa) using the mutation primers containing the cutting sites of *BamH* I and *Xho* I (Table S1). Libraries from random mutagenesis were then cloned into pET30 and transformed into *E. coli* BL21 (DE3) that were pre-transformed with the BV-generating plasmid pACYC-ho1 [23]. After induction with 0.05 mM Isopropyl β-D-thiogalactopyranoside (IPTG) at 37°C for 12 hours and subsequently at 17°C for 24 hours, the most strongly fluorescing colonies on the agar plates were collected via the channel of excitation around 603-668 nm and fluorescence around 667-732 nm using a model Qpix 420 colony picking system (Molecular Devices). These collected highly bright clones in each library served as templates for the next round of error-prone PCR.

For over-expression of BV-PAS-GAF variants (Table S2 and S3), these pET30-derived expression vectors (Table S4) were transformed into *E. coli* BL21(DE3) (Novagen) containing the BV-generating plasmid, pACYC-ho1 [23]. The doubly transformed cells were cultured at 18 °C in Luria-Bertani (LB) medium supplemented with kanamycin (20 μg/ml) and chloromycetin (17 μg/ml). After induction with isopropyl β-D-thiogalactoside (1 mM) for 16 hours, the cells were centrifuged at 12,000 × g for 3 min at 4 °C, washed twice with water and stored at -20 °C until use. For isolating the BV-PAS-GAFs, the cell pellet was resuspended in ice-cold potassium phosphate buffer (KPB, 20 mM, pH 7.0) containing 0.5 M NaCl and disrupted by sonication for 5 min at 50 W (JY92-II, Scientz Biotechnology, China). The suspension was centrifuged at 12,000 × *g* for 15 min at 4 °C and the supernatant purified *via* Ni^{2+}-affinity chromatography on chelating Sepharose (Amersham Biosciences), developed with KBP. Bound proteins were eluted with KBP containing, in addition, imidazole (0.5 M). After collection, the sample was dialyzed against KBP.

Protein analyses and bioinformatics

Protein concentrations were determined by the Bradford assay [24], using bovine serum albumin for calibration, and SDS-PAGE was performed with the buffer system of Laemmli [25]. Proteins gels were stained with Coomassie brilliant blue. Proteins carrying a covalently bound chromophore were identified by Zn^{2+}-induced fluorescence [26]. Homology modeling was performed with the SWISS-MODEL server [27, 28] using the reported crystal structure (pdb: 4E04) of PAS-GAF of Bph from *Rhodopseudomonas palustris* [29] as template. The structural comparison was performed with the Swiss-PDBViewer version 4.1 [30].

Spectrophotometric analyses

Irradiations were carried out with a fibre optical cold-light source (Volpi, 150 W, Germany) equipped with appropriate filters (15 nm full width at half maximum (fwhm), light intensity 30 $\mu mol \cdot m^{-2} \cdot s^{-1}$). Unless otherwise stated, samples were irradiated for 5 min at each irradiation wavelength using a cycle of 420-650-420 nm. In case of poorly photoconverting samples, saturation was verified by irradiation for up to 30 min. Spectra were recorded before irradiation, and immediately after each saturating irradiation.

All chromoproteins were investigated by UV-Vis absorption spectroscopy (Beckman-Coulter DU 800). BV chromophore content in BV-PAS-GAFs was quantified after denaturation with acidic urea (8 M, pH 2.0) by its absorption at 388 nm (ε = 39'900 $M^{-1}cm^{-1}$ [10]). Fluorescence spectra were recorded with a F320 spectrofluorimeter (TianJin GangDong Sci and Tech Development Company, China) or with a Fluorolog 3 spectrofluorimeter (HORIBA). Fluorescence lifetimes were determined with this Fluorolog 3 fluorophotometer equipped with a single photon-counting unit (HORIBA). Fluorescence quantum yields, Φ_{fl}, were determined in KPB (50 mM, pH 7.2), using the known value Φ_{fl} = 0.06 of RGS [31] from *Synechocystis* sp. PCC6803 as standard.

Fluorescence microscopy

NIR fluorescence from *E. coli* expressing the BV-PAS-GAF variants was detected from the respective strains after induction and growth for several hours, until the fluorescence was already visible with the naked eye. After harvesting, the cells were deposited in low melting point agarose on a glass slide, and micrographs were taken with a fluorescence microscope (Immersol 518F, Carl Zeiss) fitted with a CCD camera (CoolSNAP HQ2, Photometrics). The micrographs were recorded through a NIR band-pass filter (685 – 735 nm), with excitation through a red band-pass filter (640 – 660 nm).

Results and Discussion

AphB is a bacteriophytochrome in *Nostoc* sp. PCC 7120

Journal Name

ARTICLE

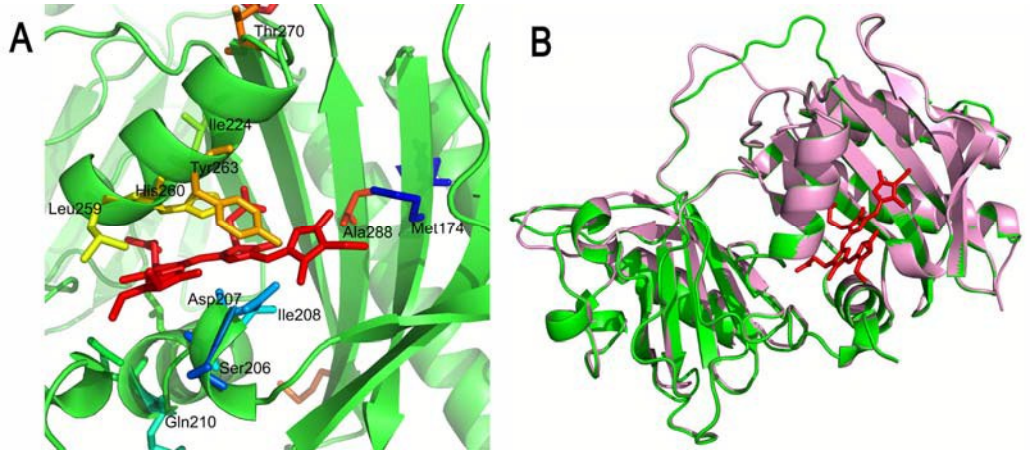


Fig. 1 Structural model of the PAS-GAF domain of AphB from *Nostoc* sp. PCC 7120. This structure (aa15-321) was modeled based on the crystal structure (pdb: 4E04) of PAS-GAF of Bph from *Rhodopseudomonas palustris*. **A**, Detailed view of the chromophore surrounding of BV-PAS-GAF of AphB, BV (sticks) in red, and apo-protein (cartoon) in green. **B**, overview on an overlay of the modelled PAS-GAF structure of AphB (green) and that of *P. palustris* (4E04, pink).

Table 1 Mutations and nomenclature of selected variants of AphB(1-321) (i.e. PAS-GAF). (a) site-directed mutations. (b) the unique variant and its ancestor. The samples shown here have been selected to show the variants with highest brightness, in comparison to the wt and the site-directed parental mutants (Table 2). The complete set of variants studied is given in Table S2.

Variants	Sequence variations
PAS-GAF-1	AphB(1-321)
PAS-GAF-2	AphB(7-321/E74V/Q89R/N92D/L139R/D207T/I208V/Y263F/T270A/A288V/I301T)
PAS-GAF-3	AphB(7-321/L37F/S100T/L139H/D207T/I208V/Y263F/N266S/I287V/A288V)
PAS-GAF-4	AphB(7-321/E74V/I90S/N116D/L139H/D207T/I208V/Y263F/M278V)
PAS-GAF-5	AphB(7-321/T51P/E74V/L139H/D207T/I208V/I234V/Y263F/A288V)
PAS-GAF-6	AphB(7-321/E74V/L139H/Q181L/D207T/I208V/Y263F/A288V)
PAS-GAF-7	AphB(7-321/S132P/K166R/D207T/I208V/Y263F/A288V)
PAS-GAF-8	AphB(7-321/V36A/E74A/L126S/K129E/V140A/V190I/K191R/D207T/I208V/Y263F/A288V/T312A)
PAS-GAF-9	AphB(7-321/E74V/K80E/L139H/M144V/L145P/D207T/I208V/Y263F/A288V)
PAS-GAF-10	AphB(7-321/E74V/L139H/K141R/I159V/S206N/D207T/I208V/Y263F/A288V)
PAS-GAF-11	AphB(7-321/L37F/V47G/E74V/L142P/K146R/V173A/D207T/I208V/Q211R/L215P/Y263H/A288V)
PAS-GAF-12	AphB(7-321/E74V/K80E/F133S/L139H/K141R/D207T/I208V/Y263F/A288V)
PAS-GAF-13	AphB(7-321/I28V/E74V/L139H/L142P/D207T/I208V/Y263F/A288V)
PAS-GAF-14	AphB(7-321/E74V/D207T/I208V/Y263F/A288V)
PAS-GAF-29 [a]	AphB(7-321/D207T/I208V/A288V)
PAS-GAF-36 [a]	AphB(7-321/D207T/A288V)
PAS-GAF-37 [b]	AphB(7-321/E74V/D207H/I208T/Y263F/A288V)
PAS-GAF-52 [a]	AphB(7-321/D207A/A288V)
PAS-GAF-53 [a]	AphB(7-321/D207H/A288V)
PAS-GAF-55 [a]	AphB(7-321/D207F/A288V)
PAS-GAF-56 [a]	AphB(7-321/D207A/S293T)
PAS-GAF-60 [a]	AphB(7-321/D207G/A288V)
PAS-GAF-64 [b]	AphB(7-321/E74V/D207H/I208T/H260Y/Y263F/A288V)

Table 2 Quantitative absorption and fluorescence data of selected NIR fluorescent biliproteins derived from the PAS-GAF fragment of AphB. Spectra were recorded in KPB (20 mM, pH 7.0) containing 0.5 M NaCl after purification by Ni^{2+} affinity chromatography. Unless otherwise stated, the fluorescence spectra were obtained upon excitation at 600 nm. Experimental traces were fitted with a bi-exponential sequential model. (a) site-directed mutations. (b) unique H260Y variant and its ancestor. (c) fluorescence spectrum obtained upon 640 nm excitation. (d) the brightness of the BV-variants relative to that of the wt BV-PAS-GAF-1 ($\varepsilon \bullet \Phi_{fl}$ = 1.5 $mM^{-1}cm^{-1}$). The samples shown here have been selected to show the variants with highest brightness, in comparison to the wt and the site-directed parental mutants. For detailed sequence information for the various constructs and for a complete list of absorbance and fluorescence parameters of all constructs see Tables S2, S3.

BV-Variants	Absorption		Fluorescence			
	λ_{max} [nm]	$\varepsilon \bullet 10^{-4}$ [M^{-1} cm^{-1}]	λ_{max} [nm]	Φ_{fl}	Brightness [d] [%]	Life time [ns] τ_1 (%), τ_2 (%); χ^2
PAS-GAF-1	697	6.4	721	0.023	100	0.48(97.7), 3.0(2.3); 1.86
PAS-GAF-2	693	8.1	717	0.060	330	0.62(98.1), 4.0(1.9); 1.94
PAS-GAF-3	691	7.7	717	0.060	314	0.61(97.6), 3.4(2.4); 1.81
PAS-GAF-4	693	8.1	717	0.059	325	0.62(98.2), 3.9(1.8); 2.11
PAS-GAF-5	693	7.9	717	0.057	306	0.61(98.0), 3.7(2.0); 1.84
PAS-GAF-6	693	8.0	718	0.057	310	0.61(97.6), 3.3(2.4); 2.15
PAS-GAF-7	693	8.0	717	0.057	310	0.61(98.1), 4.1(1.9); 1.82
PAS-GAF-8	693	7.9	718	0.057	306	0.60(98.1), 4.1(1.9); 1.86
PAS-GAF-9	693	8.4	717	0.057	326	0.61(97.8), 3.4(2.2); 2.00
PAS-GAF-10	694	8.1	718	0.057	314	0.60(97.8), 3.6(2.2); 1.84
PAS-GAF-11	693	7.9	717	0.056	301	0.59(97.9), 3.6(2.1); 1.85
PAS-GAF-12	693	7.9	717	0.056	301	0.62(98.1), 3.8(1.9); 2.00
PAS-GAF-13	693	8.1	716	0.056	308	0.61(97.8), 3.5(2.2); 2.16
PAS-GAF-14	694	8.2	718	0.056	312	0.60(98.6), 5.4(1.4); 1.83
PAS-GAF-29 [a]	694	7.0	718	0.050	238	0.57(98.0), 4.1(2.0); 1.65
PAS-GAF-36 [a]	695	7.0	719	0.049	233	0.59(98.0), 4.1(2.0); 1.51
PAS-GAF-37 [b]	693	7.0	719	0.048	228	0.59(97.6), 3.6(2.4); 1.70
PAS-GAF-52 [a]	698	7.6	722	0.045	232	0.55(98.0), 4.0(2.0); 1.56
PAS-GAF-53 [a]	698	7.2	722	0.044	215	0.56(98.0), 5.8(2.0); 1.62
PAS-GAF-55 [a]	699	6.2	723	0.043	181	0.58(97.5), 6.8(2.5); 1.69
PAS-GAF-56 [a]	700	7.2	723	0.042	206	0.56(98.4), 4.4(1.6); 1.48
PAS-GAF-60 [a]	701	6.4	723	0.040	174	0.54(98.2), 4.8(1.8); 1.69
PAS-GAF-64 [b]	606/ 692	2.2	648/ 712	0.031	46	0.27(96.7), 8.4(3.3); 2.28
PAS-GAF-64 [b,c]	606/ 692	2.2	718	0.017	25	0.27(96.8), 9.1(3.2); 2.39

As *Calothrix* sp. PCC7601, also in *Nostoc* sp. PCC7120 has two phytochrome-encoding genes *aphA* and *aphB* [5]. While AphA is a canonic phytochrome binding covalently phycocyanobilin [16], AphB is annotated as a Bph that binds covalently BV at Cys-17 in the PAS domain [17]. The crystal structure of PAS-GAF of AphB has not yet been reported, but those of orthologous PAS-GAF constructs and of mutated NIR fluorescent variants of Bphs from *Deinococcus radiodurans* (pdb: 3S7Q [9], 4O8G [15]) and from *Rhodopseudomonas palustris* (pdb: 4E04) [29] have been resolved at high resolution. A sequence alignment between the PAS-GAF domains of *D. radiodurans*, *R. palustris*, and AphB from *Nostoc* sp. PCC7120 yielded the highest degree of similarity (70%) of AphB to the *R. palustris* sequence. Accordingly, a modeling approach (Fig. 1) yielded better coincidence of the backbone and side chains with the *R. palustris* than with the *D. radiodurans* structure. The simulated AphB structure showed moderate deviations from the *R. palustris* structure only for a random coil region (aa124-136).

The PAS-GAF-coding segment of *aphB* (i.e. *aphB(1-321)*) was expressed together with *ho1* encoding heme oxygenase 1 (HO1) in *E. coli* in a two plasmid approach. This set-up generated BV from heme that was autocatalytically attached to AphB(1-321) (Fig. 2), as demonstrated on SDS-PAGE by Zn^{2+}-induced fluorescence [26]. BV-AphB(1-321) absorbed maximally at 697 nm, and fluoresced weakly at 721 nm (Φ_{fl} = 0.023, Table 1 and 2). This construct had moderate photochemistry (Fig. S1): it showed a photochemistry characteristic for Bphs with a Pr peak at λ_{max} = 697 nm that upon irradiation at 650 nm decreased, accompanied by formation of a shoulder around 750 nm indicative for Bphs Pfr absorption. As observed also for other Bphs, the conversion into the Pfr form can be accomplished only to a small extent, even upon extended

irradiation. No functional explanation to this result can be given.

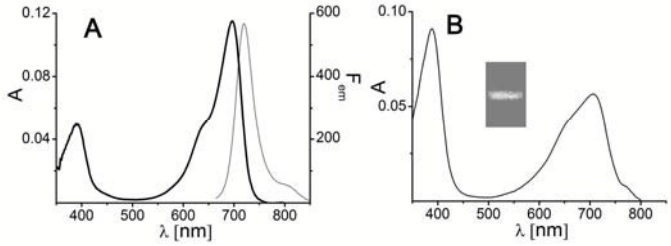


Fig. 2 Characterization of the starting BV-PAS-GAF-1 (BV-AphB(1-321)) A: Absorption (bold line) and fluorescence spectra (λ_{exc} = 640 nm, thin line) (20 mM KPB, pH 7.0 containing 0.5 M NaCl). **B**: Absorption spectrum after denaturation of the sample (8 M urea in acidic KPB buffer, pH 2.0). Covalent attachment of BV to PAS-GAF was demonstrated by Zn^{2+} induced fluorescence on SDS-PAGE (insert). The sample was generated in *E. coli*, and spectra were obtained after purification by Ni^{2+} affinity chromatography.

Design of site-directed mutations

Driven by the long wavelength absorbance and the similarly red-shifted, yet only weak fluorescence of BV-AphB(1-321), we aimed at increasing the fluorescence yield by first performing mutations of selected amino acids, and then subjecting these site-directed variants of AphB to several rounds of random mutagenesis. The mutants were screened automatically for high fluorescence in the NIR. Several variants showed in *E. coli* cultures a bright fluorescence (λ_{exc} = 630 nm) that is well visible by the naked eye, which is quite insensitive in this spectral region.

Investigation on PAS-GAF of *Dr*BphP from *Deinococcus radiodurans* [9, 10] and *Rp*BphP2 from *Rhodopseudomonas palustris* [12] aiming at increasing the fluorescence quantum yield had identified an instrumental aspartate residue (position 207 in AphB) that upon mutation leads to distinct changes of the absorption and fluorescence properties of Bphs. This aspartate residue stabilizes the pyrrole rings A, B, and C of the chromophore either directly via its side chain carboxylate or through an intercalating water molecule: the type of interaction depends on the protein under investigation. Accordingly, D207 was mutated into either T, A, H, F, or G, together with an N-terminal truncation of residues aa1-6. The full-length construct PAS-GAF-1 (AphB(1-321)), the N_6-truncated protein, and also the D207H variant 58 (Tables 1, 2) exhibited very similar absorption and fluorescence properties, thus AphB(7-321) served as the origin for the site directed mutants (see Tables 1, 2 and, Tables S2, S3 for full information. Following the parental protein, PAS-GAF-1 (= WT AphB(1-321), the various multiple mutants are numbered and listed in order of decaying fluorescence quantum yield). Starting with D207T/A/H/F/G single mutants, double and triple mutations were introduced (listed number in brackets): D207T/I208V/A288V (#29), D207T/A288V (#36), D207A/A288V (#52), D207H/A288V (#53), D207F/A288V (#55), D207A/S293T (#56), D207G/A288V (#60), D207A/V256F/A288V (#41), D207H/K295Y (#57), D207H (#58), D207H/Q292C (#61), D207A/A288V/E306A (#62). These site-directed mutants were then subjected to several rounds of random mutagenesis.

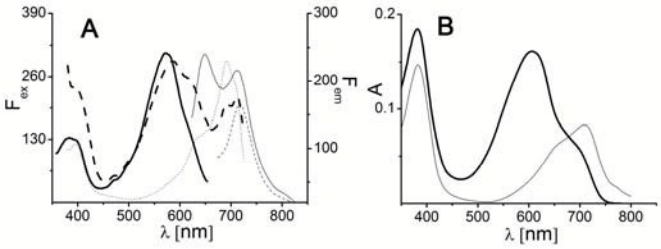


Fig. 3 Spectral heterogeneity of the unique PAS-GAF-64 (E74V/ D207H/I208T/H260Y/Y263F/A288V) variant. A: Fluorescence excitation (bold lines) and emission spectra (thin lines). Emission spectra were measured with excitation at 600 nm (solid) and 640 nm (dashed), excitation spectra with detection at 670 nm (solid) and 740 nm (dashed). For comparison, we show the excitation spectrum of variant BV-PAS-GAF-37 (E74V/D207H/I208T/Y263F/A288V) (λ_{em} = 740 nm, dotted line). **B**: Absorption spectra of the native sample (20 mM KPB at pH 7.0 containing 0.5 M NaCl, bold line) and after denaturation with 8 M acidic urea (pH 2.0, thin line). The samples were generated in *E. coli*, and then purified by Ni^{2+} affinity chromatography.

Spectral properties of BV-PAS-GAF variants of AphB

The site-directed and random mutations of the BV-bound PAS-GAF variants of Bph AphB yielded 65 BV-PAS-GAF variants with altered or improved fluorescence properties that were spectrally analyzed in greater detail (Table 1 and 2, Table S2 and S3). The properties of the site-directed variants concurred with reported data obtained with related Bphs [7, 15, 32]. As a general trend, mutations leading to an increased fluorescence lifetime also showed a larger fluorescence quantum yield (Table 2, S3). Compared with the wild type protein (PAS-GAF-1, aa 1-321) showing a low quantum yield of 2.3 %, and a short fluorescence lifetime of $\tau_1 = 0.48$ ns (fastest kinetic component), the major lifetime component for all other proteins (τ_1) was longer than 0.52 ns, except for PAS-GAF-64 (Tables 1 and 2) and PAS-GAF-65 (Tables S2 and S3). These two proteins showed significantly shorter fluorescence lifetime (τ_1 0.27 and 0.36 ns, respectively), and correspondingly had a smaller fluorescence yield quantum (Φ_{fl} = 0.017 and 0.031 for PAS-GAF-64 and PAS-GAF-65 (Φ_{fl} = 0.025). Variant 64 showed two fluorescence maxima with different quantum yields (Fig. 3A, discussed below).

Among these NIR fluorescent biliproteins, maximal absorbance (A_{max}) was around 695 nm and maximal fluorescence (F_{max}) around 720 nm, yielding a Stokes shifts of around 25 nm (Table S3). Usually, PAS-GAF mutants with λ_{Amax} < 695 nm and λ_{Fmax} < 720 nm showed a higher fluorescence quantum yield, and reversely PAS-GAF mutants with a weaker fluorescent had λ_{Amax} > 695 nm and λ_{Fmax} > 720 nm (Table 2 and S3). There are three variants deviating from this trend, viz.

PAS-GAF-37, -64, and -63: PAS-GAF-37 and -63 show identical mutation patterns (E74V/I208T/Y263F/A288V) except for the site-directed mutation at position 207 (D207H and D207W, respectively, Tables 1, S2). PAS-GAF-64 differs from variant 37 only by the additional mutation of H260Y. This mutation causes heterogeneity in the protein, which will be discussed later (see below).

Based on the simulated structure of AphB (Fig. 1), amino acids Asp207, Ile208, Tyr263 and Ala288 are all in close vicinity to the chromophore and have all been identified as important to modify the NIR fluorescence [6, 9-12]. All these amino acids are situated near to the BV chromophore (3.0 - 4.2 Å), while Glu74 that was mutated in half of highly fluorescent variants (Table 2 and S3) is found at a larger distance (~38 Å). The mutations of D207H and Y263F have a strong positive effect on the NIR fluorescence [9], and also the mutations D207T, I208V and A288V showed increased NIR fluorescence (Tables 1, 2). The mutation of I208T seems to be neutral, and the insertion of a tryptophan (D207W) results in very low fluorescent proteins like BV-PAS-GAF-63 (Table S2 and S3). Inspection of the simulated structure (Fig. 1) offers an explanation for this result: Trp at position 207 (D207W variant) probably interacts electronically with the pyrrole rings of the chromophore, and also would form multiple strong H-bonds with Pro 204 and Tyr 263, resulting in a substantial change of the chromophore pocket with a strong impact on the fluorescence properties of the BV chromophore.

The absorption and fluorescence spectra of the BV-PAS-GAF variants are of similar shape as the WT (Fig. 2) for practically all constructs. With a single exception: a critical residue for the maintenance of the spectral properties is apparently His at position 260. The variants #37 and #64 differ only in the presence of an additional mutation (H260Y) in variant 64, but their absorption and fluorescence spectra are changed remarkably (Fig. 3). BV-PAS-GAF-64 is unique in this study. It absorbed maximally at 606 nm with a shoulder absorbance around 692 nm, and also the fluorescence emission shows two peaks (λ_{Fmax} = 648 and 712 nm) instead of a single one (λ_{Fmax} = 718 nm +/- 4 nm) found for all other constructs. The emission at 712 nm is apparently comparable to the normal one at 718 nm, whereas the peak at 648 nm appears only for this mutation. Although a lower fluorescence quantum yield results, we maintain this variant here to highlight the eminent role of this conserved histidine residue for the functional properties of phytochromes. This variant also demonstrates the sensitivity of chromophore-protein interactions in phytochromes in general: a comparison of the excitation and emission spectra indicate an underlying heterogeneity of the chromophore conformations as a cause for the obtained results (Fig. 3): the 648 nm emission results from the component absorbing at ~ 580 nm, and the 718 nm emission is elicited from excitation at ~ 580, 615, 690 and 710 nm components. The shorter absorbance (λ_{Amax} = 606 nm) and fluorescence wavelengths (λ_{Fmax} = 648 nm) are similar to those of phycocyanobilin in phycobiliproteins, requesting identification of the chromophore structure. After denaturation under acidic urea, the maximal absorbance was at 709 nm (Fig. 3), corresponding to that of the covalently bound BV (Fig. 2). The maximal absorbance of covalently bound phycocyanobilin, e.g., in canonical cyanobacterial phytochromes, Cph1, is at 660 nm [33]. This would mean that although the chromophore is still BV, the conjugation system is effectively shortened. Probably, one pyrrole ring in this unique variant is twisted (partially) out of conjugation.

A structural heterogeneity upon His260 mutation is also supported by the photochemistry: the 580 and 615 nm components photobleached and the 690 and 710 nm components were hardly photoactive (Fig. S1). In the simulated structure (Fig. 1), the imidazole ring of His260 is only about 3.2 Å away from ring C of the BV chromophore, so exerting strong effects on the conformation of the BV chromophore. When the His260 is changed to Tyr, its phenolic ring gets very close to ring B instead of ring C (as is the case for the H260 ancestor). As a result, it might twist the conformations around ring B due to the resulting crowding and thus generate large structural changes resulting in the particular spectral features of BV-PAS-GAF-64. As this mutation appears only once in the entire random mutagenesis experiment, one could imagine that this position is a hotspot also for the protein stability, generating upon exchange unstable proteins that do not sufficiently maturate as holoproteins or are incapable of chromophore binding.

Certain PAS-GAF variants of AphB show improved NIR fluorescence

Besides identifying the structural effects caused by the mutations, this study aimed at increasing the NIR fluorescence of the PAS-GAF part of ApcB. Several literature studies identified a number of point mutations yielding an increase of the NIR fluorescence [6, 9-12]. The most effective amino acids were Asp207, Ile208, Val256, Ala288, and Glu306 [9, 15, 34], with Asp207, reported to have the strongest effect on the NIR fluorescence. It was chosen as starting point for further site-directed and random mutagenesis (Table 1 and 2, Table S2 and S3). Mutation of Asp207 to Ala, Phe, Gly, His or Thr (Table 1 and S2) gave especially for the D207T mutant a remarkably enhanced fluorescence quantum yield (Φ_{fl} = 0.05, compared to values <0.05 for all other D207 mutations). Hydrophobic aa residues such as Ala, Gly and Phe at position 207 show red-shifted spectra by several nanometers. While e.g., glycine at position 207 significantly reduces both fluorescence quantum yield and extinction coefficient, D207T seemed the best candidate of this series of Asp207 point mutations (Table 2 and S3). Although the above mentioned variants showed increased fluorescence, some of them also retained moderate photochemistry including a substantial bleaching under red light (Fig. S1). To circumvent the reduced stability during irradiation and to further improve the fluorescence intensity, several rounds of random mutagenesis were performed. In most cases the D207T/A288V double variant was taken as parental protein, except for exchanges of aspartate207 into histidine, D207H (variants 37, 45, 53, 57, 58, 61, 64, Table S2), alanine, D207A (variants 41, 44, 48-52, 54, 56, 59, 62), phenylalanine, D207F (variant 55), glycine, D207G (variant 60),

or tryptophan, D207W (variant 63). In nearly all cases, the isolated variants showed better photochemical stability and higher fluorescence quantum yield (Φ_{fl}). The highest Φ_{fl} of 0.06 was determined for variants 2 and 3 (Table 2), corresponding to an increase in brightness by a factor >3. These two proteins originating from D207T/A288V show additional mutations at variable positions, however, they coincide for positions 208 and 263 (I208V, Y263F).

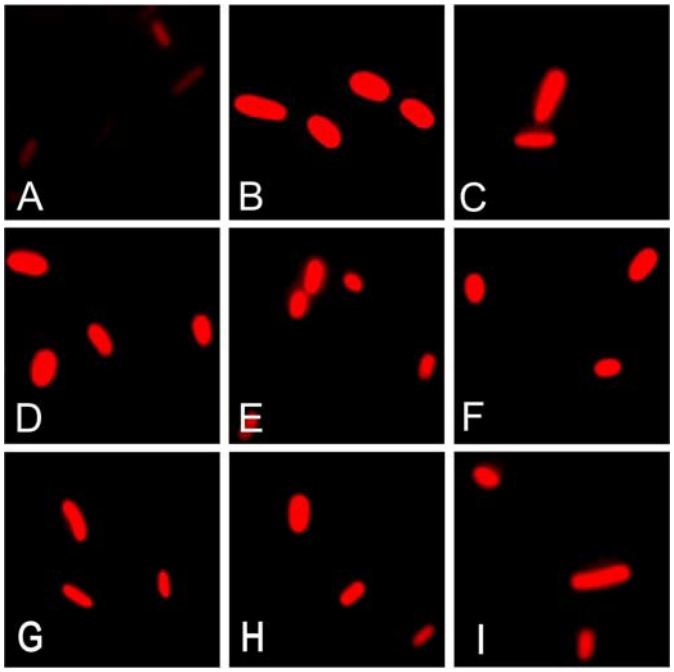


Fig. 4 Micrographs of *E. coli* cells expressing wt BV-PAS-GAF-1 (A), and the most brightly fluorescing variants 2 (B), 3 (C), 4 (D), 5 (E), 6 (F), 7 (G), 8 (H) and 9 (I). The micrographs were recorded through a NIR band-pass filter (685 – 735 nm), with excitation through a red band-pass filter (640 – 660 nm).

The effective aa mutations found to improve fluorescence (Table 1 and 2, Table S2 and S3) were not totally the same as the reported ones [6, 9-12], indicating the potential to improve BV-PAS-GAF as NIR fluorescent biliproteins by various exchanges. The data support the trend that a stronger red-shift of the absorption causes a lower extinction coefficient and also a reduced fluorescence quantum yield [35]: BV-PAS-GAF-65 shows the most red-shifted fluorescence emission wavelength (724 nm), but a fluorescence quantum yield of only 0.025. This finding might be explained by two effects: on the one hand, the energy gap between the energy levels of molecular orbitals of electron transition is changed, which impacts on the electronic properties in the excited state possibly involving charge-transfer contributions. The π-electron conjugation of the bilin chromophore can also be effectively increased by a lesser degree of twisting, which might again affect the electronic properties, but also the conformational flexibility of the chromophore in the protein's binding pocket: restricting the conformational flexibility of the chromophore increases its fluorescence capacity [15]. This trend is also followed by the various bilins found as chromophores in biliproteins, viz. phycourobilin, phycoerythrobilin, phycocyanobilin, phytochromobilin, and BV. With an increase of their π-electron conjugation and the concomitant red-shifted absorption, their fluorescence quantum yield becomes lower. It should also be noted that the even more red-shifted metastable states of phytochromes with a 15E-configured chromophore show practically no fluorescence. As a general finding, the reported data from the literature and results obtained in this work, indicate for BV in chromoproteins a fluorescence quantum yield <0.1 if the fluorescence maximum is >700 nm. Another relevant factor that is only poorly investigated in the current literature is the degree of chromophorylation [11, 36]. Judging from the relative absorptions of the chromophore and the protein, and their known (Table 2) and calculated extinction coefficients, all variants discussed in this work are chromophorylated to >50%.

Several variants were found particularly promising as NIR fluorescent biomarkers, such as PAS-GAF-2 through variant #9 (Table 1 and S2). Their brightness was more than threefold compared to the parental PAS-GAF-1 (Table 2). *E. coli* cells expressing these optimized PAS-GAF variants showed very bright NIR fluorescence that was already visible to the naked eye despite being quite insensitive in this spectral region (Fig. 4). Among these PAS-GAF variants, some have a slightly red-shifted fluorescence compared to the literature reports [6, 9-12], but none had a λ_{Fmax} beyond 725 nm. Although one BV-GAF is reported to fluoresce maximally at 730 nm, its fluorescence quantum yield is extremely low (only 0.003) [37], supporting the aforementioned trend. Seemingly, BV-chromoproteins might not exceed the limit of 730 nm at maximal fluorescence, and random mutagenesis approaches as the one applied here might indicate the upper limit for emission wavelength and quantum yield for phytochrome-derived NIR fluorescence markers. Despite the good expression and stabilization properties of the protein moieties and the ready availability of their biliverdin chromophore, one should, therefore, include alternative systems for NIR detection into the functional assay. Promising candidates are the phycobiliprotein antenna pigments of cyanobacteria due to their natural strong fluorescence and their comparably high extinction coefficients as the phytochromes [38], particularly phycobilisome core-membrane linkers, whose variants absorb and fluoresce >650 nm [39].

Abbreviations

AphB	bacterial phytochrome of *Nostoc* sp. PCC 7120
BV	biliverdin
Bph	bacteriophytochrome
GAF	protein domain (SMART acc. no. SM00065; acronym of common protein domain from cGMP phosphodiesterase, adenylyl cyclase and FhlA)
HO	heme oxygenase
KPB	potassium phosphate buffer
NIR	near infrared

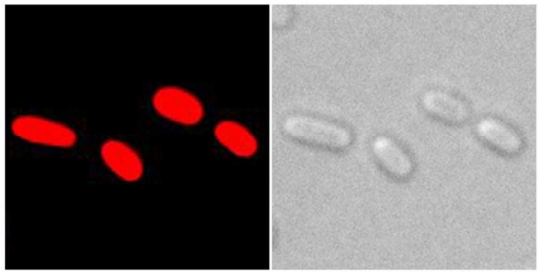
PAS protein domain (SMART acc. no. SM00091; acronym of common protein domain from period circadian protein, Ah receptor nuclear translocator protein and single-minded protein domain)

Acknowledgements

HS and KHZ (grant 31110103912), KHZ (grants 31270893 and 21472055) and MZ (grant 31370777) are grateful for support by the National Natural Science Foundation of China. WG is grateful to the Max-Planck-Society for generous continuing financial support.

References

1. N. C. Rockwell, Y. S. Su and J. C. Lagarias, *Annu. Rev. Plant Biol.*, 2006, **57**, 837-858.
2. S. H. Wu and J. C. Lagarias, *Biochemistry*, 2000, **39**, 13487-13495.
3. X. Yang, Z. Ren, J. Kuk and K. Moffat, *Nature*, 2011, **479**, 428-432.
4. R. Tasler, T. Moises and N. Frankenberg-Dinkel, *FEBS J.*, 2005, **272**, 1927-1936.
5. B. Quest, T. Hübschmann, S. Sharda, N. Tandeau de Marsac and W. Gartner, *FEBS J.*, 2007, **274**, 2088-2098.
6. K. D. Piatkevich, F. V. Subach and V. V. Verkhusha, *Nat. Commun.*, 2013, **4**, 2153.
7. K. C. Toh, E. A. Stojkovic, I. H. M. van Stokkum, K. Moffat and J. T. M. Kennis, *Proc. Natl. Acad. Sci. USA.*, 2010, **107**, 9170-9175.
8. A. J. Fischer and J. C. Lagarias, *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 17334-17339.
9. M. E. Auldridge, K. A. Satyshur, D. M. Anstrom and K. T. Forest, *J. Biol. Chem.*, 2012, **287**, 7000-7009.
10. X. Shu, A. Royant, M. Z. Lin, T. A. Aguilera, V. Lev-Ram, P. A. Steinbach and R. Y. Tsien, *Science*, 2009, **324**, 804-807.
11. D. M. Shcherbakova and V. V. Verkhusha, *Nat. Methods*, 2013, **10**, 751-754.
12. G. S. Filonov, K. D. Piatkevich, L. M. Ting, J. Zhang, K. Kim and V. V. Verkhusha, *Nat. Biotechnol.*, 2011, **29**, 757-761.
13. D. Yu, M. A. Baird, J. R. Allen, E. S. Howe, M. P. Klassen, A. Reade, K. Makhijani, Y. Song, S. Liu, Z. Murthy, S. Q. Zhang, O. D. Weiner, T. B. Kornberg, Y. N. Jan, M. W. Davidson and X. Shu, *Nat. Methods*, 2015, **12**, 763-765.
14. F. Velazquez Escobar, T. Hildebrandt, T. Utesch, F. J. Schmitt, I. Seuffert, N. Michael, C. Schulz, M. A. Mroginski, T. Friedrich and P. Hildebrandt, *Biochemistry*, 2014, **53**, 20-29.
15. S. Bhattacharya, M. E. Auldridge, H. Lehtivuori, J. A. Ihalainen and K. T. Forest, *J. Biol. Chem.*, 2014, **289**, 32144-32152.
16. K. H. Zhao, Y. Ran, M. Li, Y. N. Sun, M. Zhou, M. Storf, M. Kupka, S. Böhm, C. Bubenzer and H. Scheer, *Biochemistry*, 2004, **43**, 11576-11588.
17. T. Lamparter, *FEBS Lett.*, 2004, **573**, 1-5.
18. Q. Ma, H. H. Hua, Y. Chen, B. B. Liu, A. L. Kramer, H. Scheer, K. H. Zhao and M. Zhou, *FEBS J.*, 2012, **279**, 4095-4108.
19. M. Ikeuchi and T. Ishizuka, *Photochem. Photobiol. Sci.*, 2008, **7**, 1159-1167.
20. N. C. Rockwell, S. S. Martin, K. Feoktistova and J. C. Lagarias, *Proc. Natl. Acad. Sci. U S A*, 2011, **108**, 11854-11859.
21. J. Sambrook, E. Fritsch and T. Maniatis, *Molecular cloning: a laboratory manual*, Cold Spring Harbour Laboratory Press, New York, 2nd edn., 1989.
22. N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer and R. Y. Tsien, *Nat. Biotechnol.*, 2004, **22**, 1567-1572.
23. K. H. Zhao, P. Su, J. M. Tu, X. Wang, H. Liu, M. Plöscher, L. Eichacker, B. Yang, M. Zhou and H. Scheer, *Proc. Natl. Acad. Sci. U S A*, 2007, **104**, 14300-14305.
24. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248-254.
25. U. Laemmli, *Nature* 1970, **227**, 680-685.
26. T. Berkelman and J. C. Lagarias, *Anal. Biochem.*, 1986, **156**, 194-201.
27. K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195-201.
28. N. Guex, M. C. Peitsch and T. Schwede, *Electrophoresis*, 2009, **30 Suppl 1**, S162-173.
29. D. Bellini and M. Z. Papiz, *Acta Crystallogr. D Biol. Crystallogr.*, 2012, **68**, 1058-1066.
30. N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714-2723.
31. J. Zhang, X. J. Wu, Z. B. Wang, Y. Chen, X. Wang, M. Zhou, H. Scheer and K. H. Zhao, *Angew. Chem. Int. Edit.*, 2010, **49**, 5456-5458.
32. K. C. Toh, E. A. Stojkovic, I. H. van Stokkum, K. Moffat and J. T. Kennis, *Phys. Chem. Chem. Phys.*, 2011, **13**, 11985-11997.
33. A. N. Glazer and S. Fang, *J. Biol. Chem.*, 1973, **248**, 659-662.
34. J. R. Wagner, J. Zhang, D. von Stetten, M. Gunther, D. H. Murgida, M. A. Mroginski, J. M. Walker, K. T. Forest, P. Hildebrandt and R. D. Vierstra, *J. Biol. Chem.*, 2008, **283**, 12212-12226.
35. O. V. Stepanenko, M. Baloban, G. S. Bublikov, D. M. Shcherbakova, O. V. Stepanenko, K. K. Turoverov, I. M. Kuznetsova and V. V. Verkhusha, *Scientific reports*, 2016, **6**, 18750.
36. K. A. Rumyantsev, D. M. Shcherbakova, N. I. Zakharova, A. V. Emelyanov, K. K. Turoverov and V. V. Verkhusha, *Scientific reports*, 2015, **5**, 18348.
37. R. Narikawa, T. Nakajima, Y. Aono, K. Fushimi, G. Enomoto, W. Ni Ni, S. Itoh, M. Sato and M. Ikeuchi, *Sci. Rep.*, 2015, **5**, 7950.
38. A. Glazer, *J. Appl. Phycol.*, 1994, **6**, 105-112.
39. K. Tang, W.-L. Ding, A. Höppner, C. Zhao, L. Zhang, Y. Hontani, J. T. M. Kennis, W. Gärtner, H. Scheer, M. Zhou and K.-H. Zhao, *Proc. Natl. Acad. Sci. U S A*, 2015, **112**, 15880-15885.



Near infrared fluorescent proteins are assembled in *E. coli*, which relies on the ability of PAS-GAFs chromophorylated with biliverdin. The construct of strong brightness is promising as biomarker.