Blue protein with red fluorescence

Swagatha Ghosh[a], Chi-Li Yu[b], Daniel J. Ferraro[c], Sai Sudha[d], Samir Kumar Pal[e], Wayne F. Schaefer[f], David T. Gibson[b,1], and S. Ramaswamy[c,d,2]

[a]National Center for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India 560065; [b]Department of Microbiology, University of Iowa, Iowa City, IA 52242; [c]Department of Biochemistry, University of Iowa, Iowa City, IA 52242; [d]Institute for Stem Cell Biology and Regenerative Medicine, Bangalore, India 560065; [e]Department of Chemical, Biological & Macromolecular Sciences, S.N. Bose National Centre for Basic Sciences, Salt Lake, Kolkata 700 098, India; and [f]Department of Biological Sciences, University of Wisconsin, Washington County, WI 53095

Edited by Ce Feng Liu, Weill Cornell Medical College, New York, NY, and accepted by Editorial Board Member Gregory A. Petsko August 21, 2016 (received for review December 30, 2015)

The walleye (*Sander vitreus*) is a golden yellow fish that inhabits the Northern American lakes. The recent sightings of the blue walleye and the correlation of its sighting to possible increased UV radiation have been proposed earlier. The underlying molecular basis of its adaptation to increased UV radiation is the presence of a protein (Sandercyanin)–ligand complex in the mucus of walleyes. Degradation of heme by UV radiation results in the formation of Biliverdin IXα (BLA), the chromophore bound to Sandercyanin. We show that Sandercyanin is a monomeric protein that forms stable homotetramers on addition of BLA to the protein. A structure of the Sandercyanin–BLA complex, purified from the fish mucus, reveals a glycosylated protein with a lipocalin fold. This protein–ligand complex absorbs light in the UV region (λ_{max} of 375 nm) and upon excitation at this wavelength emits in the red region (λ_{max} of 675 nm). Unlike all other known biliverdin-bound fluorescent proteins, the chromophore is noncovalently bound to the protein. We provide here a molecular rationale for the observed spectral properties of Sandercyanin.

Sandercyanin | walleye | blue protein | red fluorescent protein | UV radiation

Sandercyanin bound to biliverdin IXα is a blue colored protein–ligand complex isolated from the skin mucus of walleye (*Sander vitreus*), a commercial and sport fish common in North America (1, 2). Walleye are known to be sensitive to light and prefer being in the shade. Sandercyanin is produced seasonally, on the dorsal side of the fish, peaking in late summer and is found only in walleye that inhabit the northern latitudes (3). It is well established that the presence of an ozone hole over the North Pole results in increased UV radiation in these latitudes. Schaefer et al. (3) have hypothesized a link between the production of Sandercyanin in walleyes in these lakes and increased UV radiation. UV radiation is known to cause breakdown of heme in blood to biliverdin IXα (BLA) that may be excreted through urine and also secreted through the skin (4, 5). In the skin mucus, BLA combines with Sandercyanin—possibly produced in sacciform cells in the epidermis of the fish (3). BLA binding results in the formation of the blue-colored tetrameric Sandercyanin. We propose that this protein complex protects the walleye from UV radiation by acting as a natural sunscreen.

Serendipitously, Sandercyanin also shows interesting fluorescence properties. Fluorescent proteins (FPs) have many applications in cell biology (6–8). Genetically expressed FPs are used as biosensors in vivo to monitor a wide range of intracellular phenomena. GFPs from *Aequoria victoria* (9) as well as other spectral homologs have presented a number of advantages in understanding cellular processes (10–12). Of these, far-red-emitting FP variants are of particular interest due to low scattering and high signal-to-noise ratio (13, 14). In the past decade, near infrared (NIR) fluorescent protein tags developed from bacterial phytochromes have been successfully used in deep tissue imaging in animals (15, 16). Kumagai et al. (17) recently reported UnaG, a green fluorescent protein from the Japanese eel with bilirubin-inducible fluorescence. Furthermore, there are newly engineered red fluorescent proteins, archearhodopsin (Arch) (18) and CRABPII (19), which bind to retinoic acid or its analog as their chromophore. Here we report Sandercyanin—the first red fluorescent protein described from a vertebrate—a fluorescent protein with endogenous noncovalent ligand-inducible far-red fluorescence that could be engineered for use in biological applications including deep-tissue imaging.

In this paper, we describe the spectral properties and molecular structures of Sandercyanin. Sandercyanin shows strong far-red fluorescence in the mucosal cells of walleye, when excited by UV radiation (λ_{max} = 375 nm). Monomeric Sandercyanin is an 18.6-kDa protein and is the smallest far-red fluorescent protein discovered to date. Sandercyanin has interesting spectral properties of having a large spectral shift, which makes it distinct from other known fluorescent proteins. Sandercyanin may be engineered as a chromophore-inducible far-red fluorescent protein of small size and large spectral shift with potential use as a fluorescent protein marker.

Results

Structure and Properties of the Native Protein Purified from Walleye. Sandercyanin was first observed in blue vesicles in the mucosa of the North American walleye (1, 3). Yu et al. (1) have previously reported the purification of the native blue protein from walleye mucus and the partial protein sequence of Sandercyanin, which suggested that it belongs to the lipocalin family of proteins (20, 21). They also showed that BLA is noncovalently bound to the protein.

Significance

Recently it has been observed that the North American walleye is turning blue. The increased blue color is an adaptation to increased exposure to UV radiation. We identified that the blue pigment (Sandercyanin) is a complex of a protein and biliverdin—a breakdown product of heme. We report here that the blue pigment shows bright red fluorescence when excited with UV light. Elucidation of crystal structures and spectral properties of Sandercyanin lead us to hypothesize that the protection to damaging UV radiation happens by absorption of the UV light and its emission in the lower energy red wavelength. Interestingly, one can think of a number of applications where ligand-induced red fluorescent proteins can be useful.

Author contributions: S.G., S.K.P., D.T.G., and S.R. designed research; S.G., C.-L.Y., D.F., S.S., S.K.P., W.F.S., and S.R. performed research; S.K.P. and S.R. contributed new reagents/analytic tools; S.G., C.-L.Y., D.F., S.K.P., and S.R. analyzed data; and S.G., S.K.P., W.F.S., and S.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. C.F.L. is a Guest Editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5F6Z, 5EZ2, and 5F1E).

[1]Deceased July 24, 2014.

[2]To whom correspondence should be addressed. Email: ramas@instem.res.in.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1525622113/-/DCSupplemental.

It was initially thought that the walleye's blue color arose either from a copper-containing compound or a commensal cyanobacterium in the scales of walleye. Several attempts to isolate possible color-producing bacteria failed. We surmised that the protein is coded by the fish genome. We prepared DNA and carried out whole-genome sequencing experiments of the walleye. We determined the full-length gene sequence of the blue protein by partial assembly of the whole genome based on mapping of known internal peptides (1). The resulting gene sequence encoded for a secretion signal peptide of 19 amino acids in the N terminus of the protein. Alignment of protein sequence with other proteins in the database shows that Sandercyanin has homology to the lipocalins, apolipoprotein D from *Larimichthys crocea* (accession no. KKF23255.1) and an unannotated peptide from *Tetraodon nigroviridis* (accession no. CAF98955.1). We discovered that the isolated blue vesicles show bright red fluorescence (Fig. 1*A*) when excited with UV light. To understand the molecular basis of binding of BLA and the observed large spectral shift in fluorescence (Fig. 1*B*), we solved the crystal structure of native Sandercyanin (Fig. 1*C*) and extended our studies to include recombinant proteins.

Recombinant Sandercyanin and Its Properties. The gene obtained from the whole-genome sequencing was synthesized without the signal peptide and cloned for bacterial expression. Preliminary expression of Sandercyanin in *Escherichia coli* resulted in low yield of soluble, functional protein in the cytosol (*SI Appendix*, Fig. S1*A*). However, a significant quantity of Sandercyanin was expressed as inclusion bodies. Inclusion bodies, when solubilized using chemical denaturants and refolded, resulted in pure, monodisperse, and functional protein. This method of protein preparation was hence used for the studies reported here. We observed that Sandercyanin not bound to BLA (hereafter called apo-Sandercyanin) predominantly exists as a small nonfluorescent colorless monomer in nature but oligomerizes to a blue-colored homotetramer of 75 kDa on binding to BLA (BLA is green in color; Fig. 2*A* and *SI Appendix*, Fig. S1 *B* and *C*). On titration of apo-Sandercyanin with increasing concentrations of BLA, the fraction of tetramer increases. However, there is no observed dimer fraction in any intermediate concentration, suggesting that Sandercyanin dimer is a transient species and that equilibrium exists between monomer and tetramer forms. Circular dichroism studies show that Sandercyanin is made of β-sheets. Binding of BLA to Sandercyanin induces chirality in the BLA as seen from appearance of significant absorbance bands at 375 and 630 nm (Fig. 2*B*) (22).

Spectral Properties of Sandercyanin. Spectroscopic properties of native and recombinant purified BLA-bound Sandercyanin shows absorbance maxima at 280, 375, and 630 nm at physiological pH (pH 7.4; Fig. 2*C*) and a strong far-red fluorescence maxima at 675 nm when excited at 375 or 630 nm (Fig. 1*B*). BLA, by itself, has a large emission peak of ∼450 nm when excited at 375 nm and negligible fluorescence in the far-red region. Binding of BLA to apo-Sandercyanin results in significantly enhanced redshift in the fluorescence of BLA in its changed environment to far-red region (*SI Appendix*, Fig. S2*A*). Molar extinction coefficient of Sandercyanin–BLA complex (holo-Sandercyanin) at 375 and 630 nm are 21,000 $M^{-1}{\cdot}cm^{-1}$ and 13,500 $M^{-1}{\cdot}cm^{-1}$, respectively, and the quantum yield is determined to be 0.016. It has been reported that the quantum yield of biliverdin dimethyl ester measured at 710 nm in ethanol is 10^{-4} (23). Further, the fluorescence spectrum is broad in the far-red region with minimal overlap with the excitation spectra in the blue region (Fig. 2*D*). Our data also show that recombinantly expressed holo-Sandercyanin has similar spectral properties (*SI Appendix*, Fig. S2*B*) to the native protein purified from the walleye mucus.

The K_d of binding of apo-Sandercyanin to BLA measured using fluorescence spectroscopy is 6.4 μM (*SI Appendix*, Fig. S2*C*). The binding curve shows positive cooperativity with a Hill constant of 2.1. Sandercyanin does not bind to other BLA-like compounds, including hemin, bilirubin, or esterified BLA derivatives (*SI Appendix*, Fig. S2*D*).

To understand the fluorescence properties of BLA in different environments, we performed experiments with free BLA in a number of solvents and pH conditions. Free BLA shows increased

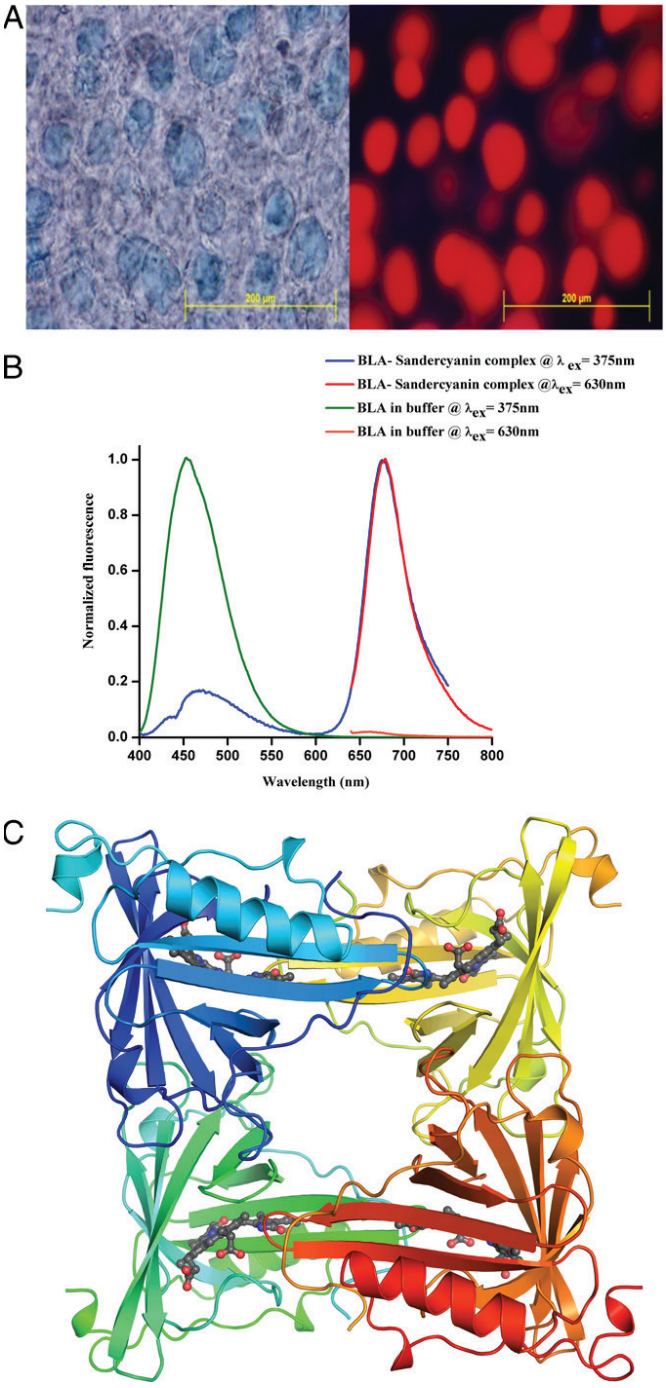


Fig. 1. Structure and properties of the native Sandercyanin purified from walleye. (*A*) Mucus from North American walleye (*Sander vitreus*) appear blue under bright field and shows intense red fluorescence on excitation at 375 nm. (*B*) Fluorescence spectra of BLA and native BLA–Sandercyanin complex on excitation at 375 or 630 nm. Unbound BLA (blue peak at 450 nm) appears as a small hump in the BLA-bound complex. (*C*) Cartoon representation of tetrameric structure of BLA-bound Sandercyanin. Each subunit of the tetramer binds to one BLA molecule (represented in ball-and-stick form).

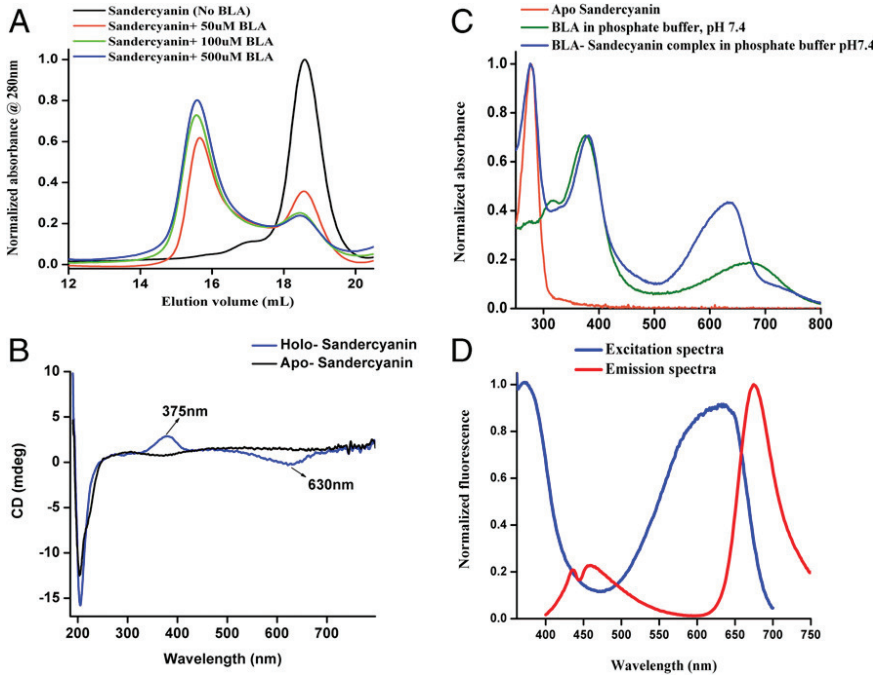


Fig. 2. Recombinant Sandercyanin and its properties. (*A*) Size-exclusion chromatography traces showing BLA-induced tetramerization in recombinant Sandercyanin. The concentration of BLA is color-coded in accordance with the chromatogram. (*B*) Circular dichroism spectra of apo-Sandercyanin showing presence of beta-strand secondary structure and BLA-induced chirality in the holo-Sandercyanin. (*C*) Absorption spectra of BLA and purified apo- and BLA-bound Sandercyanin. The spectra of apo-Sandercyanin and BLA–Sandercyanin complex are normalized to 280 nm. The BLA spectra are normalized to BLA–Sandercyanin complex at 380 nm. (*D*) Data of overlap between excitation and emission spectra of holo-Sandercyanin, showing minimal overlap.

far-red fluorescence on changing its environment from aqueous to hydrophobic medium (*SI Appendix*, Fig. S3*A*). A similar trend was observed on increasing the viscosity of the medium with polyethylene glycol (*SI Appendix*, Fig. S3*B*) and in the pH range of 8.8–9.5 (*SI Appendix*, Fig. S3*C*).

Crystal Structures of Apo- and Holo-Sandercyanin Reveal Molecular Basis of Ligand Binding. To correlate the biochemical and photophysical properties of Sandercyanin with the atomic structure, we crystallized and determined structures of native and recombinant proteins by X-ray crystallography (Figs. 1*C* and 3 *A* and *D*). The different forms of the protein crystallized in different conditions of buffer, salt, and precipitant concentrations (*SI Appendix*, Fig. S4*A*). Initially, we determined a structure of native holo-Sandercyanin at 2.2 Å resolution using a multiple-wavelength anomalous dispersion (24, 25) method because there was no homologous model available. Native protein crystals were soaked in $AuCl_3$, and data were collected at the Au–LIII edge. Further, structures of recombinant holo and apo forms of Sandercyanin were determined at 1.8 and 2.6 Å resolution, respectively, by the molecular replacement method using the native holo-Sandercyanin structure as a template. All crystal structures show that Sandercyanin is a tightly packed tetramer (Fig. 1*C*). In the native and the recombinant BLA complexes, each monomer binds noncovalently to one BLA molecule. Similar to known structures of lipocalins, Sandercyanin structure is an eight-stranded antiparallel β-barrel, capped by a long loop closing the barrel and an external α-helix (Fig. 3*A*) (20, 21). The barrel encloses a hydrophobic environment around the ligand (Fig. 3*B*). Further, there are two conserved intramolecular disulphide bonds (Fig. 3*A*) at the N and C termini. We also found that native Sandercyanin is *N*-glycosylated at position (Asn) 83 (*SI Appendix*, Fig. S4*C*). Similar to many secreted proteins in eukaryotes, glycosylation may be essential for its stability in the external environment (26, 27).

A closer look at the crystal structure shows that BLA is accommodated at the center of the barrel and assumes a ZZZssa configuration (Fig. 3*C*) (28, 29). The vinyl groups of rings A and D are buried deep in the cavity, whereas the propionate side chains of ring B and C are located near the entrance of the barrel. The ligand is mostly planar and is stabilized by steric and stacking interactions with aromatic amino acids (Fig. 3*D*). The residues Phe-55 and His-108 stack with BLA pyrrole rings B and C, respectively. Mutation of Phe-55 to alanine abolished BLA binding, suggesting that this aromatic stacking interaction plays a crucial role in binding. D-ring rotation in Sandercyanin, which has been extensively studied in bacteriophytochromes (29–32), is hindered by Tyr-116 and Tyr-142. We also observed interaction of propionate groups with Lys-57 and Lys-87, which may play a significant role in stabilizing the chromophore in the binding pocket. BLA also forms water-mediated hydrogen bonds (Fig. 3*E*) with His-108, Asn-77, and Tyr-65 through well-ordered water molecules. To understand the role of the water in the observed fluorescence properties, we carried out experiments to test the change in fluorescence on replacing the proton (H) in the water with Deuterium (D) (H_2O to D_2O). An increase in protein fluorescence intensity (quantum yield) on titration of the protein solution with D_2O was observed (*SI Appendix*, Fig. S5 *A* and *B*).

To further investigate the structural changes at the ligand-binding pocket, we determined a structure of apo-Sandercyanin. Although apo-Sandercyanin exists as a monomer in solution, it crystallizes as a tetramer, likely due to the high concentration of protein in the crystalline form. The overall structure is highly similar to BLA-bound protein tetramer, without any significant changes at the oligomerization interface. However, in the absence of BLA, the residues of the closing loops Lys-54–Lys-57 are disordered, which supports the observation that BLA–Phe-55 stacking interaction is essential to keep this loop ordered. Further, we observed conformational changes in the residues in the ligand-binding pocket (*SI Appendix*, Fig. S5*C*) near the D-ring and B-ring propionate of BLA; however, their interaction with their neighboring residues in the protein stabilizes them in the absence of the ligand.

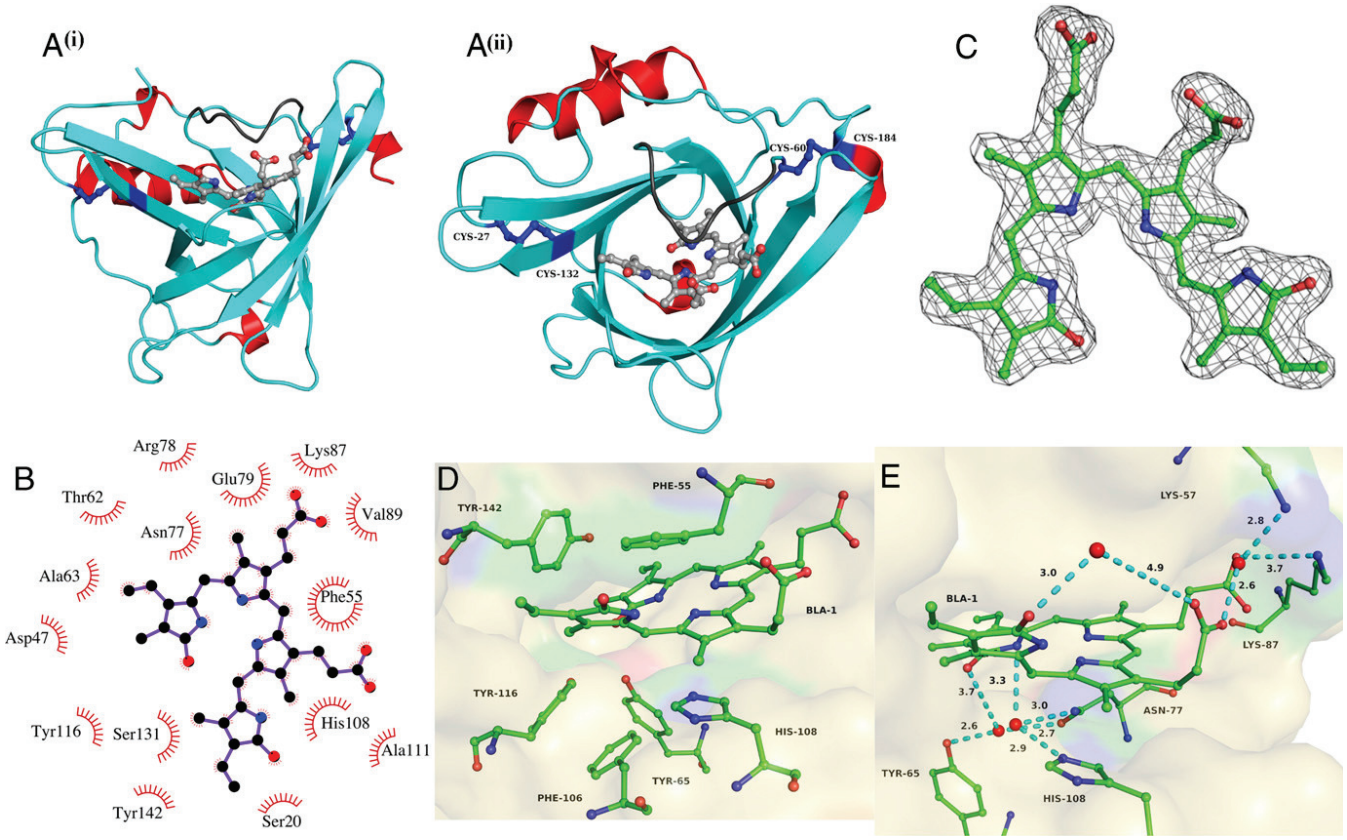


Fig. 3. Crystal structures of apo- and holo-Sandercyanin reveal molecular basis of ligand binding. (*A*) Structures of single subunit of Sandercyanin in two perpendicular orientations with BLA binding at the center of the lipocalin fold and enclosed by a long loop (shown in dark gray ribbon) on the top. Disulphide bonds between C27–C132 and C60–C184 are marked as blue sticks. (*B*) Residues surrounding BLA in the ligand-binding pocket shown in LigPlot representation. Most residues in the periphery of BLA are hydrophobic in nature. (*C*) 2_{Fo-Fc} electron density maps contoured at 1 rms around BLA in the refined structure of holo-Sandercyanin showing the ZZZssa configuration. (*D*) Interaction of aromatic residues with pyrrole rings of BLA in the binding pocket. (*E*) Detailed view of ionic and water-mediated H-bond interactions of BLA with its surrounding residues in the ligand- binding pocket.

On comparing the crystal structures of native Sandercyanin purified from walleye mucus to recombinantly expressed holoprotein, we observed conformational changes in the first two residues at the N-terminal Met-20 and Phe-21 (*SI Appendix*, Fig. S4*B*). In native Sandercyanin, Ser-20 is positioned toward the D-ring, whereas Met-20 in recombinant protein is directed outward, flipping the aromatic ring of Phe-21 toward the ligand. However, conformation of BLA remains the same. There are no significant changes in the overall secondary structure and position of the residues involved in glycosylation (*SI Appendix*, Fig. S4*C*). These results suggest that binding of BLA and fluorescent properties of Sandercyanin are minimally perturbed by changes in the N terminus and/or presence of glycosylation.

There are a number of interactions mediated by BLA that seems to stabilize one of the dimeric interfaces (*SI Appendix*, Fig. S4*D*). The Ser-138 and Leu-135 backbone forms water-mediated H-bonds with C-ring carboxylate and D-ring carbonyl groups, respectively. Moreover, the vinyl group of D-ring coordinates with the hydrophobic residues of the neighboring subunit. These interactions could possibly favor BLA-induced oligomerization in Sandercyanin. The tetramer is formed as a dimer of dimers. The dimer–dimer interface is made up of only protein–protein interaction; this is also stabilized by H-bonding via solvent molecules and hydrophobic interaction between amino acids. Overall, both interfaces present a two-fold symmetrical arrangement of residues (Fig. 1*C*). It is not very obvious from the structures why one would not observe a stable dimer form in solution.

Discussion

Sandercyanin, present in the North American walleye, binds BLA (a breakdown product of heme by UV radiation), which results in the mucus on the dorsal side acquiring a blue color; it exists as homotetramer, with each subunit having a size of 18.6 kDa and binds BLA in a noncovalent fashion. Our solution-state experiments reveal that oligomerization of Sandercyanin from monomeric apoprotein is promoted by addition of BLA. The observed Hill constant of 2.1 also suggests that BLA binding and oligomerization are interrelated cooperative processes. This protein shows interesting fluorescence properties, and is, by far, the smallest (in sequence), ligand-inducible far-red fluorescent protein reported from a vertebrate. Sandercyanin has one of the largest spectral shifts observed to date among fluorescent proteins characterized, with excitation at 375 nm and emission maximum at 675 nm. The functional and structural properties of native and recombinant Sandercyanin are similar, suggesting that glycosylation has minimal effect on the binding of BLA and spectral properties of Sandercyanin.

Solution studies with free BLA show enhanced red fluorescence with decreased polarity and increased viscosity of its surrounding media. We also observed that BLA fluorescence is pH dependent. Our high-resolution crystal structures of Sandercyanin reveal that BLA binds in a hydrophobic pocket in the center of the barrel. There is substantial pi-stacking interaction and water-mediated H-bonding between the protein and BLA. Our results from the solution studies of BLA, together with functional and structural data of Sandercyanin reported here, lead us to propose that hydrophobicity, specific stacking interactions, and H-bonding network

in the ligand-binding pocket enable BLA to dissipate the energy on excitation and generate far red fluorescence with a large spectral shift. Also, previous reports on red fluorescent proteins with intrinsic chromophore suggest that H- bonding and pi–pi stacking interaction play significant role in shifting fluorescence spectra of protein (12, 14). For instance, mCherry, mKate, and DsRed red fluorescent proteins are engineered for longer emission wavelength by perturbing the interactions between chromophore and the protein (33).

Lipocalins were initially characterized as lipid-binding proteins. However, a few biliverdin-binding lipocalins have been discussed in literature (1). These are known to bind biliverdin IX gamma isoform as the chromophore and impart blue color. It is not known, however, if these biliverdin-binding proteins have any fluorescent properties. Sandercyanin is the first lipocalin where biliverdin-inducible fluorescence is demonstrated. A comparison of the structure of Sandercyanin with previously reported Insecticyanin (PDB ID code 1BBP) structure (34) and bilin-binding protein (PDB ID code 1Z24) from *Pieris brassicae* (35) shows that the interactions between the protein and chromophore are conserved. One would predict that these proteins also show similar changes to biliverdin fluorescence.

Sandercyanin differs from previously reported fluorescent bilin-binding phytochromes (36, 37) on its chromophore binding modes. In phytochromes, one of the pyrrole rings of the chromophore associates covalently with a cysteine of the apoprotein (28, 29). Sandercyanin structures neither reveal the presence of any cysteine within close proximity to biliverdin nor show any other covalent association. Topologically, the closest residue in Sandercyanin to this cysteine residue in the phytochromes is Asp-47. It is important to note that these proteins have a completely different fold, and hence the mode of binding is very different. Bacteriophytochromes are well-studied photoswitches; their mechanism of photoconversion and structures of biliverdin in red (Pr) to far-red (Pfr) absorption (29–31) have been determined by time-resolved (30, 38) and pump-probe methods (39). Despite the difference in topology of the binding site, it would be interesting to study whether Sandercyanin has photoswitching properties similar to bacteriophytochromes. In bacterial phytochromes, it has been proposed that proton transfer and hydrogen bond interactions play a significant role in determining their fluorescence quantum yield (28). Excited-state proton transfer (ESPT) in GFP (40–43) and its variant proteins has been known for decades, which is critical in the observed redshift in their fluorescence emission (12, 14, 44).

Deuterium concentration-dependent quantum yield of the holo-Sandercyanin suggests that a solvent-mediated proton-relay tautomerization (a variety of ESPT) mechanism may be one of the excited nonradiative processes operative in the fluorescence properties (45). Replacement of exchangeable hydrogen by slower deuterium is expected to reduce the nonradiative ESPT process leading to increase in fluorescence intensity of Sandercyanin. There are several networks that can be observed from the crystal structure (Fig. 3*E*); this can be either water-mediated, given the presence of well-ordered water molecules in the structure (as seen in bacterial phytochromes), or direct interactions with the protein. Saha et al. (46) have proposed the presence of electron transfer pathways as a possible mechanism in redshifted emission of GFP. Understanding of the electron transfer pathways in Sandercyanin would be interesting because they play a role in energy loss during fluorescence, giving rise to large spectral shifts. Future studies will explore proton transfer pathways in Sandercyanin through systematic mutagenesis.

It has been postulated that the interaction of the dipole of the chromophore with asymmetric charge distribution can lead to splitting of energy levels, leading to changes in the fluorescence properties of the chromophore. This effect is called the Stark effect (47–50). To elucidate if the Stark effect plays a role in the cause of redshift in the fluorescence of BLA on binding to Sandercyanin, we calculated the electrostatic potential of the Sandercyanin structure using the Adaptive Poisson–Boltzmann Solver. Our calculations reveal significant charge differences in the BLA-binding pocket of Sandercyanin (*SI Appendix*, Fig. S6). Linke et al. (51) have used quantum mechanical calculations to describe the direction of the transient dipole of BLA in its different isomerization states. The conformation of the ligand in the structures determined here corresponds to what is described in the paper as $\nu(C_1 = O)$ in the ZZZssa geometry. There is a significant electrostatic field that is not perpendicular to the dipole of the BLA. These observations support the idea that the internal Stark effect in Sandercyanin, similar to what has been described for tryptophan in proteins (47), may contribute to the redshift in the fluorescence of BLA on binding to the protein.

In the future, we aim to engineer stable monomers of Sandercyanin with increased brightness for use as fluorescent protein fusion tags (6–8, 13, 52). Low quantum yield and tetramerization of Sandercyanin are undesirable characteristics for in vivo imaging experiments. Further, folding of Sandercyanin in the cytosol may be challenging due to the secretory nature of protein and the presence of multiple disulphide bonds. We are currently attempting to express Sandercyanin in eukaryotic cellular systems. Knowledge of the structure and the residues involved in the interactions with the chromophore will allow us to design mutations that show increased quantum yield. Knowledge of the residues involved in the stabilization of the protein–protein interfaces will allow us to rationally engineer monomeric proteins. This approach is not unique, because the most used fluorescent proteins also started as oligomers and with much lower brightness and were engineered for applications with the knowledge of the structure.

In the context of its natural occurrence in the North American walleye, Sandercyanin present in the mucus absorbs damage-causing UV radiation; it then dissipates this energy by emitting it in lower-energy red wavelength. This phenomenon may be a protective mechanism for walleyes in northern latitudes against increased levels of UV radiation due to recent arctic ozone depletion (3). An alternate explanation for the advantage provided by the blue color is the possible role of Sandercyanin in camouflaging and countershading in walleyes to avoid predation by northern pike (*Esox lucius*) (3). Our preliminary studies indicate that Sandercyanin shows antioxidant properties and may help walleyes to protect themselves from free radicals in the mucus produced on exposure to UV radiation (*SI Appendix*, Fig. S7). Though the exact mechanism is not known, previous reports reveal that biliverdin is a potential antioxidant in the biliverdin/bilirubin cycle in the presence of biliverdin reductase and helps cytoprotection (53–56). However, Dorazio et al. (57) have recently shown that biliverdin itself may have antioxidant properties, which might be an added advantage for walleyes for survival in their native environment. Schaefer et al. (3) have shown that the number and level of sandercyanin in the mucus of blue walleye increases seasonally, in the summer, as solar UV input increases. This is a remarkable example of nature protecting itself from human pollution by using the very product (biliverdin) formed from that pollution in self-defense.

Materials and Methods

Native Sandercyanin was extracted and purified from the mucus of blue forms of walleye from Northwest Ontario by chromatographic techniques as described previously (1). A putative amino acid sequence of the protein was determined from crystal structure of the native protein, confirmed, and corrected after partial genome sequencing of blue walleye. The derived sequence was synthesized and cloned into bacterial expression systems for further studies. Published methods were used for cloning, expression, purification, structure determination, and fluorescence measurements, and are described in detail in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Karthi Sivaraman in the Next Generation Sequencing Facility at Center for Cellular and Molecular Platforms, Bangalore; Kruttika Phalnikar, C. M. Nithyakala, Linda Westphal-Buth, and Wendie Dockstader for contributions at the early stages of this work; Hassan Behrali, Babu (BM14), and Rosmerie Friemann for data collection; Girish T. S. for critical reading of the manuscript; and the Industrial Macromolecular Crystallography

Association - Collaborative Access Team (CAT) and the General Medical Sciences and Cancer Institutes Structural Biology Facility at the Advanced Photon Source-CAT beam line at Advanced Photon Source and the Molecular Biology Consortium beam line (especially Jay Nix) for making available beam time for this project. Special thanks to Lokesh Gakhar for help at different stages. Support for this work was provided from Department of Biotechnology India Grant BT/PR5801/INF/22/156/2012, and partial support from the University of Iowa and the University of Wisconsin (to S.R.). We dedicate this work to the memory of Prof. David T. Gibson, a colleague and friend.

1. Yu CL, et al. (2008) Purification and properties of Sandercyanin, a blue protein secreted in the mucus of blue forms of walleye, *Sander vitreus*. *Environ Biol Fishes* 82(1): 51–58.
2. Hart JL (1973) *Pacific Fishes of Canada* (Fisheries Research Board of Canada, Ottawa).
3. Schaefer WF, Schmitz MH, Blazer VS, Ehlinger TJ, Berges JA (2014) Localization and seasonal variation of blue pigment (Sandercyanin) in walleye (*Sander vitreus*). *Can J Fish Aquat Sci* 72(2):281–289.
4. Brown BE, Bythell JC (2005) Perspectives on mucus secretion in reef corals. *Mar Ecol Prog Ser* 296:291–309.
5. Gagnon MM (2006) Serum biliverdin as source of colouration upon sexual maturation in male blue-throated wrasse *Notolabrus tetricus*. *J Fish Biol* 68(6):1879–1882.
6. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 90(3):1103–1163.
7. Day RN, Davidson MW (2009) The fluorescent protein palette: Tools for cellular imaging. *Chem Soc Rev* 38(10):2887–2921.
8. Shaner NC, Patterson GH, Davidson MW (2007) Advances in fluorescent protein technology. *J Cell Sci* 120(Pt 24):4247–4260.
9. Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509–544.
10. Müller-Taubenberger A, Anderson KI (2007) Recent advances using green and red fluorescent protein variants. *Appl Microbiol Biotechnol* 77(1):1–12.
11. Verkhusha VV, Lukyanov KA (2004) The molecular properties and applications of Anthozoa fluorescent proteins and chromoproteins. *Nat Biotechnol* 22(3):289–296.
12. Wiehler J, von Hummel J, Steipe B (2001) Mutants of Discosoma red fluorescent protein with a GFP-like chromophore. *FEBS Lett* 487(3):384–389.
13. Shcherbo D, et al. (2009) Far-red fluorescent tags for protein imaging in living tissues. *Biochem J* 418(3):567–574.
14. Subach FV, Piatkevich KD, Verkhusha VV (2011) Directed molecular evolution to design advanced red fluorescent proteins. *Nat Methods* 8(12):1019–1026.
15. Shu X, et al. (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* 324(5928):804–807.
16. Filonov GS, et al. (2011) Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat Biotechnol* 29(8):757–761.
17. Kumagai A, et al. (2013) A bilirubin-inducible fluorescent protein from eel muscle. *Cell* 153(7):1602–1611.
18. Mclsaac RS, et al. (2014) Directed evolution of a far-red fluorescent rhodopsin. *Proc Natl Acad Sci USA* 111(36):13034–13039.
19. Yapici I, et al. (2015) "Turn-on" protein fluorescence: In situ formation of cyanine dyes. *J Am Chem Soc* 137(3):1073–1080.
20. Flower DR (1996) The lipocalin protein family: Structure and function. *Biochem J* 318(Pt 1):1–14.
21. Flower DR, North ACT, Sansom CE (2000) The lipocalin protein family: Structural and sequence overview. *Biochim Biophys Acta* 1482(1-2):9–24.
22. Allenmark S (2003) Induced circular dichroism by chiral molecular interaction. *Chirality* 15(5):409–422.
23. Braslavsky SE, Holzwarth AR, Lehner H, Schaffner K (1978) The fluorescence of biliverdin dimethyl ester. *Helvetica Chimica Acta* 61(6):2219–2222.
24. Smith GD, Lemke CT, Howell PL (2007) Substructure determination in multiwavelength anomalous diffraction, single anomalous diffraction, and single isomorphous replacement with anomalous scattering data using Shake-and-Bake. *Methods Mol Biol* 364:183–196.
25. Son SK, Chapman HN, Santra R (2011) Multiwavelength anomalous diffraction at high X-ray intensity. *Phys Rev Lett* 107(21):218102 1.
26. Uzman A, Lodish H, Berk A, Zipursky L, Baltimore D (2000) The molecules of life. *Molecular Cell Biology*, eds Lodish H, et al. (Freeman, New York), 4th Ed, section 1.2.
27. Ferguson MAJ, Kinoshita T, Hart GW (2009) *Essentials of Glycobiology* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY), 2nd Ed.
28. Toh KC, Stojkovic EA, van Stokkum IHM, Moffat K, Kennis JTM (2010) Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome. *Proc Natl Acad Sci USA* 107(20): 9170–9175.
29. Salewski J, et al. (2013) Structure of the biliverdin cofactor in the Pfr state of bathy and prototypical phytochromes. *J Biol Chem* 288(23):16800–16814.
30. Takala H, et al. (2014) Signal amplification and transduction in phytochrome photosensors. *Nature* 509(7499):245–248.
31. Samma AA, Johnson CK, Song S, Alvarez S, Zimmer M (2010) On the origin of fluorescence in bacteriophytochrome infrared fluorescent proteins. *J Phys Chem B* 114(46):15362–15369.
32. Seibeck S, et al. (2007) Locked 5Zs-biliverdin blocks the Meta-RA to Meta-RC transition in the functional cycle of bacteriophytochrome Agp1. *FEBS Lett* 581(28):5425–5429.
33. Chica RA, Moore MM, Allen BD, Mayo SL (2010) Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc Natl Acad Sci USA* 107(47):20257–20262.
34. Holden HM, Rypniewski WR, Law JH, Rayment I (1987) The molecular structure of insecticyanin from the tobacco hornworm *Manduca sexta* L. at 2.6 A resolution. *EMBO J* 6(6):1565–1570.
35. Huber R, et al. (1987) Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 A resolution. *J Mol Biol* 198(3):499–513.
36. Murphy JT, Lagarias JC (1997) The phytofluors: A new class of fluorescent protein probes. *Curr Biol* 7(11):870–876.
37. Bhattacharya S, Auldridge ME, Lehtivuori H, Ihalainen JA, Forest KT (2014) Origins of fluorescence in evolved bacteriophytochromes. *J Biol Chem* 289(46):32144–32152.
38. Moffat K (2014) Time-resolved crystallography and protein design: Signalling photoreceptors and optogenetics. *Philos Trans R Soc Lond B Biol Sci* 369(1647):20130568.
39. Yang X, Ren Z, Kuk J, Moffat K (2011) Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature* 479(7373):428–432.
40. Meech SR, Tonge PJ (2009) Excited state dynamics in the green fluorescent protein. *J Photochem Photobiol A* 205(1):1–11.
41. Henderson JN, et al. (2009) Excited state proton transfer in the red fluorescent protein mKeima. *J Am Chem Soc* 131(37):13212–13213.
42. Beekman L, et al. (1997) Characterization of the light-harvesting antennas of photosynthetic purple bacteria by stark spectroscopy. 1. LH1 antenna complex and the B820 subunit from *Rhodospirillum rubrum*. *J Phys Chem B* 101:7284–7292.
43. Chattoraj M, King BA, Bublitz GU, Boxer SG (1996) Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer. *Proc Natl Acad Sci USA* 93(16):8362–8367.
44. Piatkevich KD, et al. (2013) Extended Stokes shift in fluorescent proteins: Chromophore-protein interactions in a near-infrared TagRFP675 variant. *Sci Rep* 3:1847.
45. Kasha M (1986) Proton-transfer spectroscopy. Perturbation of the tautomerization potential. *J Chem Soc Faraday Trans 2* 82(12):2379–2392.
46. Saha R, et al. (2013) Light driven ultrafast electron transfer in oxidative redding of green fluorescent proteins. *Sci Rep* 3:1580.
47. Vivian JT, Callis PR (2001) Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys J* 80(5):2093–2109.
48. Köhler M, Gafert J, Friedrich J, Vanderkooi JM, Laberge M (1996) Stark effect experiments in cytochrome c-type proteins: Structural hierarchies. *Biophys J* 71(1):77–85.
49. Moore LJ, Zhou H, Boxer SG (1999) Excited-state electronic asymmetry of the special pair in photosynthetic reaction center mutants: Absorption and Stark spectroscopy. *Biochemistry* 38(37):11949–11960.
50. Franzen S, Moore LJ, Woodruff WH, Boxer SG (1999) Stark-effect spectroscopy of the heme charge-transfer bands of deoxymyoglobin. *J Phys Chem B* 103:3070–3072.
51. Linke M, et al. (2013) Electronic transitions and heterogeneity of the bacteriophytochrome Pr absorption band: An angle balanced polarization resolved femtosecond VIS pump-IR probe study. *Biophys J* 105(8):1756–1766.
52. Stadler C, et al. (2013) Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat Methods* 10(4):315–323.
53. Baranano DE, Rao M, Ferris CD, Snyder SH (2002) Biliverdin reductase: A major physiologic cytoprotectant. *Proc Natl Acad Sci USA* 99(25):16093–16098.
54. Stocker R, McDonagh AF, Glazer AN, Ames BN (1990) Antioxidant activities of bile pigments: Biliverdin and bilirubin. *Methods Enzymol* 186:301–309.
55. Jansen T, Daiber A (2012) Direct antioxidant properties of bilirubin and biliverdin. Is there a role for biliverdin reductase? *Front Pharmacol* 3:30.
56. McDonagh AF (2010) The biliverdin-bilirubin antioxidant cycle of cellular protection: Missing a wheel? *Free Radic Biol Med* 49(5):814–820.
57. Dorazio SJ, et al. (2015) Singlet oxygen oxidation products of biliverdin IXα dimethyl ester. *Bioorg Med Chem* 23(24):7671–7675.

SUPPLEMENTARY MATERIAL

DETAILED MATERIALS AND METHODS:

Cloning, recombinant expression and purification of Sandercyanin.

Native Sandercyanin was extracted and purified from the mucus of blue forms of walleye from Northwest Ontario by chromatographic techniques as described previously (1). A putative amino acid sequence of the protein was determined from crystal structure of the native protein, confirmed and corrected after partial genome sequencing of blue walleye. Genome sequence revealed the presence of a secretion signal sequence which was not observed in the native crystal structure. The gene (without the signal peptide) was synthesized from GeneScript (Invitrogen) and cloned into pET21a bacterial expression vector between NdeI and HindIII cloning sites. Recombinant Sandercyanin was initially co- expressed with chaperons (cloned into pG-KJE8 plasmid from Clontech, Takara) in BL21*(DE3) cells. Cells were grown in minimal 1X M9 media till an OD_{600} of 0.6-0.7 and induced with 0.2mM isopropyl- thiogalactoside (IPTG) for protein expression. We maintained a temperature of $18^{o}C$ throughout the growth process. Since, this method gave very low yield of functional protein in the cytosol, we expressed Sandercyanin as inclusion bodies (IBs) in LB media and purified by refolding methods. The cell-pellet are re-suspended in 50mL of IB- wash buffer (20mM Tris.HCl, pH 7.5, 10mM EDTA and 1 % TritonX) and sonicated using macro-probe (Fisher Scientific) for 3 cycles of 3 min each with 10s on and 30 off pulses at 50% amplitude. The cell-lysate is centrifuged for 30 min at 13,000 r.p.m in Avanti J-26 XP centrifuge and JA17 rotor from Beckmann Coulter to obtain pure IBs (white residue). This is re-suspended in a solution containing 5M Guanidine.HCl, 50mM CAPS, 0.5mM phenylmethylsulfony fluoride (PMSF) and 1mM DTT, pH 7.5 and incubated at room temperature to solubilize the IBs. The denatured protein (20mg) is rapidly diluted in 25mL of buffer containing 1.1M Guanidine. HCl, 50mM Tris-HCl, pH 7.5, 50mM NaCl, 0.88mM KCl,10% glycerol, redox containing 5mM/1mM of reduced/ oxidized L- cysteine and 1µM BLA-hydrochloride (Santa Cruz Biotechnology, USA). This is then dialyzed using 3.5kDa MWCO tubing (Fisher Scientific) overnight in 2 L of the same buffer without guanidine.HCl. The refolded protein is concentrated using 3kDa Centricon (Millipore) and passed through a Superdex 200 analytical size exclusion column (GE Healthcare). Blue-colored protein fractions,

corresponding to size of 75kDa are collected, concentrated to 8mg/mL and used for crystallization. Apo- Sandercyanin is purified using the same method in buffer solutions without BLA and collected as monomeric protein by size- exclusion chromatography. The column is calibrated with standard proteins (Bio-rad).

UV-Visible Spectroscopy.

All experiments were performed with purified Sandercyanin samples (apo and holo forms) at pH 7.5 and room temperature. UV-Visible absorbance spectra of native and recombinant Sandercyanin were recorded from 200 to 800nm with ultraspec 2100 pro spectrophotometer from Amersham Biosciences. CD spectra were measured on JASCO J-815 Spectropolarimeter. Steady state fluorescence, excitation and binding studies were monitored on Horiba Jobin Yvon Fluoromax-4 fluorimeter. Data analysis was done using Origin6 and Origin8 software. For calculation of binding affinity, the data points were fitted using Graph Pad prism 6 using the equation given below:

$$\mathbf{Y = Bmax^{*}\ [L]^{h} / (K_d^{h} + [L]^{h})}$$

Bmax is the maximum value of specific binding, K_d is the ligand concentration [L] at half-maximum binding at equilibrium and h is the Hill slope.

Hydrogen was exchanged with Deuterium by increasing concentrations of D_2O to a standard protein concentration. This deuteriated protein were prepared in the same buffer, incubated for 15 minutes and monitored for its spectral properties.

Antioxidant assay.

The well established polyacrylamide gel- based method is used for demonstrating anti-oxidant property(2). In short, protein samples are run on 8% Tris- glycine native gel at low voltage (40V) to avoid protein denaturation. The gel is then incubated in 1% nitroblue tetrazolium (NBT) solution for 15 minutes in the absence of light. This is followed by incubation in riboflavin solution. The gel is exposed to blue light for 5 minutes. A change in color to blue (that off nitroblue tetrazolium (NBT)) is observed in presence riboflavin. Proteins with radical-scavenging property do not show any color change, which remain unstained on the gel. We used

superoxide dismutase (SOD) and bovine serum albumin (BSA) as positive and negative controls respectively, to monitor anti-oxidant activity.

Crystallization, data collection and structure refinement.

Crystallization of Sandercyanin (native and recombinant) were carried out at 4˚C in hanging drops vapor diffusion method using mosquito high- throughput crystallization system from TTP life sciences. Protein crystals were obtained in a number of conditions **(Fig. S4A)** and flash cooled after soaking in 10% ethylene glycol as cryo-protectant. MAD datasets for native Sandercyanin crystal data were collected from crystals soaked in Au (at ALS beam line 4.2.2). The structure was solved using the SOLVE-RESOLVE package (3). Recombinant apo- and holo- Sandercyanin datasets were collected at ID-29 and BM14 respectively in European Synchrotron Radiation Facility (ESRF, Grenoble, France). All images were indexed, integrated and scaled using HKL2000 (4) or d*TREK (5) or MOSFLM (6). Molecular replacement for recombinant proteins was performed using native Sandercyanin structure as template model and the structures refined with PHENIX. 2Fo-Fc map showed presence of positive density indicating presence of ligand in the core of each monomer subunit. Model building was done with Coot (7) and all structural illustrations were generated with PyMol (8). All parameters of data collection and refinement statistics are summarized in Supplementary - **Table1**.

APBS calculations. Adaptive Poisson-Boltzmann Solver (APBS) (9) calculations were done in PyMol using both apo and BLA bound forms of Sandercyanin.

REFERENCES TO SUPPLIMENTARY INFORMATION

1. Yu CL, et al. (2008) Purification and properties of Sandercyanin, a blue protein secreted in the mucus of blue forms of walleye, Sander vitreus. *Environ Biol Fishes* 82(1):51–58.
2. Jumee Na HI, Lee and K (2011) Expression and Purification of Recombinant Superoxide Dismutase (PaSOD) from Psychromonas arctica in Escherichia coli. *Bull Korean Chem Soc* 32(7):2405–2409.
3. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 55(4):849–861.
4. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D Biol Crystallogr* 67(4):235–242.
5. Pflugrath JW (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr Sect D Biol Crystallogr* 55(10):1718–1725.
6. Leslie AGW (2006) The integration of macromolecular diffraction data. *Acta Crystallogr Sect D Biol Crystallogr* 62(1):48–57.
7. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr* 60(12 I):2126–2132.
8. DeLano WL (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC wwwpymolorg* Version 1.:http://www.pymol.org.
9. Baker N a, Sept D, Joseph S, Holst MJ, McCammon J a (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 98(18):10037–41.

Table 1. X- ray data collection and refinement statistics

Crystal	Native Sandercyanin with BLA (PDB code 5F6Z)	Recombinant Sandercyanin with BLA (PDB code 5EZ2)	Recombinant Apo-Sandercyanin (PDB code 5F1E)
Data collection	R-AXIS IV (University of Iowa, Protein crystallography Facility)	BM14, ESRF	ID-29, ESRF
Space group	$P4_12_12$	$P6_322$	$P6_322$
Unit cell dimensions	93.4, 93.4, 246.4	159.3, 159.3, 84.2	158.8, 158.8, 84.8
Resolution range (A^0)	19.8-2.2 (2.3-2.2)*	35.99- 1.85 (1.9- 1.85)*	53.5-2.7 (2.8-2.7)*
Total reflections	180844	215082	654304
Unique reflections	52052 (5127)*	53181 (5194)*	17810 (1741)*
Multiplicity	3.4	4.0	36.7
Completeness (%)	98.6 (98.7)*	99.5 (97.5)*	99.98 (99.94)*
Mean I/sigma(I)	5.9 (3.7)*	15.5 (1.9)*	31.5 (7.8)*
Wilson B-factor	23.7	26.4	44.5
R-sym	0.155	0.066	0.13
R-factor	0.24 (0.37)*	0.19 (0.29)*	0.19 (0.23)*
R-free	0.31 (0.43)*	0.21 (0.30)*	0.24 (0.25)*
Number of atoms:			
Macromolecules	5079	2594	2556
Ligands	228	106	0
Water	324	313	83
Protein residues	672	338	334
RMS(bonds)	0.009	0.019	0.010
RMS(angles)	1.30	1.56	1.19
Ramachandran favored (%) (Outliers in parenthesis)	97(0)	98(0)	98(0)
Clashscore	11.23	6.78	7.15
Average B-factor:	40.1	25.2	25.4
Macromolecules	40.1	24.1	25.5
Solvent	38.6	34.5	21.6

*Statistics for the highest-resolution shell are shown in parentheses.

SUPPLEMENTARY FIGURES

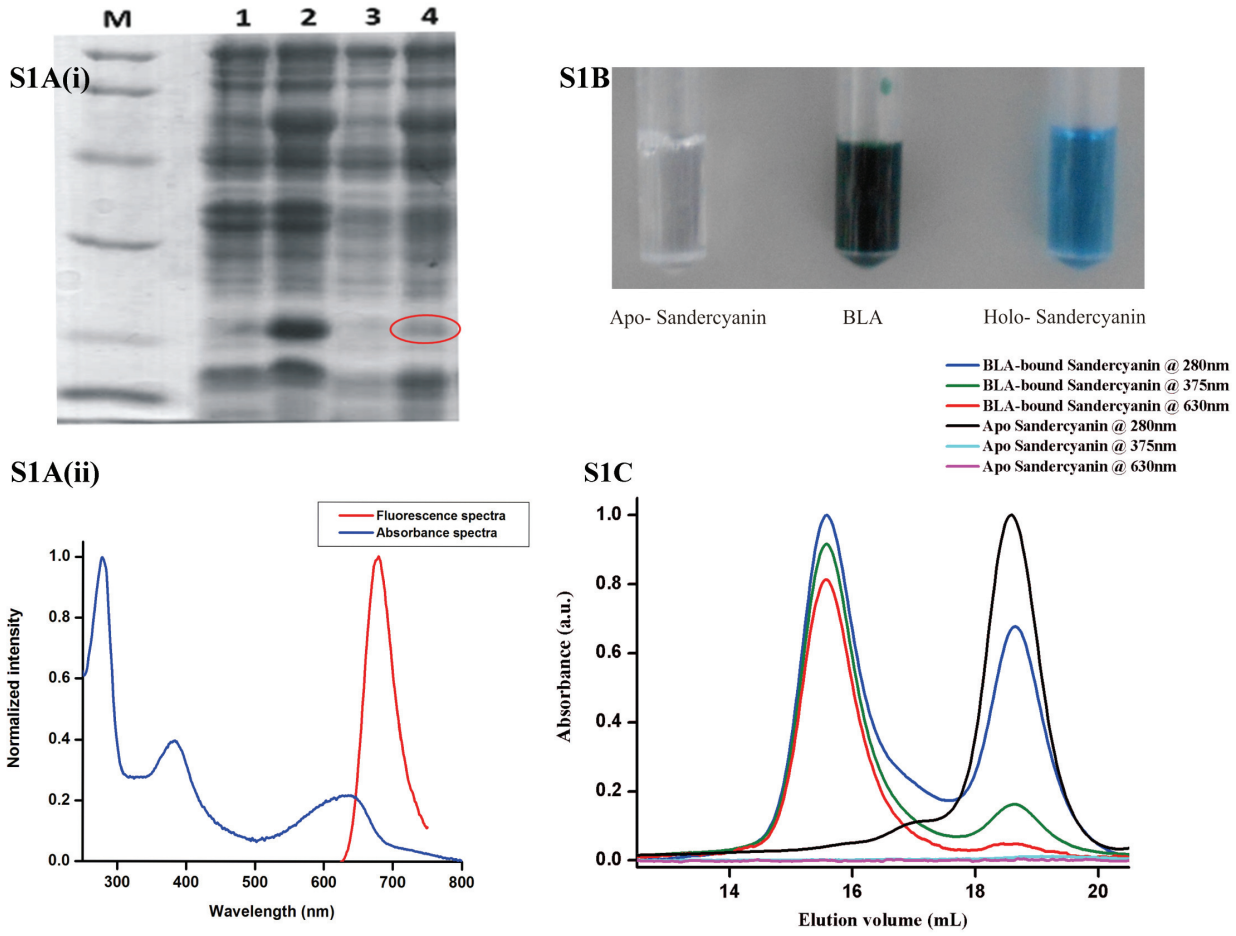


Fig. S1. Recombinant Sandercyanin and its properties. (A) (i) SDS-PAGE shows cytosolic expression (red- circle) of recombinant Sandercyanin in *E. coli*. Expression of IPTG- induced protein (lanes 2 and 4) were checked against un-induced samples (lanes 1 and 3). (ii) Absorbance and fluorescence emission spectra (excitation at 375 nm) of soluble fraction of recombinant Sandercyanin-BLA complex – same as in lane 4 in (i). (B) Generation of blue color of Sandercyanin on adding BLA (green) to apo- protein (colorless). (C) Size- exclusion chromatograms of BLA- bound Sandercyanin monitored at different wavelengths. Apo protein does not absorb at 375nm or 630nm. The chromatograms are normalized to the tetramer peak. The peak around elution volume of 15-16mL corresponds to the tetramer peak. The peak that elutes at about 19ml corresponds to the monomer peak.

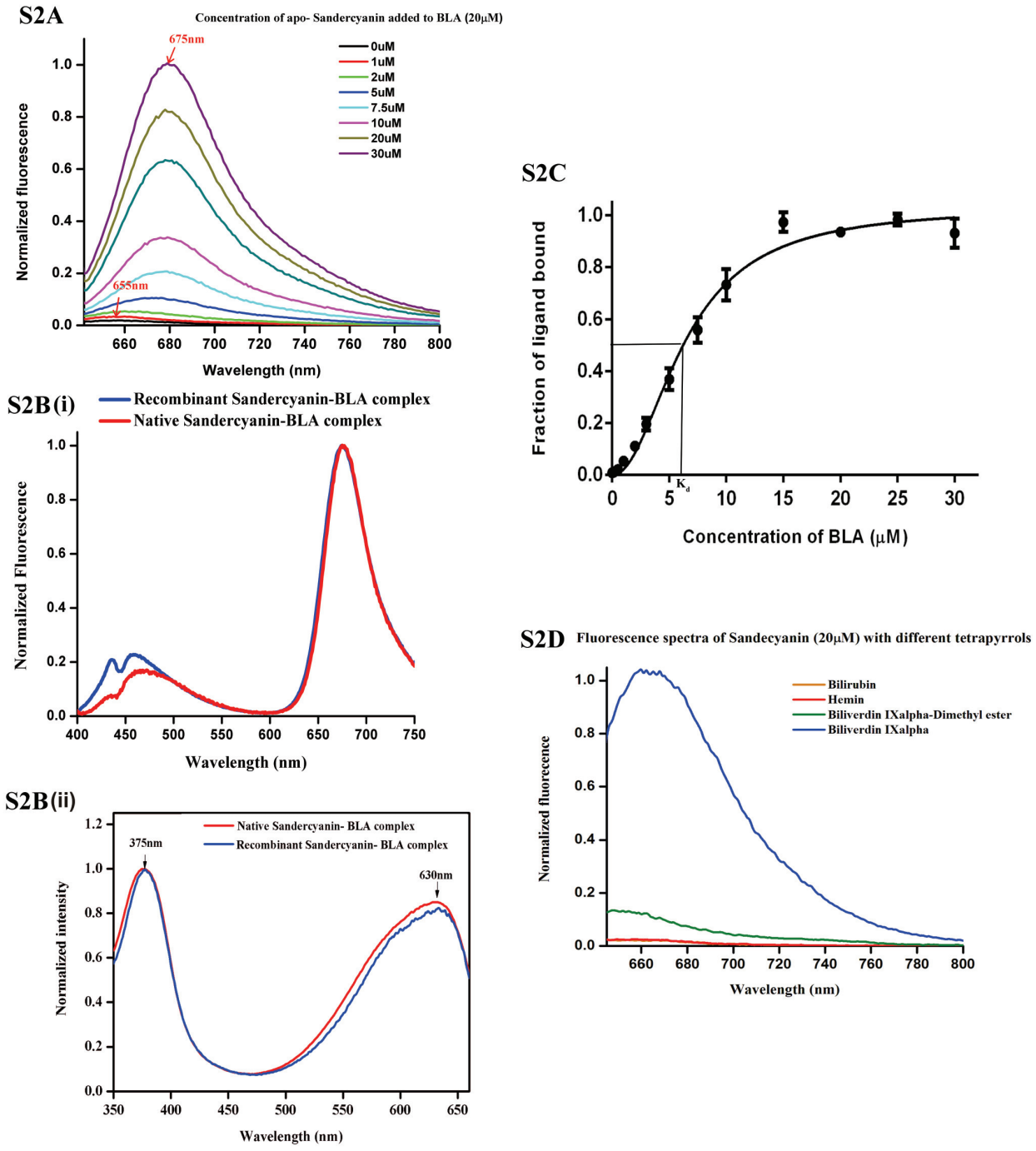
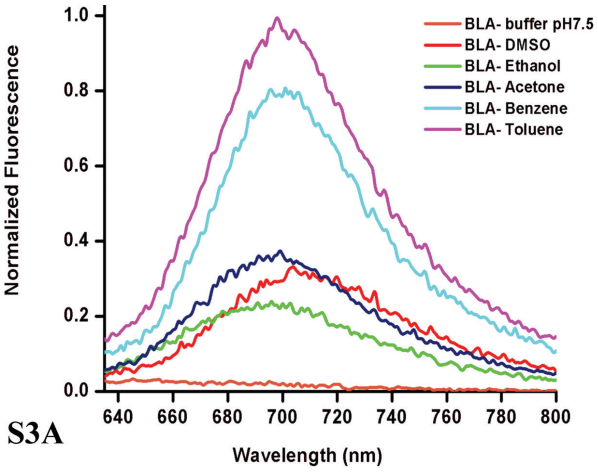
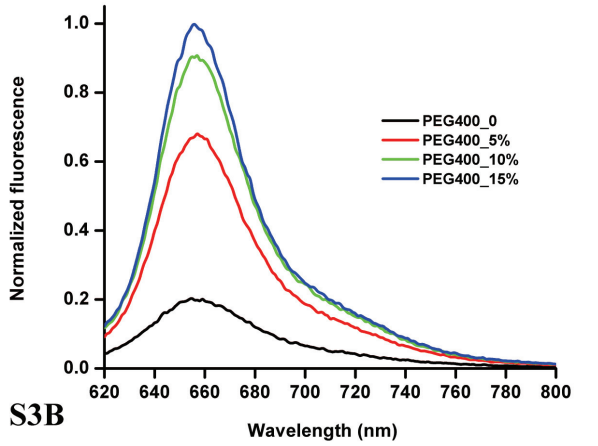


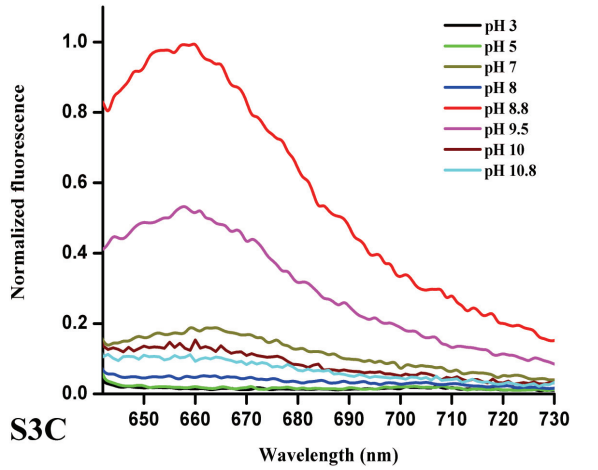
Fig. S2. Spectral properties of Sandercyanin. (A) Titration of BLA (20uM) with increasing apo-Sandercyanin concentration show enhanced red-shifted fluorescence of BLA with maxima at 675 nm on excitation at 375 nm. Similar spectra were observed on excitation at 630 nm. (B) Comparison of (i) emission and (ii) excitation spectra of native (blue) and recombinant (red) BLA- Sandercyanin complex show that they are similar. (C) Titration of apo-Sandercyanin (20μM) with BLA measured at 675nm shows K_d of 6.4μM and Hill slope of 2.1. (D) Comparison of emission spectra of Sandercyanin on binding with different tetrapyrrols shows specificity towards BLA. All measurements were taken in phosphate buffer pH 7.4 at room temperature under similar experimental conditions.



S3A



S3B



S3C

Fig. S3. Fluorescence properties of free BLA in solution. Effect of (A) non-polar solvent (B) increased viscosity of solvent (PEG 400) and (C) pH, on fluorescence spectra of free- BLA in solution (10μM) monitored at excitation at 375nm. Similar spectra were observed on excitation at 600nm. BLA shows increased far-red fluorescence on increased hydrophobicity and viscosity. It shows a maximum fluorescence around pH 8.8 – pH 9.5.

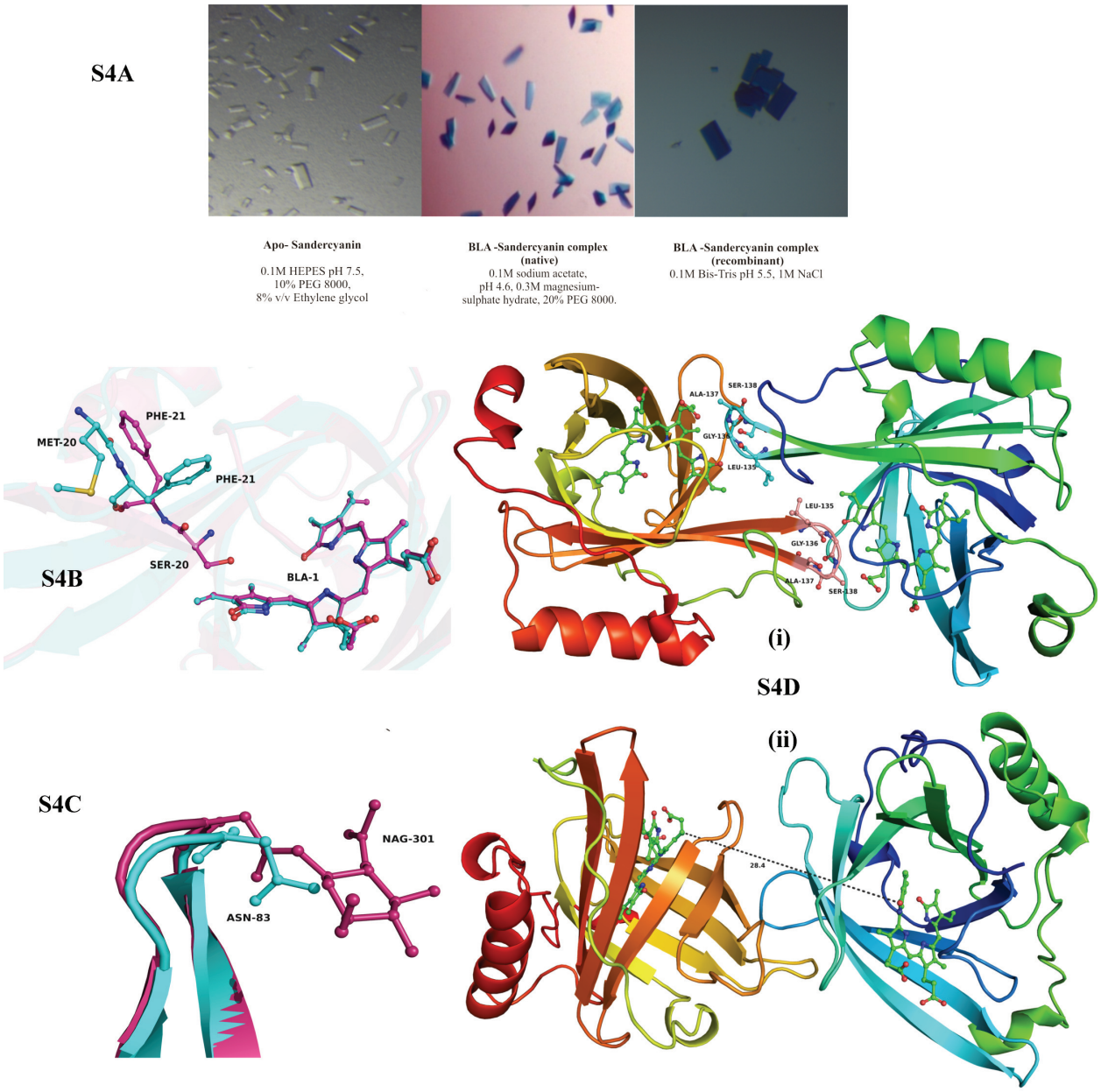


Fig. S4. Crystal structures of apo and BLA – bound native and recombinant Sandercyanin. (A) Crystals and crystallization conditions of apo (colorless) and BLA-bound (blue) forms of native and recombinant Sandercyanin. Comparison of structures of native (magenta) and recombinant (cyan) Sandercyanin-BLA complex, showing (B) rotamer changes of Phe21 at the N-terminus and (C) position of glycosylation in native protein bound to BLA. (D) Close up view of (i) BLA- protein interaction at one dimer interface and (ii) protein- protein interactions at the dimer-dimer interface.

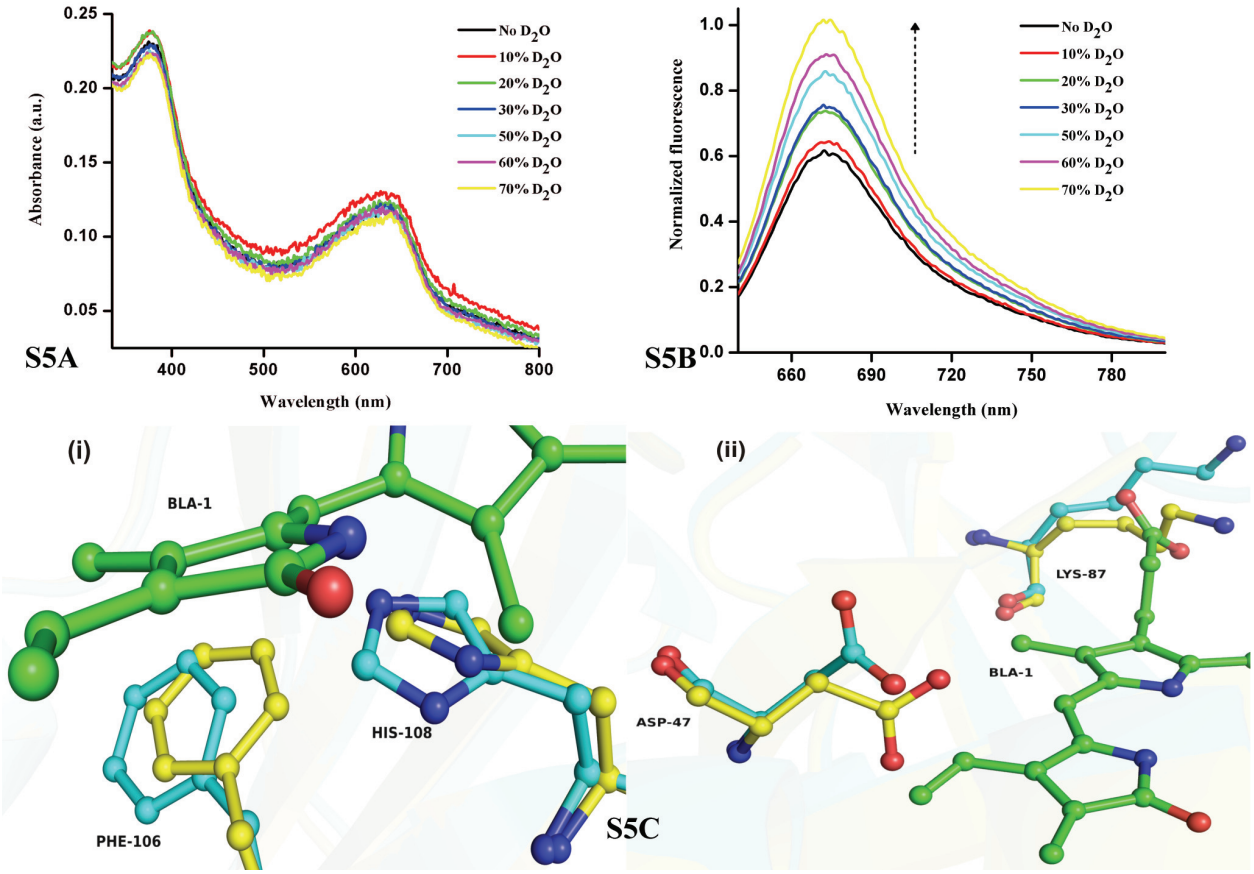


Fig. S5. Role of excited state proton transfer (ESPT) in perturbing fluorescence properties of Sandercyanin. Absorbance (A) and emission spectra (B) of Sandercyanin monitored on excitation at 375nm at different D_2O concentration, suggesting influence of proton transfer on the fluorescence properties of Sandercyanin (20μM). Similar observation was made on excitation at 630nm. (C) Structural insights (i and ii) of apo (yellow) and holo (cyan) Sandercyanin showing conformational changes in amino acids around the pyrrole rings of BLA.

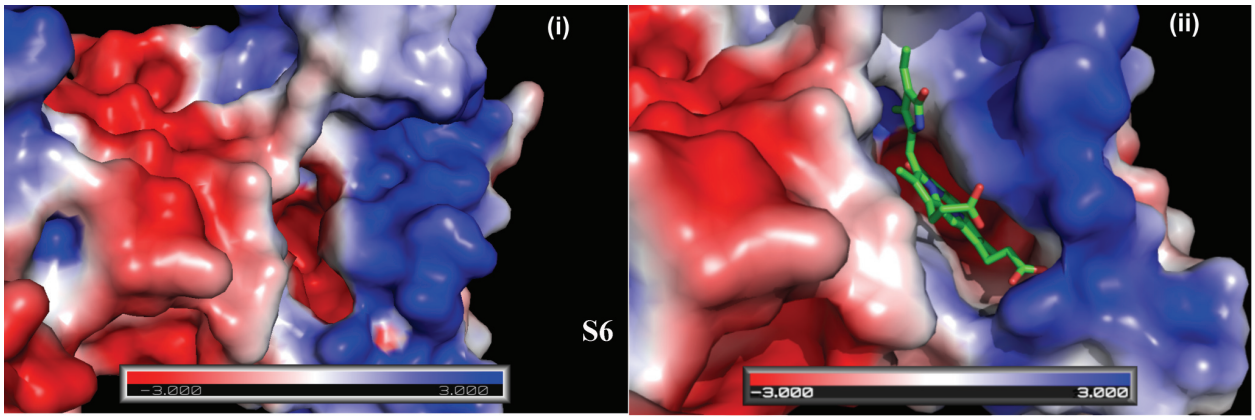


Fig. S6. APBS calculations. Electrostatic potential of Sandercyanin shows asymmetric charge distribution in the binding pocket of Sandercyanin. (i) shows the empty pocket and (ii) with bound BLA.

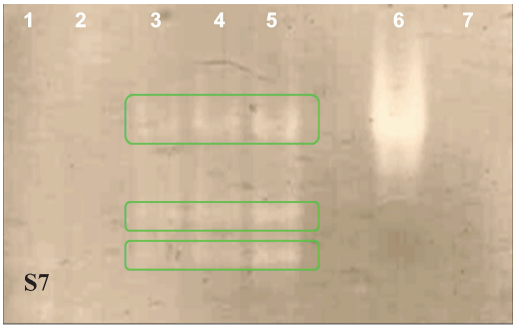


Fig. S7. Anti-oxidant properties of Sandercyanin. Native PAGE showing radical scavenging activity of BLA- bound Sandercyanin (lanes 3, 4 and 5) in presence of riboflavin, Tetra methyl ethylenediamine (TEMED) and nitroblue tetrazolium (NBT). Apo Sandercyanin (lanes 1 and 2) do not show radical scavenging activity. The antioxidant activity is compared with standard human superoxide dismutase (SOD from sigma, lane 6) and bovine serum albumin (lane 7) as positive and negative controls respectively.