

SUPPORTING INFORMATION

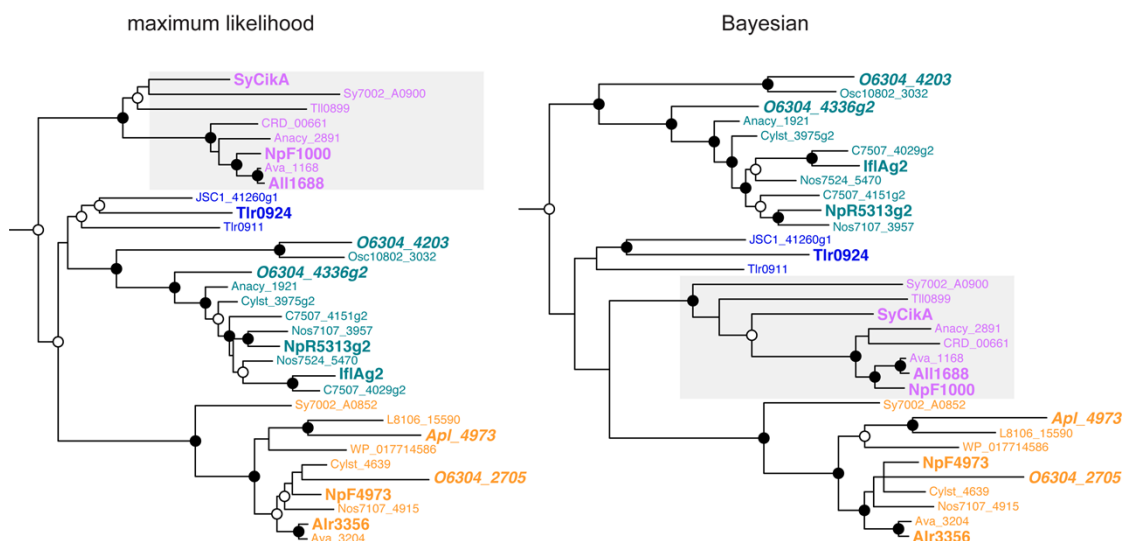


Figure S1. Phylogenetic clustering of SyCikA, NpF1000, and All1688. Maximum likelihood (*left*) and Bayesian (*right*) analyses are shown. Both are subsets of the larger analyses presented in Fig. 3.

SPLC1_S541230	ATGAACTTCCCTAATATATATTCGGTGATTCAACCAATAGCAACCCTAATCCTAACAGTGAT
ARTHRO_30350	ATGAACTTCCCTAATATATATTCGGTGATTCAACCAATAGCAACCCTAATCCTAACAGTGAT
Apl_4973	ATG GGG CTTCCCTAATATATATTCGGTGATTCAACCAATAGCAACCCTAATCCTAACAGTGAT
	*** *****
SPLC1_S541230	CCACGCTTTCAACAATTTGTTGAACGTTTACAACAGGTGGAAGATCGCGATCGCCTAATT
ARTHRO_30350	CCACGCTTTCAACAATTTGTTGAACGTTTACAACAGGTGGAAGATCGCGATCGCCTAATT
Apl_4973	CCACGCTTTCAACAATTTGTTGAACGTTTACAACAGGTGGAAGATCGCGATCGCCTAATT

SPLC1_S541230	CAGGAAACCCTGAATGGGTTACGAAAACACCTGAAAAGCGATCGGATAGTCTTATACTAC
ARTHRO_30350	CAGGAAACCCTGAATGGGTTACGAAAACACCTGAAAAGCGATCGGATAGTCTTATACTAC
Apl_4973	CAGGAAACCCTGAATGGGTTACGAAAACACCTGAAAAGCGATCGGATAGTCTTATACTAC

SPLC1_S541230	TTTTATAAGCAGTGGCGCGGACAAGTTACCTTTGAATCACTGCGTTATATCGACTATTTCG
ARTHRO_30350	TTTTATAAGCAGTGGCGCGGACAAGTTACCTTTGAATCACTGCGTTATATCGACTATTTCG
Apl_4973	TTTTATAAGCAGTGGCGCGGACAAGTTACCTTTGAATCACTGCGTTATATCGACTATTTCG

SPLC1_S541230	ATTTATGGCTCAACTGGGCTGATGATTGTTTCAATGATCAATATGCTAGACTCTATTTA
ARTHRO_30350	ATTTATGGCTCAACTGGGCTGATGATTGTTTCAATGATCAATATGCTAGACTCTATTTA
Apl_4973	ATTTATGGCTCAACTGGGCTGATGATTGTTTCAATGATCAATATGCTAGACTCTATTTA

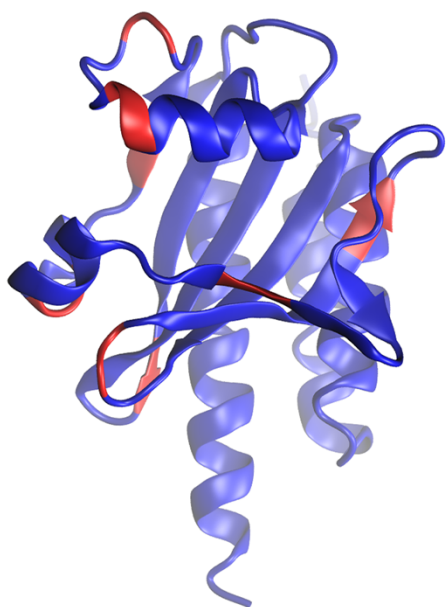
SPLC1_S541230	GAAGGGAGAATATCGGCAATTTTCAGATATTGATGAAGCTGACATTGACTCCTGTCACCGC
ARTHRO_30350	GAAGGGAGAATATCGGCAATTTTCAGATATTGATGAAGCTGACATTGACTCCTGTCACCGC
Apl_4973	GAAGGGAGAATATCGGCAATTTTCAGATATTGATGAAGCTGACATTGACTCCTGTCACCGC

SPLC1_S541230	GATTTTCTGAAAAGCATCAGAGTTAAGGCAAATTTAGCGGTTCCAGTGGGTTAAAAATCAA
ARTHRO_30350	GATTTTCTGAAAAGCATCAGAGTTAAGGCAAATTTAGCGGTTCCAGTGGGTTAAAAATCAA
Apl_4973	GATTTTCTGAAAAGCATCAGAGTTAAGGCAAATTTAGCGGTTCCAGTGGGTTAAAAATCAA

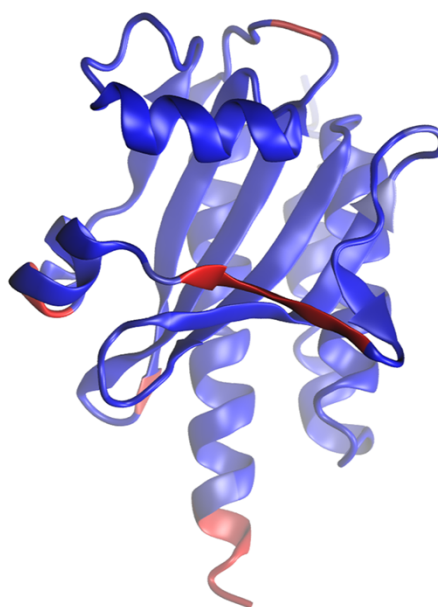
SPLC1_S541230	CAATTATGGGGATTACTAATTGCCCACTATTGTCAAACCTCCCACTCTTGGGTTAATGCC
ARTHRO_30350	CAATTATGGGGATTACTAATTGCCCACTATTGTCAAACCTCCCACTCTTGGGTTAATGCC
Apl_4973	CAATTATGGGGATTACTAATTGCCCACTATTGTCAAACCTCCCACTCTTGGGTTAATGCC

SPLC1_S541230	GAAATTGAAACCATGAAACAGGCGGCTAAATTATTATCAGAATCACCATTTTGGACTAA
ARTHRO_30350	GAAATTGAAACCATGAAACAGGCGGCTAAATTATTATCAGAATCACCATTTTGGACTAA
Apl_4973	GAAATTGAAACCATGAAACAGGCGGCTAAATTATTATCAGAATCACCATTTTGGAC---

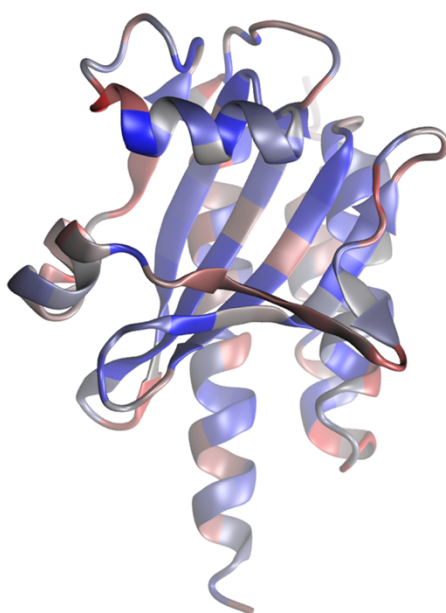
Figure S2. Sequence data for Apl_4973. Nucleotide sequence alignment of the PCR product spanning the candidate *Apl_4973* gene with identical sequences from known *Arthrospira* genomes. PCR primers used in amplification (see Methods) introduced a Gly residue at position two (red highlight) and deleted the stop codon to facilitate cloning and expression.



AnPixJg2 gaps
relative to others



AnPixJg2 insertions
relative to others



BLOSUM62 similarity
relative to AnPixJ

Figure S3. Structure-based analysis of the multiple sequence alignment. The crystal structure of AnPixJ in the red-absorbing *15Z* dark state (PDB accession 3W2Z) is shown colored by gap frequency relative to the sequences in Supporting File 1 (*top left*), insertion frequency relative to the sequences in Supporting File 1 (*top right*), and

similarity of the sequences in Supporting File 1 relative to AnPixJg2 using a BLOSUM62 similarity matrix (*bottom*). Gap and insertion frequency are scaled from blue to red with increasing frequency. Gaps are defined as positions at which AnPixJg2 lacks an amino acid but ≥ 1 other CBCR has one and are scored on the flanking positions. Insertions are defined as positions at which AnPixJg2 has an amino acid but ≥ 1 other CBCR lack one and are scored directly on each position. Similarity is colored from dark blue (conserved) through white to bright red (highly variable).

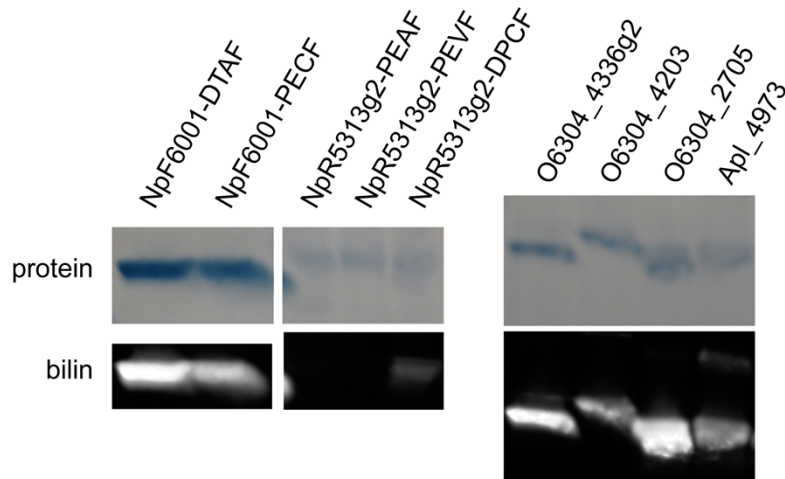


Figure S4. Characterization of previously uncharacterized purified CBCR proteins.

Total protein (*top*) and bilin content (*bottom*) are shown for the indicated proteins.