Supplementary information's to **'Allosteric communication between DNA-binding and light-responsive domains of diatom class I aureochromes'** -by Banerjee et al.

**Ankan Banerjee[1,†], Elena Herman[2,†], Manuel Serif[3], Manuel Maestre Reyna[4], Sebastian Hepp[1], Richard Pokorny[5], Peter G. Kroth[3], Lars-Oliver Essen[1*], Tilman Kottke[2].**

**SUPPLEMENTARY FIGURE LEGENDS**

**Figure S1. *Pt*AUREO forms homo- and heterodimers. (A)** *Pt*AUREO1a and *Pt*AUREO1c interact with 1:1 stoichiometry. About five-fold higher concentration of the labeled protein (*Pt*AUREO1c) to the estimated $K_D$ ~0.24 ± 0.09 µM was used to determine the stoichiometry. The fraction bound to the molar ratio of *Pt*AUREO1a and *Pt*AUREO1c was plotted and fitted with the hill equation (red line) provided in Origin 8.0. The stoichiometry (n) was determined calculating the intercept between saturation and binding. (**B**) SDS-PAGE of the purified aureochrome variants are shown in the inset. **(C)** Size exclusion chromatography of full-length aureochrome1a at injected concentrations of 0.1 (top) and 1.0 mM (bottom) in the dark (black lines) and under continuous illumination (red lines). Protein with bound chromophore was detected at 448 nm (thin line) and 390 nm (thick line). At an injected concentration of 0.1 mM, *Pt*AUREO1a eluted at 12.8 mL (170 kDa). At a ten times higher concentration, the elution peak shifted to 12.2-12.3 mL (230-240 kDa). The oligomeric state of full-length aureochrome is not dependent on the light conditions, but dependent on the concentration. With a theoretical molecular mass of the monomer of 43.7 kDa, these species are assigned to higher oligomers than a dimer, in which the N-terminal extension serves as an oligomerization site. The void volume was 8.1 mL and is marked by an arrow (red). The elution volumes of protein standards used for calibration are further marked by arrows (black). From left to right: 1, ferritin (440 kDa); 2, aldolase (158 kDa); 3, conalbumin (75 kDa); 4, ovalbumin (44kDa); 5, ribonuclease A (13.7 kDa). **(D)** Helical wheel interaction model to predict homo- and heterodimerization ability of *Pt*AUREOs, generated by DrawCoil. Blue dashed lines show stabilizing positive interactions via salt bridges of positively and negatively charged amino acid side chains in close proximity in the coiled coil, whereas red lines (not present) would show destabilizing negative interactions.

**Figure S2. UV/Vis and FT-IR spectroscopy of aureochrome variants** (**A**) UV/Vis spectrum of the full length (red line) and bZIP-A´α-LOV-Jα (black line) of *Pt*AUREO1a in the dark. The absorbance spectra show the typical fine structure of protein-bound oxidized flavin with an absorption maximum at 448 nm. (**B**) *Top*. Light-*minus*-dark FT-IR difference spectra of bZIP-A´α-LOV-Jα in $D_2O$ (red line) and $H_2O$ (black line) as well as of A´α-LOV-Jα in $D_2O$ (blue line) taken from (1). The efficient exchange of hydrogen to deuterium in bZIP-A´α-LOV-Jα is demonstrated by the shift of $C_4O$ bands of flavin from 1727, 1713, and 1693 $cm^{-1}$ to 1719, 1704, and 1684 $cm^{-1}$, respectively, which resemble those in A´α-LOV-Jα in $D_2O$. *Bottom*. Segment-resolved double difference spectra obtained by subtraction of the difference spectrum of A´α-

LOV-Jα from that of bZIP-A´α-LOV-Jα in $D_2O$ (red line) and $H_2O$ (black line), respectively. Bands at 1666 and 1649 $cm^{-1}$ in $D_2O$ are assigned to amide I bands of corresponding amide I bands in $H_2O$ at 1671 and 1653 $cm^{-1}$ (highlighted in gray). In this spectral region, such weak shifts are typical for amide I modes. The strong difference band at 1609 (-) /1578 (+) $cm^{-1}$ is assigned to the asymmetric stretch of aspartate or glutamate in $D_2O$, supporting the assignment of the 1387 (-) / 1404 (+) $cm^{-1}$ signal to the corresponding symmetric stretch. The double difference spectrum of bZIP-A´α-LOV-Jα–*minus*–bZIP-A´α-LOV-Jα in $D_2O$ (gray line) illustrates the sample-to-sample variation of independent preparations. (**C**) Light-*minus*-dark FT-IR difference spectrum of AUREO1aΔJα (red line) in comparison to that of bZIP-A´α-LOV-Jα (black line) and to that of A´α-LOV (gray line) taken from ref. (1). Dimerization of the LOV domains is significantly suppressed in AUREO1aΔJα, which is assigned by the strongly decreased absorbance in the band at 1627 $cm^{-1}$ in comparison to the marker band at 1630 $cm^{-1}$ in bZIP-A´α-LOV-Jα (highlighted in gray). The comparison to the difference spectrum of the monomeric A´α-LOV demonstrates the absence of the formation of a lit state dimer in AUREO1aΔJα.

**Figure S3. The bZIP domain of *Pt*AUREO1a bound to ACGT sequence independent of light. (A)** DNA sequences used in the EMSA assays. Aureo box and G-box hold ACGT motif in the promoter core. **(B)** bZIP-A´α-LOV-Jα is specific to ACGT box. bZIP domain of *Pt*AUREO1a does not bind a non-related promoter motif, the STAT box. **(C)** Aureochrome binds specifically to the TGACGT motif and forms the super shift complex independent of DNA length (31 bp ~3turn; 20 bp ~2turn; 14 bp~ 1.5 turn DNA). **(D)** Competitive EMSA of 0.5 µM bZIP-A´α-LOV-Jα bound to 2nM labeled aureo box and an increasing concentration of either WT aureo box (i), or C->T mutated aureo box (aureomut) (ii), or G-box (iii)unlabeled oligonucleotide. Minus (-) indicates oligonucleotide without protein, plus (+) indicates 0.5 µM bZIP-A´α-LOV-Jα with 2 nM aureo box. DNA formed both LS, lower shift and SS, super shift. The super shift complex is easily accessible to unlabeled oligonucleotide, whereas the lower shift (1:1 complex) is more stable. **(E)** Competitive EMSA using labeled G-box and increasing concentration of WT aureo box (i) and mutated aureo box (ii). **(F)** bZIP-A´α-LOV-Jα does not bind to single-stranded aureo box containing 32 bp long oligonucleotide. **(G)** Increment of helicity was monitored by CD spectra of DNA bound and unbound bZIP-A´α-LOV-Jα in 0.5X TBE buffer. **(H)** Isolation of aureo box-bound bZIP-A´α-LOV-Jα in 0.5X TBE buffer. The aureo box eluted at 12.2 and 13.2 mL. bZIP-A´α-LOV-Jα eluted at 12.5mL and the complex of dsDNA-bZIP-A´α-LOV-Jα eluted at 11mL.

**Figure S4. Effect of DNA on recovery kinetics and FT-IR spectra of *Pt*AUREO1a and bZIP-A´α-LOV-Jα.** (**A**) Recovery kinetics of bZIP-A´α-LOV-Jα without (black line) and with equimolar concentration of aureo box (red line) at 100 mM NaCl. A moderate influence of the aureo box on the recovery kinetics is observed. (**B**) Representative FT-IR absorption spectrum of bZIP-A´α-LOV-Jα in the presence of dsDNA (aureo box). Bands assigned to the dsDNA are indicated. (**C**) Light-induced difference spectra of full-length *Pt*AUREO1a (red) and bZIP-A´α-LOV-Jα (black) in the presence of aureo box. The similarity points to a minor influence of the N-terminal extension on the light response of *Pt*AUREO1a. (**D**) Difference spectra of bZIP-A´α-LOV-Jα in the presence of aureo box (red), G-box (blue), STAT box (green) and without DNA (black). The selected spectral region contains mostly contributions from the flavin chromophore and shows very little influence by the bound DNA. (**E**) Double difference spectra generated from the difference spectra of bZIP-A´α-LOV-Jα *minus* A´α-LOV-Jα in the presence of aureo box (red), G-box (blue), and STAT box (green) in comparison to without DNA. The structural response of the bZIP region to light in the presence of dsDNA is dependent on the DNA sequence and different from that without DNA. The signals with and without DNA cannot be matched by an inversion pointing to differences in the origin of the structural responses.

**Figure S5. SAXS data and MD simulations of bZIP-A´α-LOV-Jα. (A)** Small angle X-ray scattering of bZIP-A´α-LOV-Jα was measured in Tris-HCl buffer (see materials and methods) at different concentrations. The datasets were merged taking into account that lower concentrations correspond to small angles and higher concentrations to wider angles. The measurement performed in lit (BL, blue) and dark (DR, red) states revealed no or very little conformational changes of the dimeric envelope (see also table S2). The P(r) distribution profiles and the radius of gyrations are shown in the upper panel. **(B)** Kratky plot of dark and lit states indicate no difference in the flexibility components. The *ab initio* envelops of DR and BL states are shown in the panel above. **(C)** Average structures of bZIP-A´α-LOV-Jα initial (start), and major conformation (max 1 and max 2) resulting from a 145 ns unrestrained MD simulation. The conformational space was explored as a function of the DNA-to-bZIP tilt (angle between the blue vectors), against bZIP twist (angle between the yellow vectors). **(D)** 2D frequency analysis of bZIP twist vs. DNA-to-bZIP tilt. Higher occupancy values are shown in red, lower values in blue. Relative positions of Max 1, Max 2 and Start (Fig. S5A) are highlighted. **(E)** Time-

dependent trajectory for the Self-guided Langevin Dynamics SAXS-restraint energy. After ~550 ps, the restraint energy has stabilized, indicating simulation convergence.

## SUPPLEMENTARY FIGURES

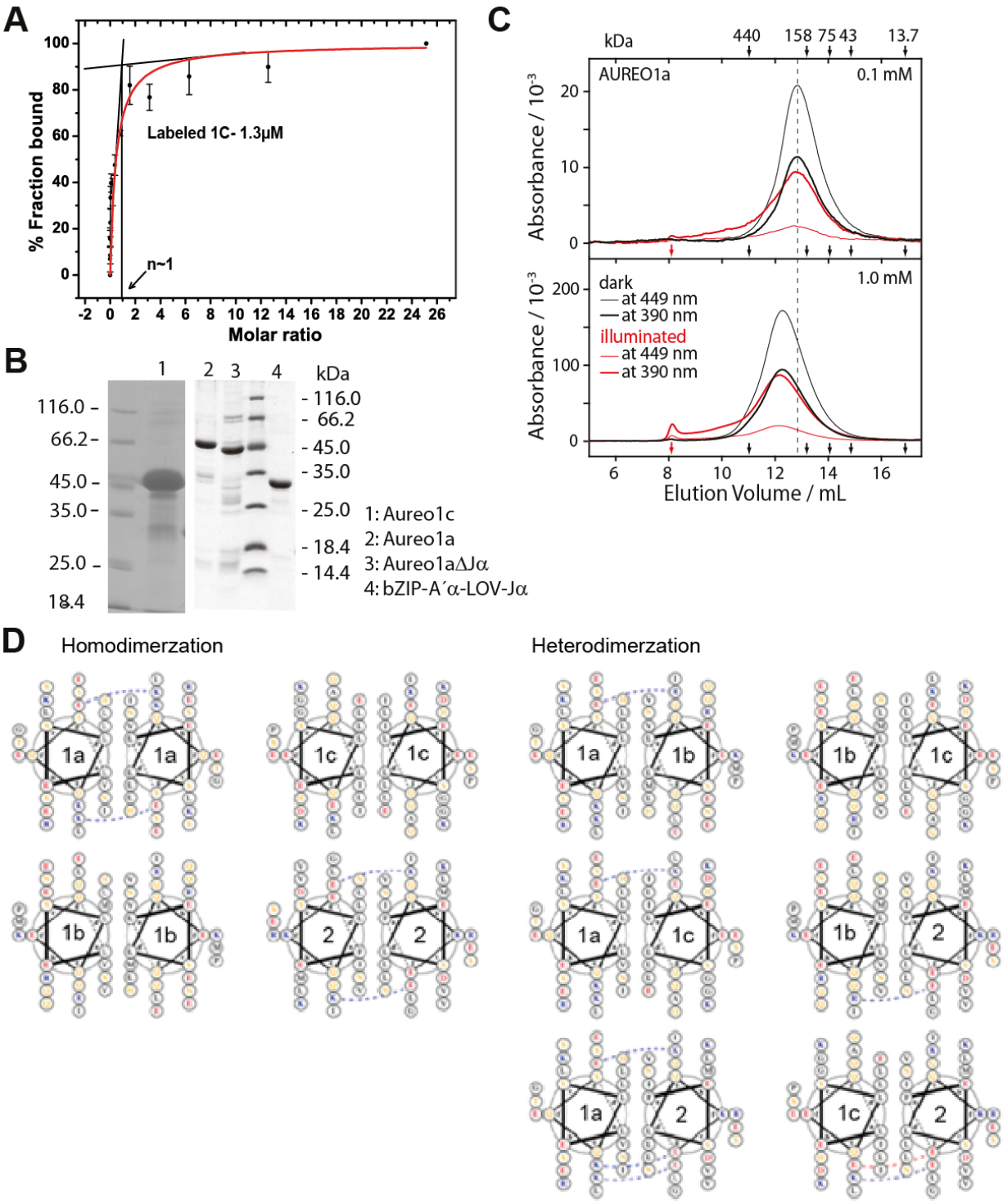


**Figure S1**

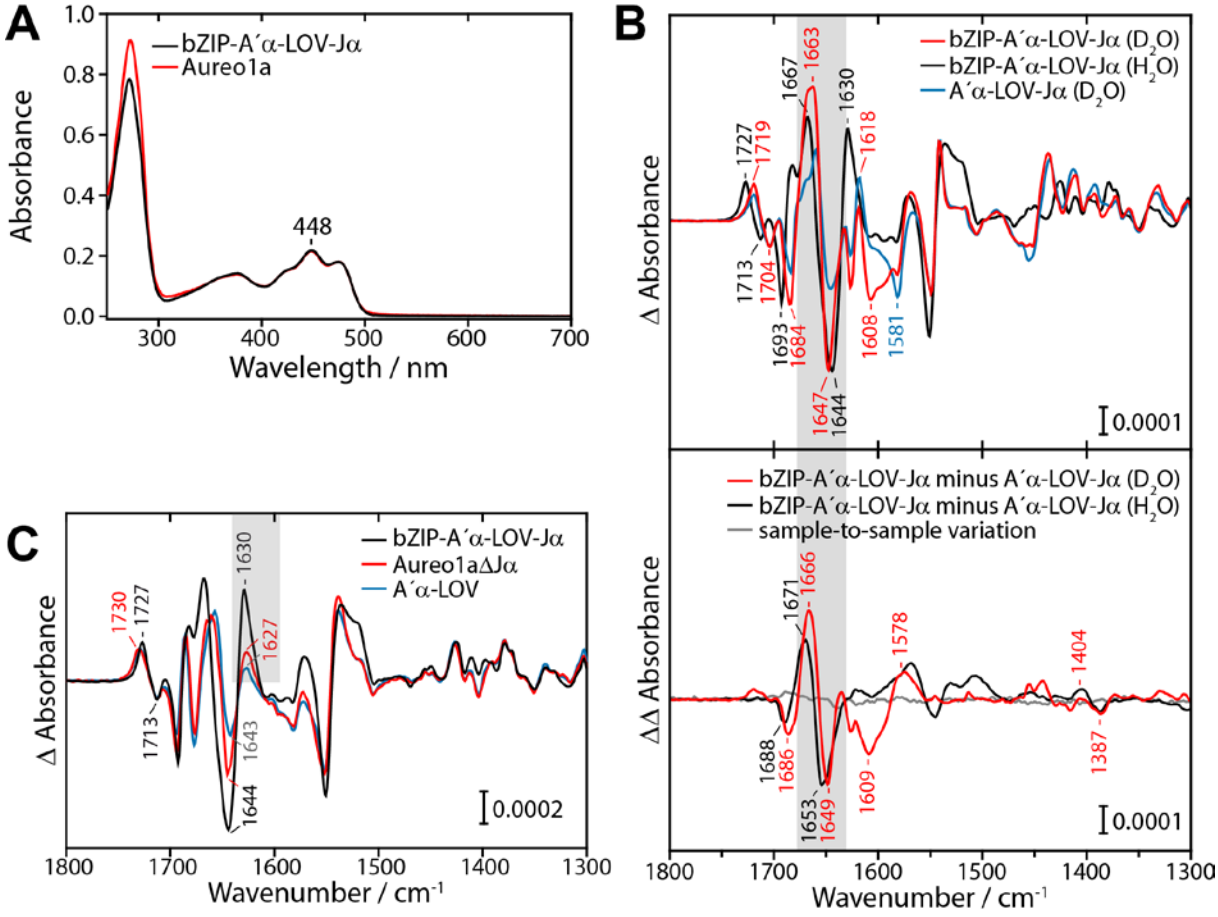
A
bZIP-A´α-LOV-Jα
Aureo1a
448
Absorbance
Wavelength / nm
B
bZIP-A´α-LOV-Jα (D2O)
bZIP-A´α-LOV-Jα (H2O)
A´α-LOV-Jα (D2O)
Δ Absorbance
0.0001
bZIP-A´α-LOV-Jα minus A´α-LOV-Jα (D2O)
bZIP-A´α-LOV-Jα minus A´α-LOV-Jα (H2O)
sample-to-sample variation
ΔΔ Absorbance
0.0001
Wavenumber / cm-1
C
bZIP-A´α-LOV-Jα
Aureo1aΔJα
A´α-LOV
Δ Absorbance
0.0002
Wavenumber / cm-1

**Figure S2**

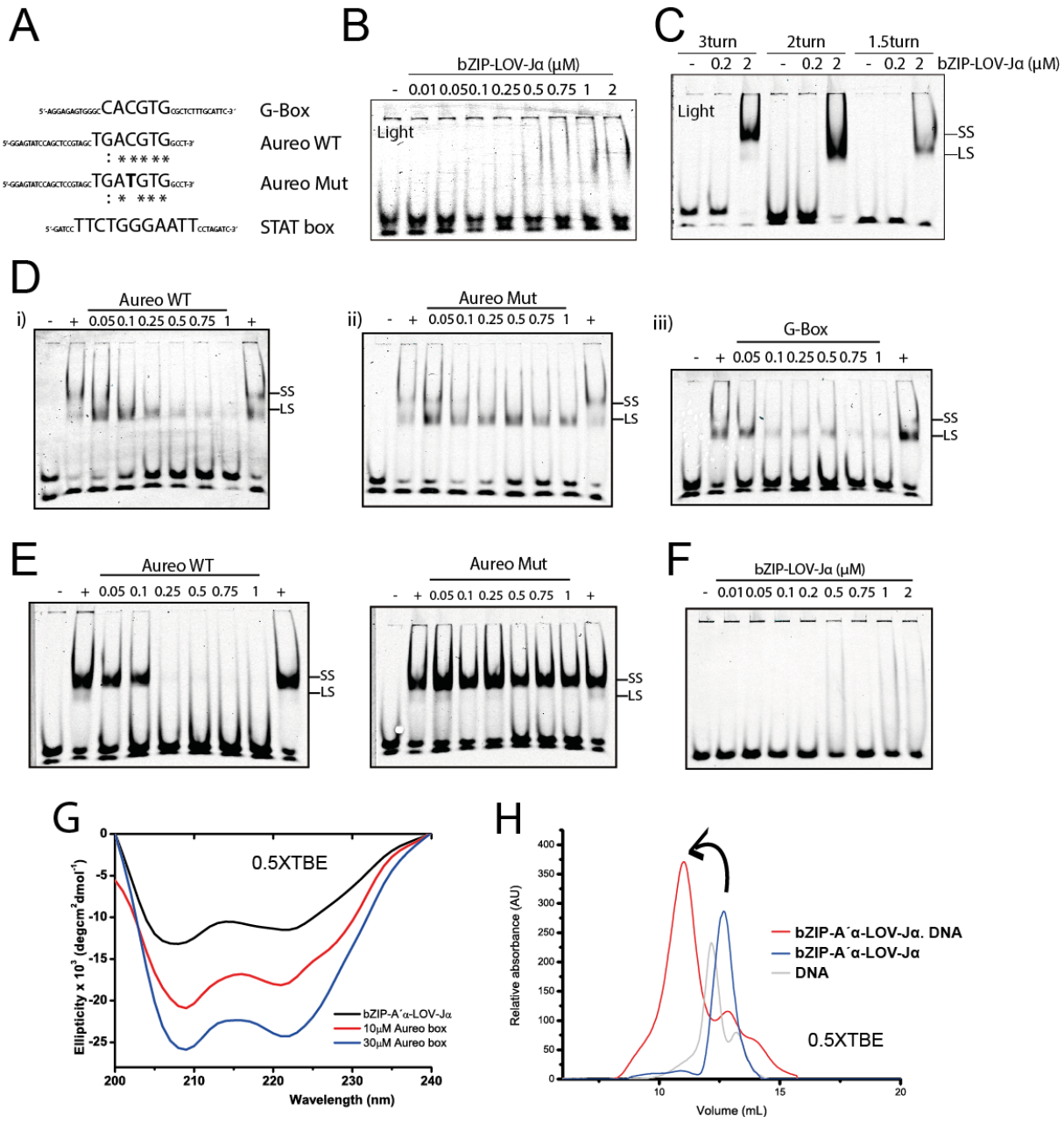
A
G-Box
Aureo WT
Aureo Mut
STAT box
B
bZIP-LOV-Jα (μM)
Light
C
3turn
2turn
1.5turn
bZIP-LOV-Jα (μM)
Light
SS
LS
D
Aureo WT
Aureo Mut
G-Box
E
F
bZIP-LOV-Jα (μM)
G
0.5XTBE
Ellipticity x 10³ (degcm²dmol⁻¹)
Wavelength (nm)
bZIP-A´α-LOV-Jα
10μM Aureo box
30μM Aureo box
H
Relative absorbance (AU)
Volume (mL)
bZIP-A´α-LOV-Jα. DNA
bZIP-A´α-LOV-Jα
DNA
0.5XTBE

**Figure S3**

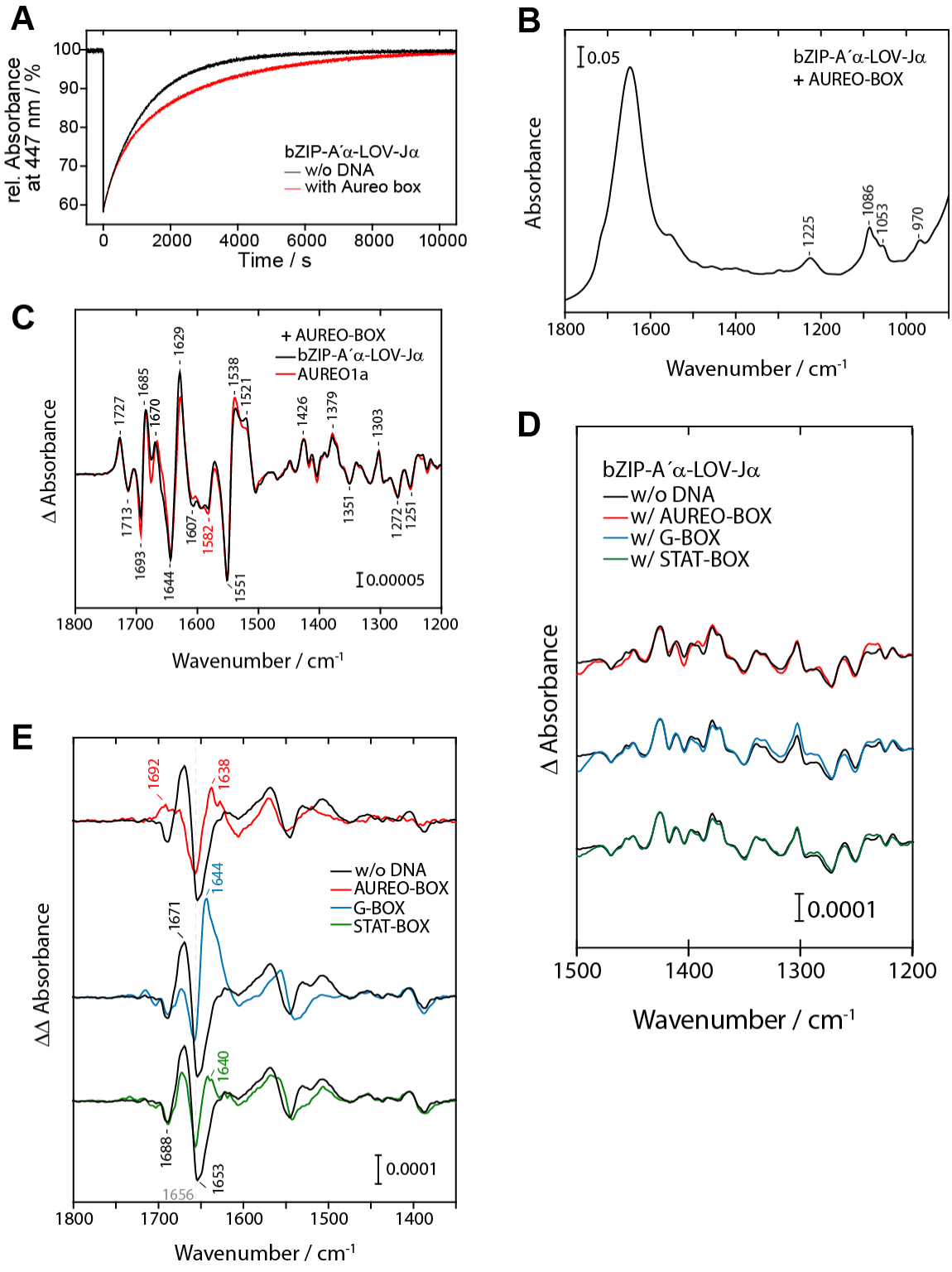
A
rel. Absorbance at 447 nm / %
Time / s
bZIP-A´α-LOV-Jα
w/o DNA
with Aureo box
B
0.05
bZIP-A´α-LOV-Jα
+ AUREO-BOX
Absorbance
1225
1086
1053
970
Wavenumber / cm-1
C
+ AUREO-BOX
bZIP-A´α-LOV-Jα
AUREO1a
Δ Absorbance
1727
1685
1670
1629
1538
1521
1426
1379
1303
1713
1693
1644
1607
1582
1551
1351
1272
1251
0.00005
Wavenumber / cm-1
D
bZIP-A´α-LOV-Jα
w/o DNA
w/ AUREO-BOX
w/ G-BOX
w/ STAT-BOX
Δ Absorbance
0.0001
Wavenumber / cm-1
E
ΔΔ Absorbance
1692
1638
1644
1671
1640
1688
1653
1656
w/o DNA
AUREO-BOX
G-BOX
STAT-BOX
0.0001
Wavenumber / cm-1

**Figure S4**

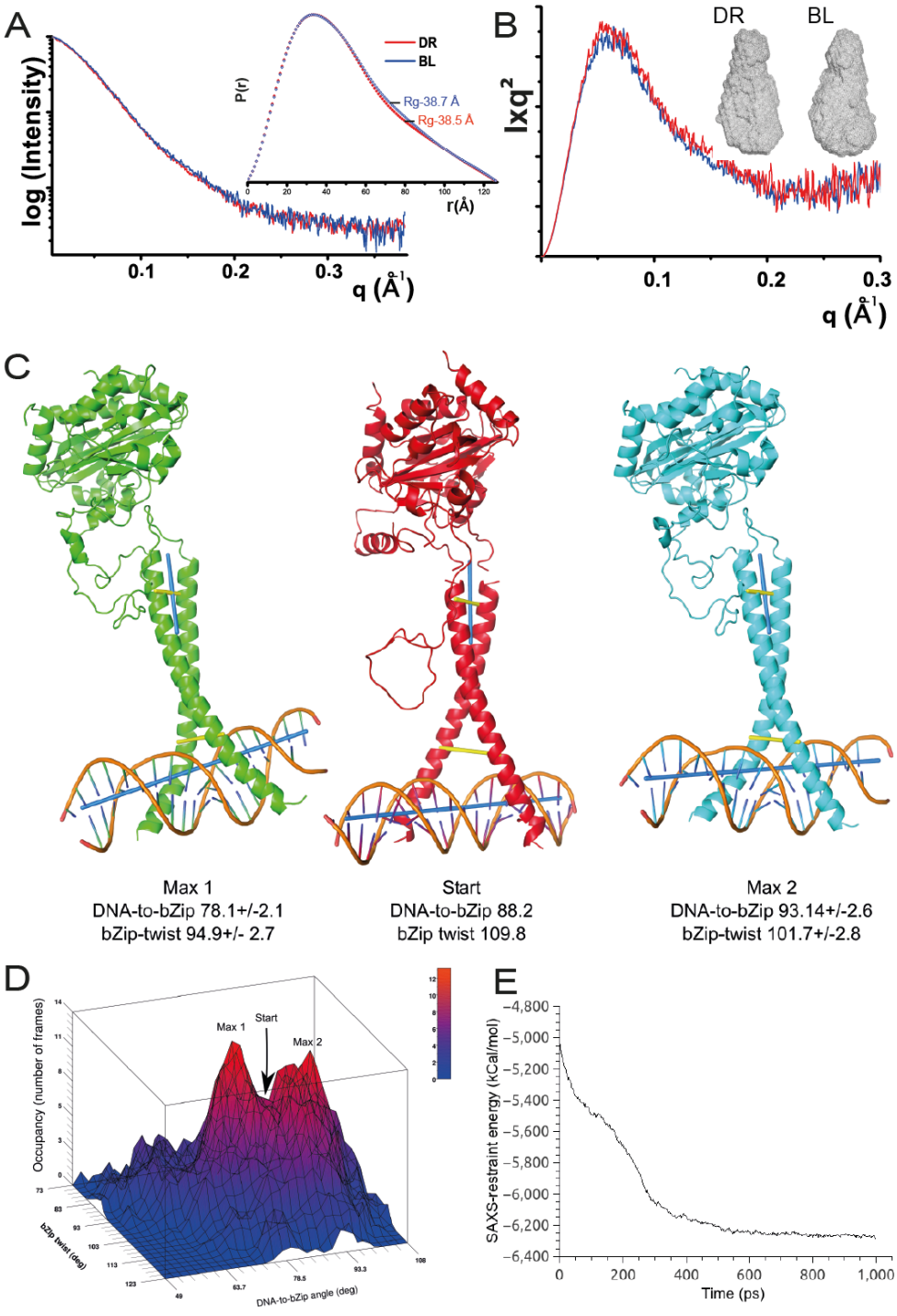
A
log (Intensity)
q (Å-1)
DR
BL
P(r)
Rg-38.7 Å
Rg-38.5 Å
r(Å)
B
DR
BL
Ixq2
q (Å-1)
C
Max 1
DNA-to-bZip 78.1+/-2.1
bZip-twist 94.9+/- 2.7
Start
DNA-to-bZip 88.2
bZip twist 109.8
Max 2
DNA-to-bZip 93.14+/-2.6
bZip-twist 101.7+/-2.8
D
Occupancy (number of frames)
Max 1
Start
Max 2
bZip twist (deg)
DNA-to-bZip angle (deg)
E
SAXS-restraint energy (kCal/mol)
Time (ps)

**Figure S5**

## SUPPLEMENTARY TABLES

**Table S1.** Changes of secondary structure contents upon dsDNA binding to bZIP-A´α-LOV-Jα, quantified by circular dichroism spectroscopy.

| | | α-Helix (%) |
|---|---|---|
| **Tris-HCl** | bZIP-A´α-LOV-Jα (38.5 µM dimer) | 21.3 |
| | bZIP-A´α-LOV-Jα+10 µM aureo box | 22.3 |
| | bZIP-A´α-LOV-Jα+20 µM aureo box | 23.5 |
| | bZIP-A´α-LOV-Jα+30 µM aureo box | 26.2 |
| | bZIP-A´α-LOV-Jα+40 µM aureo box | 25.5 |
| | bZIP-A´α-LOV-Jα+50 µM aureo box | 27.5 |
| **TBE** | bZIP-A´α-LOV-Jα (38.5 µM dimer) | 15.3 |
| | bZIP-A´α-LOV-Jα+10 µM aureo box | 18.3 |
| | bZIP-A´α-LOV-Jα+30 µM aureo box | 23.0 |

**Table S2.** Concentration dependence in SAXS experiments of bZIP-A´α-LOV-Jα (lit state) in Tris-HCl versus TBE buffer. bZIP-A´α-LOV-Jα estimated molecular weight is 57.2 kDa. $R_g$, radius of gyration was calculated from the linear region defined by Guinier approximation. $I_0$ was derived by using BSA as a standard. $D_{max}$ (maximal particle diameter) and Porod volumes were taken from the P(r) distribution.

| Sample | Concentration ($mgmL^{-1}$) | Rg (Å) | $I_0$ | $D_{max}$ (Å) | Volume ($nm^3$) |
|---|---|---|---|---|---|
| **bZIP-A´α-LOV-Jα (TBE)** | 0.625 | 36.4 | 62.35 | 120.73 | 110.67 |
| | 1.25 | 34.5 | 57.29 | 126.4 | 112.94 |
| | 2.5 | 35.0 | 61.07 | 130 | 118.95 |
| **bZIP-A´α-LOV-Jα (Tris-HCl)** | 0.625 | 36.3 | 41.72 | 120.58 | 115.87 |
| | 1.25 | 43.0 | 65.81 | 150.06 | 140.54 |
| | 2.5 | 43.5 | 81.82 | 150.22 | 168.54 |

**Supplementary video1** : **Molecular dynamic simulation traces of dsDNA bound bZIP-A´α-LOV-Jα.** An overview of the post-SAXS restrained 140 ns molecular dynamics simulation. In the video, the aureochrome-DNA complex starts in its fully SAXS-restrained form (green). In the simulation, the complex relaxes within 2.2 ns into the meta-stable state (blue), corresponding to the first 22 snapshots. The meta-stable state is maintained for 61.1 ns, although several significant deviations occur within this time (grey, for example at 12.2 ns, i.e. snapshot 122), perhaps suggesting unsuccessful attempts to reach the stable state. While still primarily in the meta-stable state, at 60.7 ns the transition towards the stable state is initiated, as the bZip helices start to kink and twist around each other. At 63.7 nanoseconds, the transition is irreversible, and leads within 3.7 ns to the stable state (red), which is maintained for most of the remainder of the simulation. However, snapshot 768 (76.8 ns) presents significant deviation from the stable state, and a higher degree of similarity to the meta-stable state (Fig. 6A), perhaps suggesting that, in the long run, the meta-stable state can be transiently achieved from the more stable final state at room temperature. In order to highlight the most important regions of the simulation (i.e. the restraint relaxation and the transition to the stable state) those two parts of the video have been slowed down ten times. In all transition snapshots, which exceeded the limits of the calculated bZip-RMSD distributions shown in Fig. 6B, present the aureochrome in grey.

## SUPPLEMENTRAY METHODS

### Construction of expression plasmids

The sequences encoding *Pt*AUREO1a and *Pt*AUREO1c (Protein-ID 49116 and 56742, Joint Genome Institute (JGI) database) from *P. tricornutum* were optimized in codon usage for *E. coli* and *de novo* synthesized (Eurofins MWG). An *Nde*I restriction site was included at the 5´-end and a stop codon followed by a *Sal*I restriction site at the 3´-end. The sequences were cloned into a pET28a (+) vector (Novagen) in such a way that the nucleotides encode a $(His)_6$ afiinity tag and a thrombin-cleavage site at the N-terminus.

The vector pET28a (+)-Aureo1a was used as a template to generate the vector encoding bZIP-A´α-LOV-Jα (amino acids 145-378) or AUREO1aΔJα by amplification using the forward primer 5´- GAC CGC AAA ATG AGC GAG CAG CAG AAA G-3´ or 5´- GTG CAA AGT ATC GGA TCA ATA AGT CGA CAA GC-3´ and the 5´-end phosphorylated reverse primer 5´-p-CAT ATG GCT ACC GCG CGG CAC-3´(Invitrogen) or 5´- TTG ATC CGA TAC TTT GCA CTG CAC ACC CA-3´ respectively and Phusion polymerase (New England Biolabs). After *Dpn*I digestion, the vector fragment was treated with T4 DNA kinase and ligated using T4 DNA ligase. The resulting sequence of the vector was verified by sequencing.

### Protein expression and purification

For UV/Vis, FT-IR and SEC, heterologous expression in *E. coli* BL21 (DE3) pLysE and purification via His-bind resin was performed as described previously (Herman *et al.*, 2013), Protein samples were reconstituted with FMN by incubation at twice the concentration of FMN with respect to protein for two hours at 4°C in the dark. Afterwards, the samples were dialyzed three times against 400 volumes of 50 mM phosphate buffer, pH 8, 300 mM NaCl, and 20 % glycerol. Reconstitution led to an increase of the chromophore occupancy from ~40 % to ~70 % for *Pt*AUREO1a and bZIP-A´α-LOV-Jα and from ~20 to 30 % for *Pt*AUREO1aΔJα. Reconstitution does not alter the response of the samples to light as determined by FT-IR difference spectroscopy (Herman & Kottke, 2015).

For EMSA, MST, CD, DNA-protein complex isolation and SAXS experiments, N-terminally $His_6$ tagged aureochrome variants were expressed in *E. coli* Bl21-De3 Gold after induction with 10 uM of isopropyl β-D-thiogalactopyranoside (IPTG) for 20 hr at 18 °C. The cells were harvested using centrifugation at 5000 xg for 30min and resuspended in 50 mM Na-phosphate buffer, 100 mM NaCl, 20% glycerol pH 7.8. To process the harvested cell an aliquot of resuspended cell pellet was lysed using French press after 30 min incubation with 5 µl lysozyme (50 mg/ml) and DNAase I. The cell free lysate was collected using centrifugation at 13000 X g, 4 °C, and 30 min and filtered through a 0.45 µm filter. The cell lysate was loaded on IMAC column. The protein

eluted 45 ml 50 % Ni-NTA B buffer (50mM Na-phosphate buffer, 100 mM NaCl, 20% glycerol, 500 mM imidazole pH7.8). Then the affinity purified proteins were concentrated using Amicon cutoff filter of 10kDa. 1 ml of concentrated protein was filtered using micro filters and further purified by size exclusion using a 120 ml superdex200HR column in 10 mM Tris-HCl, 300 mM NaCl pH8 buffer. Mono-disperse peaks were separated and analyzed on SDS-PAGE. The fractions containing protein of interest were concentrated and the purity was further analyzed on a SDS-PAGE.

**Electrophoretic mobility shift assay**

IRdye700 labeled oligonucleotides (see Fig.S3) were purchased from MWG Eurofins and equimolar concentration of complimentary oligos were mixed under green light in 10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1mM EDTA buffer and annealed in dark using the protocol obtained from Sigma-Aldrich and stored at -20 °C for future use. 2 nM final concentration of IRdye700 labeled dsDNA was mixed in aliquots with increasing concentration of dark adapted protein starting from 10 nM to 2 µM in dark (under safe light $\lambda_{max}$ ~ 650 nm) in binding buffer containing 10 mM Na-phosphate pH8, 50 mM NaCl, 0.5 mM EDTA, 5 mM $MgCl_2$, 1 mM DTT, 20 µg/ml BSA, 6 % glycerol (adapted from (2)). The light state of bound aureochrome-dsDNA was achieved by using blue light pulse, $\lambda_{max}$ ~ 470nm, 100µmol $m^{-2}s^{-1}$ for 5-10 min. A 4% native PAGE was used to separate dsDNA bound complexes. The PAGE ran under complete dark on 0.5X TBE (50 mM Tris pH 8, 50 mM boric acid, 1 mM EDTA) as electrophoresis buffer. Competition experiments were performed using a fixed concentration of protein 0.5 µM, 2 nM labeled probe and a serial increment of non-labeled dsDNA concentration (50 nM-1 µM). The bound complex was visualized on IR-scanner Odyssey (LI-COR).

**Micro scale thermophoresis**

Bio-molecular interaction of binding partners was analyzed with Micro scale thermophoresis (MST) using a protocol described earlier (3). Fluorescently labeled proteins with amine-reactive Monolith NT$^{TM}$ protein labeling kit NT-647 was analyzed for infra-red light induced thermophoresis with its selective unlabeled interaction counterparts on Monolith NT.115 (NanoTemper Technologies GmbH). The binding experiment was performed in a binding buffer containing 10 mM Tris-HCl pH8, 300 mM NaCl and 0.1 % tween20. Dark state experiment was performed under safe red light ($\lambda_{max}$~ 650 nm). To determine the binding affinity of either full-length aureochrome (*Pt*AUREO1a/ *Pt*AUREO1c) and/or truncated variants, 66.7 nM labeled protein was mixed at steady state with an increasing concentration of unlabeled interaction partners. The measurements were performed at 25 °C, in either light or dark state with 20 % LED power and 80 % infrared-laser power using hydrophilic capillaries. At least three replicates

were evaluated using NTanalysis 1.4.27 (NanoTemper Technologies GmbH) and representative fitting curves were generated using the Hill equation in Origin8. To determine the stoichiometry of *Pt*AUREO1a-*Pt*AUREO1c interaction, 1.6 µM labeled *Pt*AUREO1c was used and titrated with increasing concentration of *Pt*AUREO1a.

**Size exclusion chromatography**

Aliquots of 60 µL with a protein concentration of 0.1 and 1.0 mM for AUREO1a and 0.1 and 1.3 mM for bZIP-A´α-LOV-Jα were analyzed with simultaneous detection at 449 nm and 390 nm. For equilibration and elution, 50 mM phosphate buffer, pH 8, 300 mM NaCl was used. Protein samples were centrifuged at 21400 *g* for 10 min at 4 °C before loading onto the column. Elution was performed with a flow rate of 0.4 mL/min. The apparent molecular mass of the sample was determined by calibration using the standard marker proteins ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa), and ribonuklease A (13.7 kDa). (GE Healthcare). The void volume was determined by analyzing thyroglobulin (669 kDa).

**UV/Vis spectroscopy**

UV/Vis spectra and kinetics were recorded on a Shimadzu UV-2450 spectrometer. The 455 nm LED for illumination had a 20 nm full width at half-maximum (fwhm) and an intensity of 20 mW $cm^{-2}$ at the sample (Philips Lumileds). Absorbance spectra were used for determination of the chromophore and protein concentrations. Extinction coefficients of 12500 L $mol^{-1}$ $cm^{-1}$ at 450 nm and 18500 L $mol^{-1}$ $cm^{-1}$ at 280 nm were used for free FMN. Extinction coefficients for the AUREO1a apoprotein of 17210, for AUREO1aΔJα of 15930 and for bZIP-A´α-LOV-Jα of 10810 L $mol^{-1}$ $cm^{-1}$ were used at 280 nm as calculated with EXTCOEF (http://www.workbench.sdsc.edu).

**FT-IR spectroscopy**

Non-reconstituted protein samples were used with concentrations of ~ 1.9 mM ($OD_{448}$~ 24) for bZIP-A´α-LOV-Jα and ~1.0 mM ($OD_{448}$~ 12) for AUREO1a with respect to the bound chromophore. FT-IR spectra were recorded on Bruker IFS 66v and IFS 66/S spectrometers. Long wave pass filters with a cut-on at 4.9 µm (OCLI) and 5.3 µm (Spectrogon), respectively, were placed in front of the detector to block stray light and improve the signal-to-noise ratio. FT-IR difference spectra with a resolution of 2 $cm^{-1}$ were recorded at 20 °C with 1024 scans before and after illumination for 2 s with a 455 nm LED (20 nm fwhm) with 20 mW $cm^{-2}$ at the sample (Philips Lumileds). Several experiments on independent preparations were averaged to a total number of 8192 scans for bZIP-A´α-LOV-Jα and 10240 scans for AUREO1a.

The ratio of DNA to protein was 1.5 to 2.0 with the concentrations of bZIP-A´α-LOV-Jα determined to ~0.98 mM ($OD_{448}$~8) with 2 mM aureo box, ~1.1 mM ($OD_{448}$~9) with ~1.6 mM G-

box and ~0.6 mM ($OD_{448}$~ 5) with ~0.9 mM STAT box. Extinction coefficients for double stranded oligo-nucleotides were calculated by using the equation $\varepsilon_D = (1 - h_{260nm})(\varepsilon_{S1} + \varepsilon_{S2})$ with $h_{260nm} = 0.287 f_{AT} + 0.059 f_{GC}$, where $f_{AT}$ and $f_{GC}$ are fractions of AT and GC base pairs, respectively (4). The results were 316742, 301456 and 382917 $M^{-1}$ $cm^{-1}$ for aureo box, G-box and STAT box, respectively. Extinction coefficients for single stranded oligo-nucleotides were calculated at the IDT Biophysics website (5). Several experiments on independent preparations were averaged to a total number of 15360 scans in presence of the aureo box, 21504 scans for the G-box and 8192 scans for the STAT box.

H/D-exchange of bZIP-A´α-LOV-Jα was performed by transferring the sample via ultrafiltration into 50 mM phosphate buffer in $D_2O$, pD 8.0 (pH 7.6), 100 mM NaCl and concentrating the sample to ~0.8 mM with respect to the bound chromophore ($OD_{448}$~ 10). 2 µL of the protein sample were applied to the $BaF_2$ cuvettes and were kept afterwards for at least 24 hours at room temperature in the dark. Several experiments on independent preparations were averaged to a total number of 17408 scans.

**Far-UV CD spectroscopy**

The bZIP domain is a stretch of two helices which holds a quasi-stable helical fold in the DNA-unbound form but when it is bound to the cognate DNA ligand (holobZIP) then attains a stable α-helix (6).To examine the conformational transition of bZIP-A´α-LOV-Jα upon dsDNA binding we used CD spectroscopy. CD spectra were recorded on JASCO (Easton, MD, USA) J-810 spectropolarimeter. Experiments were performed in 1 mm path length cells in light state (only) at 25 °C. Pure bZIP-A´α-LOV-Jα was mixed with aureo box DNA in different stoichiometric ratios in the binding buffer, either in 10 mM Tris-HCl pH8, 300 mM NaCl or in 0.5 XTBE. To ensure that no interference arises from the 'buffer only' spectrum, buffer and the samples were diluted 20 fold. Information for the structural transition was obtained by acquisition of far-UV (200–240 nm) CD spectra. Far-UV spectra are presented as the smoothed mean of three accumulations. Secondary structure prediction analysis was performed using the supplied J-810 software, using the Yang statistical algorithm (7). All recorded CD spectra were baseline-corrected by subtraction of the 'buffer only' spectrum.

**Small angle X-ray scattering data collection and analysis**

SAXS datasets were collected at the synchrotron Bio-SAXS beamline BM29 (ESRF) Grenoble, France (8). Measurements were performed in the dark (sample prepared under safe light $\lambda_{max}$~ 650nm and measured under complete dark condition) and blue light-illuminated states, respectively. The wavelength λ= 1.0 Å and the sample-to-detector distance of 2.43 m resulted in

scattering vectors, q, ranging from 0.0025 $Å^{-1}$ to 0.50 $Å^{-1}$. The scattering vector is defined as $q = 4\pi \sin\theta/\lambda$, where $2\theta$ is the scattering angle. All experiments were performed at 4 °C and data was processed by the ATSAS software package (9). SAXS data collection was performed in 10 mM Tris-HCl pH8, 300 mM NaCl and in 0.5 % TBE (50 mM Tris pH 8, 50 mM boric acid, 1 mM EDTA) buffer. One dimensional datasets were subtracted from the buffer only spectrum then merged analyzed in Primus (10). The light-dependent structural flexibility of the 1D-scattering of different concentrations were merged, where appropriate, given that the lower concentrations better represent low q data points. The radius of gyration was calculated by AutoRg (11). The $I_0$ was calculated considering bovine serum albumin (BSA) as standard, where the $R_g$ was 32.4±0.06 Å and $I_0$ 70.49±0.03. P(r) distance distribution functions were calculated by the program Gnom (12). Distance r, where P(r) functions approach zero probability, identifies the maximal dimension ($D_{max}$) of the macromolecule. The *ab initio* models of molecular envelopes were generated by Dammin (13) using P1 symmetry, which is a averaged and filtered representation of 20 different envelopes. Crysol (14) was used to generate $\chi$ values of the structural fits to the one dimensional scattering data. *Ab initio* envelopes and structures were finally visualised by PyMol (15). Theoretical atomic models of the double-stranded DNA in straight B-DNA was generated from the modelling server Haddock and the bZIP domain was modeled from Swiss modeling server (16, 17). The data sets are deposited in SASBDB (18) with accession codes: SASDB24 (bZIP-LOV module in Tris.HCl buffer, Dark state), SASDB34(bZIP-LOV module in Tris.HCl buffer, Light state), SASDBX3 (20bp long aureo box containing dsDNA), SASDBY3 (bZIP-LOV module in 0.5X TBE buffer, light state) and SASDBZ3 (20bp long dsDNA bound bZIP-LOV module in 0.5X TBE buffer, light state).

**Calculation of domain movement and Molecular dynamic simulation**

Changes of relative domain orientations within LOV domain dimers upon dark/light transitions have been calculated by the lsqkab programme within the CCP4 suite using PDB entries 5A8B/5DKK for PtAureo1a and 2PR5/2PR6 for YtvA, respectively.

All molecular dynamic simulations were performed with the AMBER14 package (Case *et al,* 2014), using the amber ff14SB force field (19), TIP3P water model, and the corresponding monovalent ion parameters. A PDB file corresponding to energy minimized SAXS model was used as input for tleap. The protonation state of the structure was predicted via the H++ protonation server (20), and used as is. FMN parameters were obtained from the amber parameter database (21). The protein-DNA complex was neutralized with sodium ions, while ten extra sodium and 10 chloride ions were added to account for a moderate ionic strength.

Minimization was performed in three to four steps. First, solute atoms were constrained with a 500 kcal/mol harmonic restraint. SHAKE-restrained (22) water and ions were allowed to relax for 500 cycles of steepest descent, followed by 4500 steps of conjugate gradient. Next, the restraints were lifted from the modeled loops, which were then subjected to 500 cycles of steepest descent and 4500 steps of conjugate gradient. In a third step, restraints on ions, water, loops, and FMN molecules were all lifted, and the molecules were allowed to relax under the same conditions as before. Fourth, the same procedure was applied, this time lifting DNA restraints. Finally, the harmonic restraints on the protein were lifted and the whole system was relaxed for the same amount of cycles.

After minimization, temperature was slowly raised over 50 ps from 0 to 300 K, using a Langevin thermostat (23) (with random seed and $\gamma$=5 $ps^{-1}$), applying weak restraints to the solute molecules. Next, another 50 ps of constant volume simulation at 300 K was performed, in order to further equilibrate the system. Finally, constant pressure, restraint-free equilibration to one atmosphere was carried out for 50 ps (Monte-Carlo barostat, pressure relaxation time 2ps).

After equilibration, Self-guided Langevin dynamics (LD, guiding effect strength set to 0.5, target guiding temperature 300 K), were run for 1.5 ns, with a moving, rigid body fitted EMAP restraint (24, 25) based on an a priori coarse SAXS model. The EMAP restraint was set with a force constant of 0.02 kcal/g and a map resolution of 2.0 Å. The LD simulation was run until it was clear that the EMAP restraint energy had converged (Fig.S5E). Keeping the same parameters as in the constant pressure equilibration step, the system was allowed to relax for 2 ns. After relaxation, the system was run for a further 138 ns.

CPPTRAJ (26) was used to extract data from the production trajectories, including RMSDs, population binning, and Boltzmann free energies. Energies were obtained via the Boltzmann equation, binning whole bZIP domain backbone RMSD populations and converting binned and smoothed data into frequencies (bin occupancy/total bin population). Maxima and minima were calculated via numerical differentiation, and values were subtracted to obtain energy barriers Line art in figures were performed with the Qtiplot software. MMGBSA (27) calculations were used to obtain per-bin interaction energies. Plotting of data was performed with Qtiplot. PyMol (15) was also used for structural figure rendering.

**SUPPLEMENTARY REFERENCES**

1. Herman,E., Sachse,M., Kroth,P.G. and Kottke,T. (2013) Blue-light-induced unfolding of the Jα helix allows for the dimerization of aureochrome-LOV from the diatom Phaeodactylum

tricornutum. *Biochemistry*, **52**, 3094–3101.

2. Takahashi,F., Yamagata,D., Ishikawa,M., Fukamatsu,Y., Ogura,Y., Kasahara,M., Kiyosue,T., Kikuyama,M., Wada,M. and Kataoka,H. (2007) AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 19625–19630.
3. Jerabek-Willemsen,M., André,T., Wanner,R., Roth,H.M., Duhr,S., Baaske,P. and Breitsprecher,D. (2014) MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.*, **1077**, 101–113.
4. Tataurov,A. V., You,Y. and Owczarzy,R. (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys. Chem.*, **133**, 66–70.
5. Cantor,C.R., Warshaw,M.M. and Shapiro,H. (1970) Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers*, **9**, 1059–1077.
6. John,M., Leppik,R., Busch,S.J., Granger-Schnarr,M. and Schnarr,M. (1996) DNA binding of Jun and Fos bZip domains: Homodimers and heterodimers induce a DNA conformational change in solution. *Nucleic Acids Res.*, **24**, 4487–4494.
7. Yang,J.T., Wu,C.S. and Martinez,H.M. (1986) Calculation of protein conformation from circular dichroism. *Methods Enzymol.*, **130**, 208–269.
8. Pernot,P., Round,A., Barrett,R., De Maria Antolinos,A., Gobbo,A., Gordon,E., Huet,J., Kieffer,J., Lentini,M., Mattenet,M., *et al.* (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.*, **20**, 660–664.
9. Petoukhov,M. V, Franke,D., Shkumatov,A. V, Tria,G., Kikhney,A.G., Gajda,M., Gorba,C., Mertens,H.D.T., Konarev,P. V and Svergun,D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr*, **45**, 342–350.
10. Konarev,P. V., Volkov,V. V., Sokolova,A. V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
11. Petoukhov,M. V., Konarev,P. V., Kikhney,A.G. and Svergun,D.I. (2007) ATSAS 2.1 - Towards automated and web-supported small-angle scattering data analysis. In *J. Appl. Cryst. **40**, s223-s228*
12. Svergun,D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
13. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, **76**, 2879–2886.

14. Svergun,D., Barberato,C. and Koch,M.H. (1995) CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.
15. DeLano,W.L. (2002) The PyMOL Molecular Graphics System. *Schrödinger LLC.*
16. Van Dijk,M., Van Dijk,A.D.J., Hsu,V., Rolf,B. and Bonvin,A.M.J.J. (2006) Information-driven protein-DNA docking using HADDOCK: It is a matter of flexibility. *Nucleic Acids Res.*, **34**, 3317–3325.
17. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
18. Valentini,E., Kikhney,A.G., Previtali,G., Jeffries,C.M. and Svergun,D.I. (2015) SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.*, **43**, D357–D363.
19. Wang,J., Wolf,R.M., Caldwell,J.W., Kollman,P.A. and Case,D.A. (2004) Development and testing of a general Amber force field. *J. Comput. Chem.*, **25**, 1157–1174.
20. Anandakrishnan,R., Aguilar,B. and Onufriev,A. V. (2012) H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.*, **40**. W537-41
21. Schneider,C. and Sühnel,J. (1999) A molecular dynamics simulation of the flavin mononucleotide-RNA aptamer complex. *Biopolymers*, **50**, 287–302.
22. Miyamoto,S. and Kollman,P.A. (1992) SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.*, **13**, 952–962.
23. Loncharich,R.J., Brooks,B.R. and Pastor,R.W. (1992) Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. *Biopolymers*, **32**, 523–535.
24. Wu,X., Milne,J.L.S., Borgnia,M.J., Rostapshov,A. V., Subramaniam,S. and Brooks,B.R. (2003) A core-weighted fitting method for docking atomic structures into low-resolution maps: Application to cryo-electron microscopy. *J. Struct. Biol.*, **141**, 63–76.
25. Wu,X., Subramaniam,S., Case,D.A., Wu,K.W. and Brooks,B.R. (2013) Targeted conformational search with map-restrained self-guided Langevin dynamics: Application to flexible fitting into electron microscopic density maps. *J. Struct. Biol.*, **183**, 429–440.
26. Roe,D.R. and Cheatham III,T.E. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J. Chem. Theory Comput.*, **9**, 3084–3095.
27. Miller,B.R., McGee,T.D., Swails,J.M., Homeyer,N., Gohlke,H. and Roitberg,A.E. (2012) MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.*, **8**, 3314–3321.