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

Photochemical & Photobiological Sciences

FULL PAPER

Photo-control of DNA binding by an engrailed homeodomain-photoactive yellow protein hybrid

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

A. Kumar,[a] A. M. Ali[a,b] and G. A. Woolley[a†]

A photo-controlled version of the engrailed homeodomain (zENG) was created by inserting the homeodomain into a surface loop of a circularly permuted version of photoactive yellow protein (cPYP). The two proteins fold independently as judged by NMR and fluorescence denaturation measurements. In the dark, the affinity of the zENG domain for its cognate DNA is inhibited >100-fold compared to wild-type zENG. Blue-light irradiation of the hybrid protein leads to enhanced conformational dynamics of the cPYP portion and a two-fold enhancement of DNA binding affinity of the zENG domain. These results suggest that insertion into a surface loop into cPYP can be a general approach for conferring an initial level of photo-control on a given target protein. Focussed mutation/selection strategies may then be used to enhance the degree of photo-control.

Introduction

Engrailed is a DNA-binding protein of the homeodomain transcription factor family.[1, 2] Engrailed proteins are found in a wide range of organisms and play critical roles in embryological development.[2-4] Since the timing and location of engrailed activity is central to normal development, a method for spatiotemporal control of engrailed activity would be useful research tool.[3, 5] As a first step in this direction, we sought to develop a genetically encoded, photo-controlled version of the engrailed homeodomain.

Previously we engineered a photo-controlled version of the DNA-binding protein GCN4 by fusing it with photoactive yellow protein (PYP).[6, 7] PYP from *H. halophila* is an extensively studied small (125 residue) water soluble protein that contains the chromophore *p*-coumaric acid and undergoes a blue-light triggered photocycle.[8-10] The dark state of PYP is very stable[11] and folds reversibly making it a good starting point for protein engineering efforts. [12] Whereas the dark state of PYP exists in a compact well-folded conformation, the light state[‡] shows extensive conformational exchange, as measured by NMR,[13] and has characteristics of a molten globule.[14, 15] The light state of PYP reverts spontaneously to the dark state on a timescale of seconds.

To produce a photo-controlled version of the transcription factor GCN4, we replaced the N-terminus of PYP by the C-terminal dimerization domain of GCN4.[6, 7] This approach was made possible by a significant degree of sequence homology between PYP and GCN4. In an effort to develop a PYP-based scaffold for photo-control that could be applied to a variety of target protein sequences, we designed a circularly permuted version of PYP (designated cPYP).[16] In cPYP, the sequence targeted for photo-control is inserted between the original N- and C-termini of wild-type PYP and new termini are generated at a surface loop corresponding to wild-type residues 114 and 115. A cPYP protein with a small flexible linker (GGSGGSGG) as the target sequence was shown to fold reversibly and undergo a photocycle very similar to that of wild-type PYP.[16]

Here we report efforts to produce a photo-controlled engrailed homeodomain by replacing the flexible linker GGSGGSGG in cPYP with the engrailed homeodomain sequence from zebrafish (zENG). The resulting hybrid protein is designated cPYP-zENG.

[a] Dept. of Chemistry, University of Toronto, 80 St. George St., Toronto Canada, M5S 3H6.
[b] Dept. of Medicinal Chemistry, Faculty of Pharmacy, Assiut University, Assiut, Egypt
† Correspondence: awoolley@chem.utoronto.ca

Photochemical & Photobiological Sciences Accepted Manuscript

Please do not adjust margins

Results and discussion

The N- and C-termini of the zENG homeodomain are separated by approximately 22 Å,[17] a distance similar to that between the ends of the insertion site in cPYP.[16, 18] Thus we expected that, in the dark, both zENG and cPYP could adopt normal folded states. However availability of the DNA binding helix of zENG was expected to be limited to some degree by fusion to cPYP. Since blue light irradiation allows cPYP to adopt a much looser conformation,[11] we expected that the irradiated state of cPYP-zENG would present fewer restrictions on zENG, allowing it to access its DNA binding site. A schematic showing the overall design concept for photo-controlled cPYP-zENG in Figure 1.

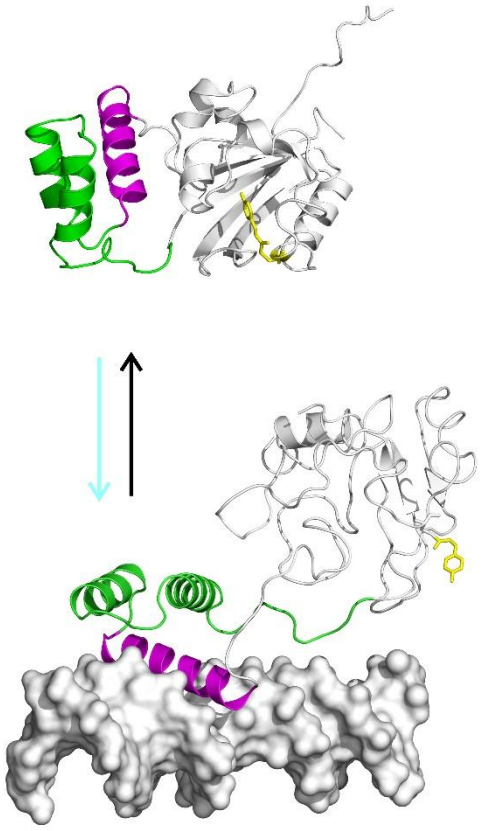

Figure 1. Design concept for photo-control of DNA binding by cPYP-zENG. The cPYP domain is coloured white and zENG domain is coloured green. The *p*-coumaric chromophore is shown as yellow sticks. DNA is shown in surface representation. (Top) In the dark-adapted state the DNA binding helix of zENG (magenta) is sterically blocked by cPYP. (Bottom) Irradiation with blue light loosens PYP enabling DNA binding by zENG. Models were built using coordinates for DNA-bound engrailed (2HOT), dark-adapted PYP (1NWZ) and light state PYP (2KX6).

We synthesized a gene encoding cPYP-zENG and expressed the protein in *E. coli* with a C-terminal His-tag to enable purification via Ni-ion affinity chromatography. Although the cPYP protein with a GGSGGSGG insert is monomeric, some inserts can lead to the production of domain-swapped dimers.[19] Size exclusion FPLC chromatography of cPYP-zENG showed that, although some dimeric species are produced during overexpression in *E. coli*, the dominant species is the monomeric form and this does not appear to interconvert with the dimeric species on the timescale of days during normal handling of the protein (see Experimental section). In addition, blue light irradiation does not cause a change in oligomeric state as detected by size exclusion FPLC (data not shown). Thus all the characterization of cPYP-zENG, and the single point mutant to be described below refers to the monomeric species.

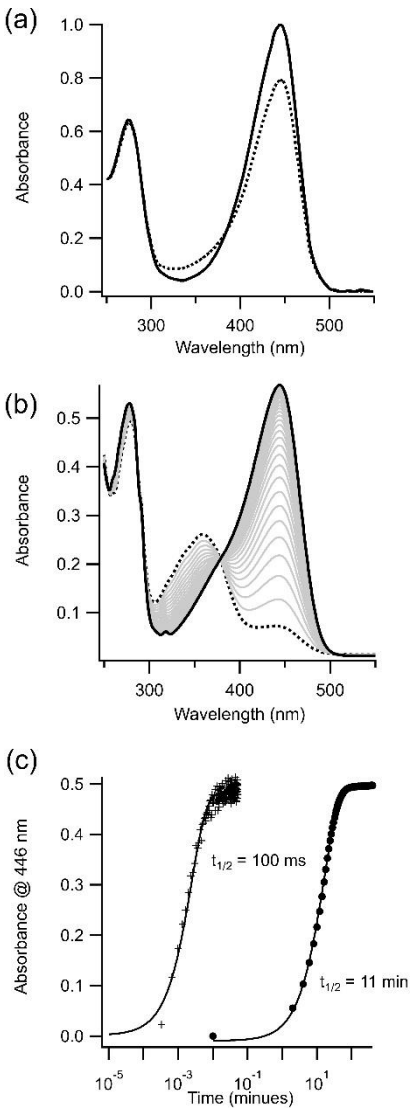

Figure 2. (a) UV-Vis spectra of dark-adapted (solid line) and blue light irradiated (dotted line) cPYP-zENG. (b) UV-Vis spectra of dark-adapted (solid line) and blue light irradiated (dotted line) cPYP-zENG-M173A. Grey lines shows spectra obtained during thermal relaxation of the protein. (c) Thermal relaxation of cPYP-zENG (+) and cPYP-zENG-M173A (●). All spectra were obtained in tris-acetate EDTA buffer, 100 mM NaCl, pH 7.5, 20°C.

Figure 2a shows the UV-Vis spectrum of cPYP-zENG under dark-adapted conditions. This is indistinguishable from the

Photochemical & Photobiological Sciences Accepted Manuscript

This journal is © The Royal Society of Chemistry 2015

Please do not adjust margins

spectrum of cPYP and wild-type PYP.[16] The spectrum is only slightly altered under continuous blue light irradiation with a high power LED (~50 mW/cm^2) (Fig. 2a, dotted line) indicating that thermal relaxation of the light state is fast. Thermal relaxation was measured directly using a gated irradiation source. The half-life observed was ~100 ms (Fig. 2c). Rapid relaxation means that very bright light sources are required for substantial steady-state population of the light state. Since high intensity irradiation can lead to unwanted heating or photochemical side reactions in downstream applications, we decided to slow the thermal relaxation process of cPYP-zENG by targeted mutation.

Mutations at the Met100 site in wild-type PYP are known to substantially slow the thermal relaxation of the light state.[20, 21] We therefore made a Met to Ala mutation at the corresponding site in cPYP-zENG (cPYP-zENG-M173A). The UV-Vis spectrum of dark-adapted cPYP-zENG-M173A is shown in Figure 2b together with spectra obtained after blue light irradiation. The half-life for thermal relaxation was observed to be 11 min, approximately 6000 times longer than cPYP-zENG. Note that the M173A mutation introduces a shoulder on the short wavelength side of the peak in UV-Vis spectrum of the dark-adapted protein. This feature, seen previously with cPYP and other PYP mutants,[16, 22] has been attributed to changes in H-bond networks that lead to the presence of a fraction of protonated trans chromophore. These changes occur without substantial change in overall protein conformation.[23, 24]

Next, we wished to test whether the zENG and cPYP domains of the hybrid proteins were both folded in the dark. As shown in Figure 3a, each domain contains a single buried Trp residue. Intrinsic protein fluorescence emission, which is primarily due to Trp residues, is usually affected by protein folding/unfolding. We monitored intrinsic protein fluorescence as a function of guanidinium HCl (Gdn) concentration. Figure 3b shows that both cPYP-zENG and cPYP-zENG-M173A undergo a two stage folding/unfolding process. With cPYP-zENG the first folding/unfolding transition occurs with a midpoint near 2 M GdnHCl and the second transition occurs near 3.5 M GdnHCl. For cPYP-zENG-M173A the first transition occurs at 1.5 M GdnHCl and the second near 3.5 M. We attribute the first transition to folding/unfolding of the cPYP domain and the second transition to the folding/unfolding of the zENG domain in each case. This is because (i) cPYP (M100) with a GGSGGSGG insert in place of zENG shows only one transition with a midpoint of 2 M (Fig. 3b), (ii), zENG alone unfolds near 3.5 M GdnHCl and (iii) the yellow colour of the proteins (monitored via absorbance at 446 nm) is lost during the first unfolding transition (Fig. 3c). As expected the M173A mutation lowers the stability of the cPYP domain.[16] These data are consistent with both cPYP and zENG domain being folded in the dark.

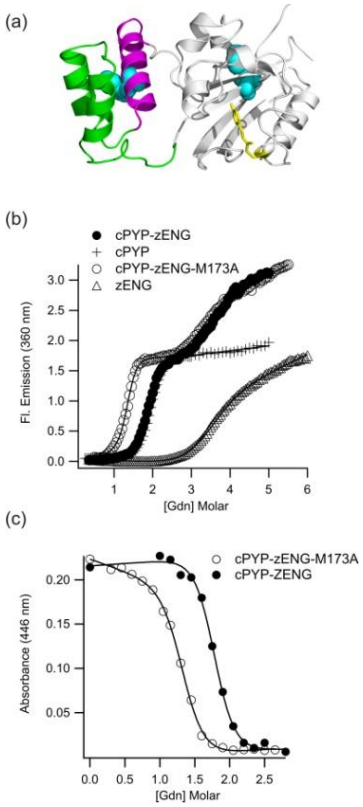

Figure 3. (a) Model of the dark-adapted cPYP-zENG hybrid showing the two Trp residues in cyan-coloured spacefill representation. (b) Fluorescence emission monitored during refolding titrations. (c) Absorbance at 446 nm as a function of denaturant concentration for cPYP-zENG and cPYP-zENG-M173A. All spectra were obtained in tris-acetate EDTA buffer, 100 mM NaCl, pH 7.5, 20^{o}C.

To further analyse the structure of the zENG-cPYP-M173A hybrid we performed solution state NMR measurements with uniformly ^{15}N labelled protein. Figure 4a shows the NH HSQC spectrum of cPYP-zENG-M173A under dark-adapted conditions. Figure 4c shows the NH HSQC spectrum of dark-adapted cPYP with a GGSGGSGG linker in place of the zENG domain as well as the spectrum of zENG alone. The sum of these spectra closely resembles the spectrum of dark-adapted cPYP-zENG-M173A (Fig. 4a). This result confirms that the zENG and cPYP domains are independently folded in the dark. When the sample was exposed to blue light delivered through a fibre optic cable into the spectrometer the spectrum in Figure 4b was obtained. As expected the majority of cross-peaks disappear from the spectrum due to conformational exchange processes on the millisecond timeframe in light state cPYP.[16, 25, 26] However a subset of cross-peaks remains, and these can be attributed primarily to the folded zENG domain. This can be seen by comparing the spectrum of irradiated cPYP-zENG-M173A (Fig. 4b) with that of zENG alone (Fig. 4d). Thus, blue light irradiation leads to enhanced conformational dynamics of the cPYP domain while leaving the zENG normally folded, and in a state competent to bind DNA.

Photochemical & Photobiological Sciences Accepted Manuscript

Please do not adjust margins

Photochemical & Photobiological Sciences

FULL PAPER

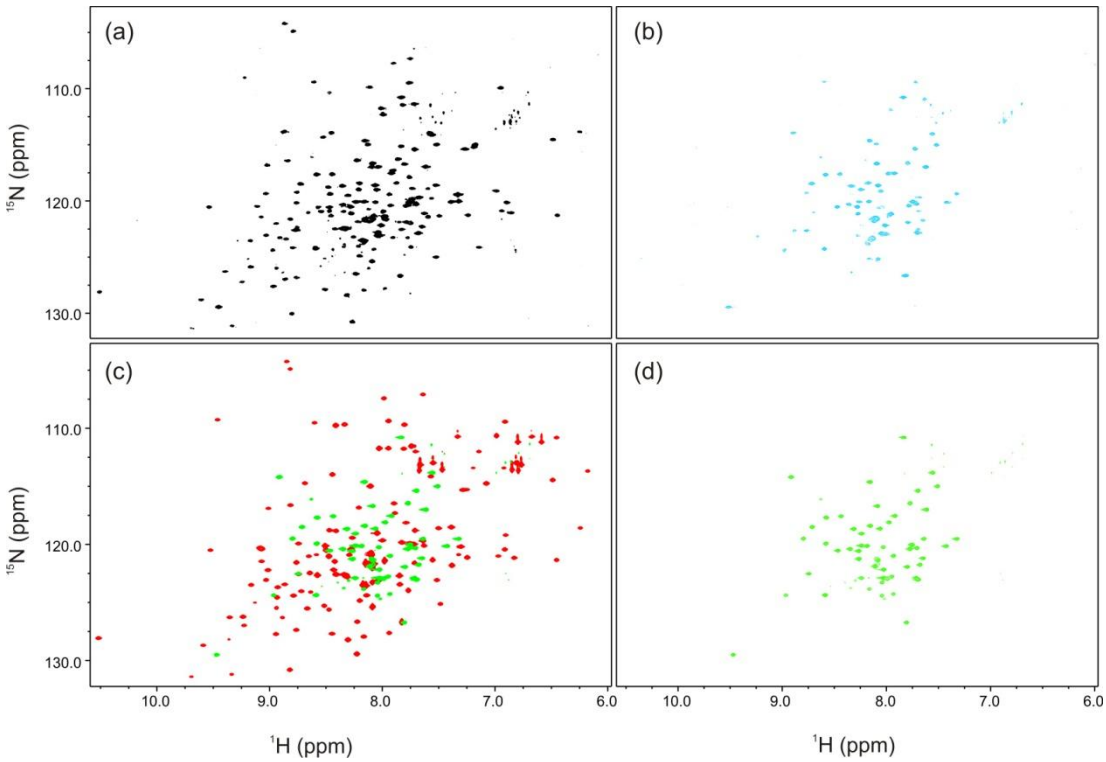

Figure 4. (a) NH-HSQC spectrum of dark-adapted cPYP-zENG-M173A. (b) NH-HSQC spectrum of blue-light irradiated cPYP-zENG-M173A. (c) NH-HSQC spectrum of dark-adapted cPYP (red) and zENG (green). (d) NH-HSQC spectrum of zENG. All spectra were obtained in tris-acetate EDTA buffer, 100 mM NaCl, pH 7.5, 20^{o}C.

To assess the light dependence of DNA binding by cPYP-zENG-M173A, we performed an electrophoretic mobility shift assay (EMSA).[27] Fluorescently labelled (Cy5) target DNA was mixed with nonspecific DNA as described in the Experimental section. Samples were prepared with increasing amounts of cPYP-zENG-M173A and incubated in the dark or irradiated with blue light. The samples we subjected to gel electrophoresis, which separates free DNA from protein-DNA complexes. By quantifying the fraction of DNA bound as a function of cPYP-zENG-M173A protein concentration, the apparent affinity of the protein for DNA can be assessed in the dark-adapted or irradiated state. A representative set of gels is shown in Figure 5 together with the derived binding curves. For reference, a binding curve of zENG alone to the same target DNA is shown. It is clear from these data that fusion of zENG to cPYP greatly reduces the apparent affinity of the zENG domain for its target sequence, by a factor of at least 100. Moreover, the affinity for the target is measurably higher under blue light irradiated conditions than it is in the dark (Fig. 5a). Thus it appears the overall strategy described here for producing a photo-controlled zENG homeodomain is feasible. Fusion to cPYP does inhibits DNA binding in the dark as intended. This inhibition is relieved somewhat in the light. By varying the length and sequence of the linker between the zENG domain and the cPYP domain it may be possible to restore higher DNA affinity to the light state selectively. The development of high throughput approaches for screening for light switchable DNA binding sequences would facilitate this effort. [28]

Conclusions

Rational design of a photo-controlled zENG homeodomain DNA binding protein was accomplished by inserting the zENG homeodomain sequence into a surface loop of a circularly permuted PYP protein. While the degree of photo-control achieved is modest, this rational design approach provides a direct route to photo-control of a given protein activity that could subsequently be improved by focussed mutation/selection strategies.

Photochemical & Photobiological Sciences Accepted Manuscript

Please do not adjust margins

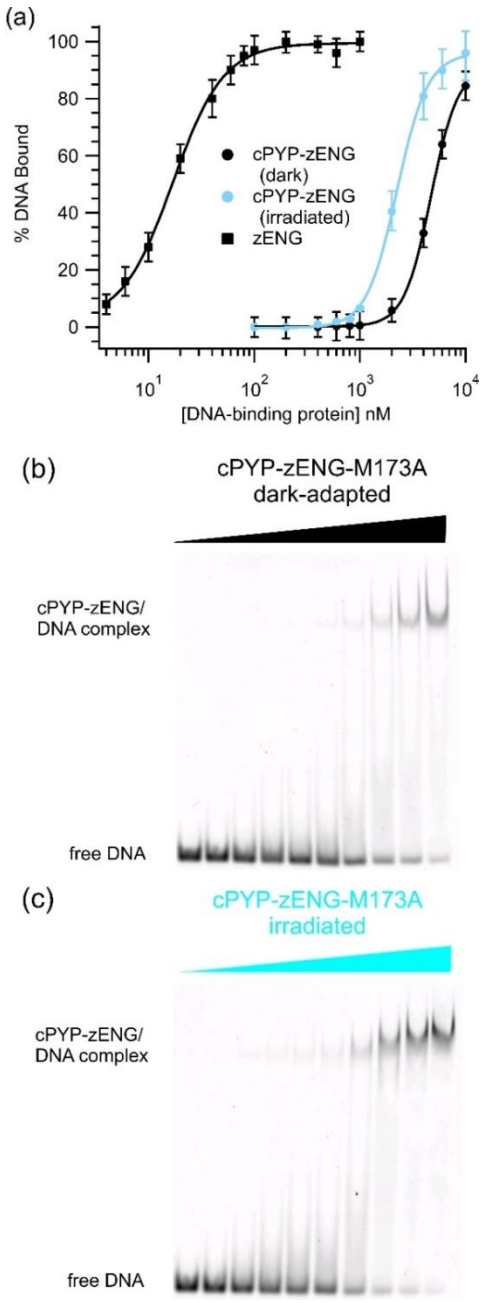

Figure 5. (a) Binding curves for zENG and cPYP-zENG-M173A (dark/light) to target DNA. EMSA data are fitted to the Hill equation. Apparent K_ds are: zENG 17±5 nM, cPYP-zENG-M173A-light: 2200±100 nM; cPYP-zENG-M173A-dark: 4700±100 nM (b) EMSA gel for showing DNA binding by increasing concentrations of dark-adapted cPYP-zENG-M173A (indicated by the black wedge). (c) EMSA gel for showing DNA binding by increasing concentrations of irradiated cPYP-zENG-M173A (indicated by the blue wedge).

Experimental

Gene synthesis and site-directed mutagenesis

DNA encoding cPYP-zENG (codon optimized for *Escherichia coli*) was synthesized by BioBasic Inc. (Toronto, ON) and inserted into the pET24b(+) vector using NdeI and HindIII restriction sites. The mutation to create cPYP-zENG-M173A was introduced using standard molecular biology protocols. Primers to introduce the mutation were designed with PrimerX (http://www.bioinformatics.org/primerx/) following the Stratagene QuickChange Protocol (Agilent, Inc.) and purchased from ACGT Corp. (Toronto, ON). Primers for the M173A mutation were:
Forward:
GAATATACTTTCGACTACCAGGCAACTCCGACCAAGGTTAAAGTC
Reverse:
GACTTTAACCTTGGTCGGAGTTGCCTGGTAGTCGAAAGTATATTC

The expressed sequence of cPYP-zENG-M173A is shown below (the M173 residue is underlined):

MKGDSYWVFVKRVEDKRPRTAFTAEQLQRLKAEFQTNRYLTEQRRQ
SLAQELGLNESQIKIWFQNKRAKIKKAMEHVAFGSEDIENTLAKMDD
GQLDGLAFGAIQLDGDGNILQYNAAEGDITGRDPKQVIGKNFFKDVA
PCTDSPEFYGKFKEGVASGNLNTMFEYTFDYQ**M**TPTKVKVHMKKAL
SKLAAALEHHHHHH

Protein expression and purification

Expression and purification of cPYP-zENG and cPYP-zENG-M173A were adapted from the work of Devanthan et al. as described previously. [16] DNA (0.2 ng) was transformed into BL21*(DE3) competent cells and plated onto Luria-Bertani (LB) agar plate containing 30 µg/mL kanamycin. The following day, a single colony was used to inoculate 25 mL of LB broth that had been supplemented with kanamycin (30 mg/mL). The 25 mL overnight culture was used to inoculate 1 L of LB supplemented with 30 µg/mL kanamycin. Cells were grown at 37 ℃ until an OD_{600} of ~0.6 was reached and then induced with 1 mM IPTG. The temperature was adjusted to 25℃, and the cells were grown for a further 1.5 h before 25 mg of activated chromophore dissolved in 1 mL of ethanol was added to the medium. The synthesis of the activated chromophore, 4-hydroxycinnamic acid *S*-thiophenyl ester, was performed as described[29] except that the product was not recrystallized. The cells were grown for a further 6 h before being centrifuged to separate the medium from the protein-containing cell pellet.

The pellet was resuspended in lysis buffer (50 mM sodium phosphate (pH 8.0), 300 mM sodium chloride, and 5 mM magnesium chloride) and stored frozen at −20°C. The resuspended cell pellet was sonicated in pulses (intervals of 10 seconds) on ice for 6 min and then centrifuged at 12K rpm for 1 h to separate the supernatant from the pellet. The protein-containing supernatant was filtered using 0.45 mm Millex® filter unit then loaded onto a Ni-NTA column that was

Photochemical & Photobiological Sciences Accepted Manuscript

equilibrated with lysis buffer. The yellow color became associated with the resin. The resin was washed with 10 column volumes (CV) of lysis buffer. The resin was subsequently washed with 5 CV of high-salt buffer (i.e., lysis buffer supplemented with 2 M NaCl) followed by a further 5 CV of lysis buffer. To elute nonspecifically bound proteins, the resin was subjected to 5 CV of lysis buffer supplemented with 5 mM imidazole. The protein was finally eluted by increasing the concentration of imidazole to 200 mM. The eluted protein was dialyzed extensively against 40 mM Tris-acetate, 1 mM EDTA, and 100 mM NaCl (pH 7.5) at 4℃. The dialyzed protein was concentrated to ~1 mL using an Amicon ultracentrifugal device (10000 Da NMWL (nominal molecular weight limit) (Millipore)).

Fast Protein Liquid Chromatography (FPLC) Analysis

The oligomeric state of cPYP-zENG and cPYP-zENG-M173A was determined by gel filtration using FPLC: GL column (GE Healthcare), Superdex 75 (10/300); buffer, 1xTAE, 100 mM NaCl, pH 7.5; flow-rate, 0.45 ml/min; concentration ≥ 300 μM; room temperature; detection at 280 and 446 nm. Protein samples were applied to the gel filtration column after Ni^{2+}-NTA chromatography. Both cPYP-zENG and cPYP-zENG-M173A eluted as two peaks from the gel filtration column with the majority of the material in the second (longer retention time) peak. This peak corresponds to the monomeric species based on comparison with standards and with cPYP.[16] Fractions corresponding to monomer were concentrated to ~ 100 μM under dark conditions and the sample was divided into two parts. One fraction was blue-light irradiated for approx. 45 minutes followed by 2 h of dark adaptation and the other fraction was dark-adapted at room temperature for up to 10 days. On re-chromatography each of these fractions remained >95% monomeric.

UV–Vis absorbance spectroscopy was used to determine which eluted fractions had the highest ratios of absorbance at 446 nm to that at 278 nm. A value of ~1.8 was typical for cPYP-zENG and 1.1 for cPYP-zENG-M173A holoprotein. The purity and identity of samples were confirmed using 12.5% sodium dodecyl sulfate–polyacrylamide gel electrophoresis and electrospray ionization mass spectrometry (ESI-MS). Each 1 L culture routinely gave >7-10 mg of purified protein.

UV/Vis spectra and photoisomerization

UV-Vis spectra were obtained using a PE Lambda 35 spectrophotometer or using a diode array UV–Vis spectrophotometer (Ocean Optics Inc., USB4000), in each case coupled to a temperature controlled cuvette holder (Quantum Northwest, Inc.). Protein concentrations were determined using an extinction coefficient at λ_{max} (~446 nm) of 45 x 10^3 $M^{-1}cm^{-1}$ for cPYP-zENG and 29 x 10^3 $M^{-1}cm^{-1}$ for the M173A mutant. Irradiation of the protein sample was carried out using a Luxeon III Star LED Royal Blue (455 nm) Lambertian operating at approximately 700 mA (~50 mW/cm^2). Changes in the absorbance spectrum at 446 nm were monitored to determine the rate constants for thermal relaxation. Data were fit to single-exponential functions to extract rate constants. All fitting was performed using IgorPro (Wavemetrics). For denaturation experiments, protein samples were prepared in 1× TAE, 100 mM NaCl buffer (pH 7.5) containing 0, 0.3, 0.45, 0.60, 0.75, 0.90, 1.05, 1,1.15, 1.3, 1.45, 1.6, 1.75, 1.90, 2.05, 2.2, 2.35, 2.5, 2.65 and 2.8 M GdnHCl (ThermoFisher). Samples were fully dark adapted before they were scanned in the UV-Vis region from 250 nm to 550 nm. Before each scan the samples were allowed to equilibrate at 20°C for 2 min before spectra were collected. The contribution from the buffer was subtracted from the protein spectra. Absorbance data at 446 nm vs. [Gdn] (Fig. 3c) were fit using the equation:

$$A_{446} = \frac{(\alpha_N + \beta_N[Gdn] + \left\{(\alpha_D + \beta_D[Gdn])\exp\left[\frac{m_{1_{D-N}}([Gdn]-[D_1]^{50\%})}{RT}\right]\right\})}{\left\{1+\exp\left[\frac{m_{1_{D-N}}([Gdn]-[D_1]^{50\%})}{RT}\right]\right\}}$$

where α_N is A_{446} in the absence of guanidinium, β_N is the slope (A_{446} vs $[Gdn]$) at the beginning of the curve, α_D is the value of A_{446} for the fully denatured state, and β_D is the slope at the end of the transition. The quantity $[D_1]^{50\%}$ is the Gdn concentration at the point where the protein is 50% unfolded, and the term $m_{1_{D-N}}$ is a constant that reflect the degree of unfolding.

Fluorescence monitored denaturation studies

Guanidinium-induced denaturation of cPYP-zENG and cPYP-zENG-M173A was monitored by measuring fluorescence using an Aviv Associates (Lakewood, NJ) model ATF 105 automated titrating spectrofluorometer. Protein samples (~2.5 μM) were prepared in 1× TAE, 100 mM NaCl buffer (pH 7.5) with or without 5.5 M GdnHCl and allowed to incubate overnight to fully dark adapt. Both folding and unfolding titrations were performed. The fluorescence of protein samples was acquired using excitation at 280 nm and emission at 360 nm with bandwidth of 2 nm. During the fluorescence experiment, the sample volume was kept constant at 2 mL, with 1.0 cm path length quartz cuvette, and the temperature was maintained at 20°C. The sample was stirred for 2 min after each addition of denaturant, and then stirring was halted before the fluorescence measurement. Fluorescence data were fit using either the equation shown above for a single transition or, for a two state transition, the equation:

$$F_{360} = \frac{(\alpha_N + \beta_N[Gdn] + \left\{(\alpha_D + \beta_D[Gdn])\exp\left[\frac{m_{1_{D-N}}([Gdn]-[D_1]^{50\%})}{RT}\right]\right\})}{\left\{1+\exp\left[\frac{m_{1_{D-N}}([Gdn]-[D_1]^{50\%})}{RT}\right]\right\}} + \frac{\left\{(\alpha_I + \beta_I[Gdn])\exp\left[\frac{m_{2_{I-N}}([Gdn]-[D_2]^{50\%})}{RT}\right]\right\}}{\left\{1+\exp\left[\frac{m_{2_{I-N}}([Gdn]-[D_2]^{50\%})}{RT}\right]\right\}}$$

Photochemical & Photobiological Sciences Accepted Manuscript

This journal is © The Royal Society of Chemistry 2015

Please do not adjust margins

where α_N is F_{360} in the absence of guanidinium, β_N is the slope (F_{360} vs $[Gdn]$) at the beginning of the curve, α_D is the value of F_{360} after the first transition, β_D is the slope at the end of the first transition, α_I is the value of F_{360} for the fully denatured state, and β_I is the slope for the fully denatured state. The quantities $[D_1]^{50\%}$ and $[D_2]^{50\%}$ are the Gdn concentrations at the midpoints of the first and second transitions, and the terms $m_{1_{D-N}}$ and $m_{2_{I-N}}$ are constants that reflect the degree of unfolding.

NMR

Labeled protein samples (^{15}N-cPYP-zENG and ^{15}N-cPYP-zENG-M173A) were prepared as described previously.[16] Solution NMR experiments were recorded at CSICOMP (Dept. of Chemistry, University of Toronto) on an Agilent DD2 700 MHz spectrometer equipped with an HFCN cold probe. Spectra were acquired with a ^{15}N NH HSQC Watergate pulse sequence from the Varian "Biopack" library. All HSQC spectra were acquired at 20°C with 2048 points spanning 8802.82 Hz in the ^{1}H dimension and 256 increments spanning 2198.07 Hz in the ^{15}N dimension. All studies were performed in 1× TAE, 100 mM NaCl buffer (pH 7.5). Sample concentrations varied from 200 to 400 μM, and the number of transients collected varied from 4 to 16 depending on the concentration available. Samples were fully dark adapted (10h) before dark-state HSQC spectra were acquired. For light state spectra, samples were exposed to blue light (Luxeon V Star LED Royal Blue (455 nm) Lambertian operating at approximately 700 mA (~80 mW/cm^2)) delivered through a fibre optic cable (Thorlabs). Spectra were processed using the NMRPipe processing suite. Typically, FID signals were zero filled to double the original data size and apodized using a squared-cosine window function prior to Fourier transformation. The NMRViewJ (One Moon Scientific) suite was employed for analyzing HSQC spectra.

Electrophoretic mobility shift assays (EMSA) for DNA binding

Samples of cPYP-zENG-M173A (5 μL each) at various dilutions (1 μM, 2 μM, 4 μM, 6 μM, 8 μM, 10 μM, 20 μM,40 μM, 60 μM, 100 μM) were incubated with 45 μL of EMSA cocktail for 1 hour at 25°C. The cocktail contained 25 μL of 2x EMSA buffer (1× EMSA buffer contains 20 mM Tris (pH 7.5), 40 mM KCl, 3 mM $MgCl_2$, 0.1% Triton X-100, 5% glycerol, and 1 mM DTT), 1 μL of 5 mg/mL BSA, 2.5 μL of 2 mg/mL of sheared salmon testes DNA, 11.5 μL H_2O, and 5 μL of 50 nM target DNA. The target DNA was the Cy5-labelled oligo: 5'-CGCAGTGTAATCCCC TCGAC-3', annealed to its complement. The first lane contained a negative control, where 5 μL of dilution buffer (20 mM Tris (pH 7.5), 100 μg/mL BSA, and 0.1% Triton X-100) was used instead of 5 μL DNA-binding protein. After the indicated incubation time, 20 μL of the sample was run on an 8% polyacrylamide gel containing 0.5× TBE (1× TBE contains 90 mM boric acid, 2 mM EDTA and 100 mM Tris, pH 8.3). The running buffer was 0.5× TBE. Gels were run at 25°C for 105 min at 300 V using an Emperor Penguin Water cooled dual-gel electrophoresis system. To assay under blue light irradiation, the sample preparation, incubation, and gel run were performed under cycles of 450 nm light illumination composed of 3 minutes of irradiation followed by 2 minutes in the dark. An array of Luxeon III Star LED Royal Blue LEDS (455 nm), each operating at approximately 50 mW/cm^2 (700 mA), was used as the light source. For the dark-adapted trials all the steps are conducted in the presence of red light. After the run time, the gels were scanned with a red-light laser on Pharos FX® plus molecular imager (Bio-Rad) and the images were recorded using QuantityOne software (Bio-Rad). The extent of DNA binding was quantified by analysis with Image Lab® software (Bio-Rad). Using Igor Pro software, three sets of data were averaged and fit to the Hill equation to determine the apparent K_d.

Protein models

The three dimensional structure of dark adapted wild type PYP (Protein Data Bank (PDB) entry 1NWZ) and re-engineered homeodomain (HD)-DNA (Protein Data Bank (PDB) entry 2HOT) obtained by high resolution X-ray crystallography, and light adapted wild type PYP obtained by solution NMR in combination with SAXS, DEER (Protein Data Bank (PDB) entry 2KX6) were used to create models of cPYP-zENG-M173A. The Accelyrs DSVisualizer suite was employed.

Acknowledgements

We are grateful to the Natural Science and Engineering Research Council of Canada for funding. NMR instrumentation at the Centre for Spectroscopic Investigation of Complex Organic Molecules and Polymers was supported by the Canada Foundation for Innovation (Project number: 19119) and the Ontario Research Fund.

Notes and references

‡ We use the term "light state" here to refer to the relatively long-lived putative signalling state of PYP. This state has also been called the I(2)' state, the PYP_M, and the pB state.

1. C. Desplan, J. Theis and P. H. O'Farrell, *Nature*, 1985, **318**, 630-635.
2. S. DiNardo, E. Sher, J. Heemskerk-Jongens, J. A. Kassis and P. H. O'Farrell, *Nature*, 1988, **332**, 604-609.
3. K. Hatta, T. F. Schilling, R. A. BreMiller and C. B. Kimmel, *Science*, 1990, **250**, 802-805.
4. R. Morgan, *FEBS Lett.*, 2006, **580**, 2531-2533.
5. S. B. Carroll, S. DiNardo, P. H. O'Farrell, R. A. White and M. P. Scott, *Genes Dev.*, 1988, **2**, 350-360.
6. S. A. Morgan, S. Al-Abdul-Wahid and G. A. Woolley, *J. Mol. Biol.*, 2010, **399**, 94-112.
7. S. A. Morgan and G. A. Woolley, *Photochem. Photobiol. Sci.*, 2010, **9**, 1320-1326.
8. R. Nakamura, N. Hamada, H. Ichida, F. Tokunaga and Y. Kanematsu, *Photochem. Photobiol.*, 2008, **84**, 937-940.

Photochemical & Photobiological Sciences Accepted Manuscript

9. S. F. El-Mashtoly, S. Yamauchi, M. Kumauchi, N. Hamada, F. Tokunaga and M. Unno, *J. Phys. Chem. B*, 2005, **109**, 23666-23673.
10. A. F. Philip, M. Kumauchi and W. D. Hoff, *Proc. Natl. Acad. Sci. U.S.A*, 2010, **107**, 17986-17991.
11. S. Ohishi, N. Shimizu, K. Mihara, Y. Imamoto and M. Kataoka, *Biochemistry*, 2001, **40**, 2854-2859.
12. K. Shirai, Y. Yamazaki, H. Kamikubo, Y. Imamoto and M. Kataoka, *Proteins*, 2007, **67**, 821-833.
13. C. Bernard, K. Houben, N. M. Derix, D. Marks, M. A. van der Horst, K. J. Hellingwerf, R. Boelens, R. Kaptein and N. A. van Nuland, *Structure*, 2005, **13**, 953-962.
14. B. C. Lee, P. A. Croonquist, T. R. Sosnick and W. D. Hoff, *J. Biol. Chem.*, 2001, **276**, 20821-20823.
15. P. L. Ramachandran, J. E. Lovett, P. J. Carl, M. Cammarata, J. H. Lee, Y. O. Jung, H. Ihee, C. R. Timmel and J. J. van Thor, *J. Am. Chem. Soc.*, 2011, **133**, 9395-9404.
16. A. Kumar, D. C. Burns, M. S. Al-Abdul-Wahid and G. A. Woolley, *Biochemistry*, 2013, **52**, 3320-3331.
17. N. D. Clarke, C. R. Kissinger, J. Desjarlais, G. L. Gilliland and C. O. Pabo, *Protein Sci.*, 1994, **3**, 1779-1787.
18. E. D. Getzoff, K. N. Gutwin and U. K. Genick, *Nat. Struct. Biol.*, 2003, **10**, 663-668.
19. J. M. Reis, D. C. Burns and G. A. Woolley, *Biochemistry*, 2014, **53**, 5008-5016.
20. J. Sasaki, M. Kumauchi, N. Hamada, T. Oka and F. Tokunaga, *Biochemistry*, 2002, **41**, 1915-1922.
21. Y. Imamoto, M. Harigai, T. Morimoto and M. Kataoka, *Photochem. Photobiol.*, 2008, **84**, 970-976.
22. T. E. Meyer, S. Devanathan, T. Woo, E. D. Getzoff, G. Tollin and M. A. Cusanovich, *Biochemistry*, 2003, **42**, 3319-3325.
23. C. P. Joshi, H. Otto, D. Hoersch, T. E. Meyer, M. A. Cusanovich and M. P. Heyn, *Biochemistry*, 2009, **48**, 9980-9993.
24. A. Kumar and G. A. Woolley, *Photochem. Photobiol.*, 2015, May 6. doi: 10.1111/php.12464.
25. C. J. Craven, N. M. Derix, J. Hendriks, R. Boelens, K. J. Hellingwerf and R. Kaptein, *Biochemistry*, 2000, **39**, 14392-14399.
26. G. Rubinstenn, G. W. Vuister, F. A. Mulder, P. E. Dux, R. Boelens, K. J. Hellingwerf and R. Kaptein, *Nat. Struct. Biol.*, 1998, **5**, 568-570.
27. S. E. Ades and R. T. Sauer, *Biochemistry*, 1994, **33**, 9187-9194.
28. M. Mazumder, K. Brechun, Y. Kim, S. Hoffmann, Y.-Y. Chen, C.-L. Keiski, K. M. Arndt, David; and G. A. Woolley, *Protein Eng. Des. Sel.*, 2015, in press.
29. P. Changenet-Barret, A. Espagne, N. Katsonis, S. Charier, J. B. Baudin, L. Jullien, P. Plaza and M. M. Martin, *Chem. Phys. Lett.*, 2002, **365**, 285-291.

Photochemical & Photobiological Sciences Accepted Manuscript

This journal is © The Royal Society of Chemistry 2015

Please do not adjust margins