

RESEARCH

Open Access



Assessing nucleic acid binding activity of four dinoflagellate cold shock domain proteins from *Symbiodinium kawagutii* and *Lingulodinium polyedra*

Bahareh Zaheri and David Morse*

Abstract

Background: Dinoflagellates have a generally large number of genes but only a small percentage of these are annotated as transcription factors. Cold shock domain (CSD) containing proteins (CSPs) account for roughly 60% of these. CSDs are not prevalent in other eukaryotic lineages, perhaps suggesting a lineage-specific expansion of this type of transcription factors in dinoflagellates, but there is little experimental data to support a role for dinoflagellate CSPs as transcription factors. Here we evaluate the hypothesis that dinoflagellate CSPs can act as transcription factors by binding double-stranded DNA in a sequence dependent manner.

Results: We find that both electrophoretic mobility shift assay (EMSA) competition experiments and selection and amplification binding (SAAB) assays indicate binding is not sequence specific for four different CSPs from two dinoflagellate species. Competition experiments indicate all four CSPs bind to RNA better than double-stranded DNA.

Conclusion: Dinoflagellate CSPs do not share the nucleic acid binding properties expected for them to function as *bone fide* transcription factors. We conclude the transcription factor complement of dinoflagellates is even smaller than previously thought suggesting that dinoflagellates have a reduced dependence on transcriptional control compared to other eukaryotes.

Keywords: Transcription factors, Cold shock domain proteins, Dinoflagellates, RNA binding domain, DNA binding domain, Transcription

Background

Dinoflagellates are an important group of unicellular eukaryotes perhaps best known for their large genomes and permanently condensed chromosomes. Surprisingly, little is known how gene expression is regulated in these organisms. Transcriptome analyses in several species, including *Lingulodinium* and *Symbiodinium*, have revealed a general paucity (typically 0.15%) of sequences

annotated as transcription factors (TF). This is in sharp contrast to the roughly 6% of genes annotated as TF in plants [1] or animals [2]. In addition, a high proportion (~60%) of the annotated dinoflagellate TF in transcriptomes are cold shock domain (CSD) containing proteins (CSPs) [3, 4] yet this class is typically less than 1% of the TF in other eukaryotes. CSDs are small (roughly 70 amino acid) nucleic acid binding domains containing two conserved RNA recognition motifs, KGFGFI and VEVHE, that are known to bind both DNA and RNA. All dinoflagellate CSPs contain the two RNA binding motifs characteristic of the CSD. Four divergent domain

* Correspondence: david.morse@umontreal.ca

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, 4101 Sherbrooke Est, Université de Montréal, Montréal H1X 2B2, Canada



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

structures have been found in *Lingulodinium* and *Symbiodinium* proteins, the most prevalent ones containing a CSD either alone or with a C-terminal G-rich domain. Less frequently observed are some structures containing a Zn-finger domain following the G-rich domain, and also examples of sequences with multiple CSDs and one or more RNA recognition motifs (RRM). Thus, many of the dinoflagellate CSPs are similar to what are found in bacteria as these typically contain only a CSD [5].

In *E. coli*, CSPs have a wide range of functions, including binding DNA as transcription factors, binding to RNA, regulating transcription, splicing, and translation, and affecting mRNA stability as RNA chaperones [6, 7]. Bacterial CSPs have a non-specific RNA binding function during cold stress, which is correlated to their chaperone activity, and this helps transcription by acting as an antiterminator [7, 8]. However, the dinoflagellate proteins may be different from their bacterial counterparts as two *Lingulodinium* CSPs, both containing a single CSD followed by a glycine-rich C-terminal region, were both unable to complement the growth of an *E. coli* strain lacking four different CSP genes at low temperature [5]. Furthermore, cold temperatures did not induce the CSP transcripts in *L. polyedra* [9]. Previous work on *L. polyedra* CSPs showed binding to both single- and double-stranded DNA as well as to RNA, but it was unclear if binding would show any sequence specificity that would be likely if they were to function as transcription factors [5]. Here we performed two experimental approaches to assess the specific nucleic acid binding activity of *L. polyedra* CSP1 (*LpCSP1*) and three *S. kawagutii* CSPs (*SkCSP1*, *SkCSP2* and *SkCSP3*). Initially, these four CSPs were expressed, purified and used in electrophoretic mobility assays (EMSAs) to

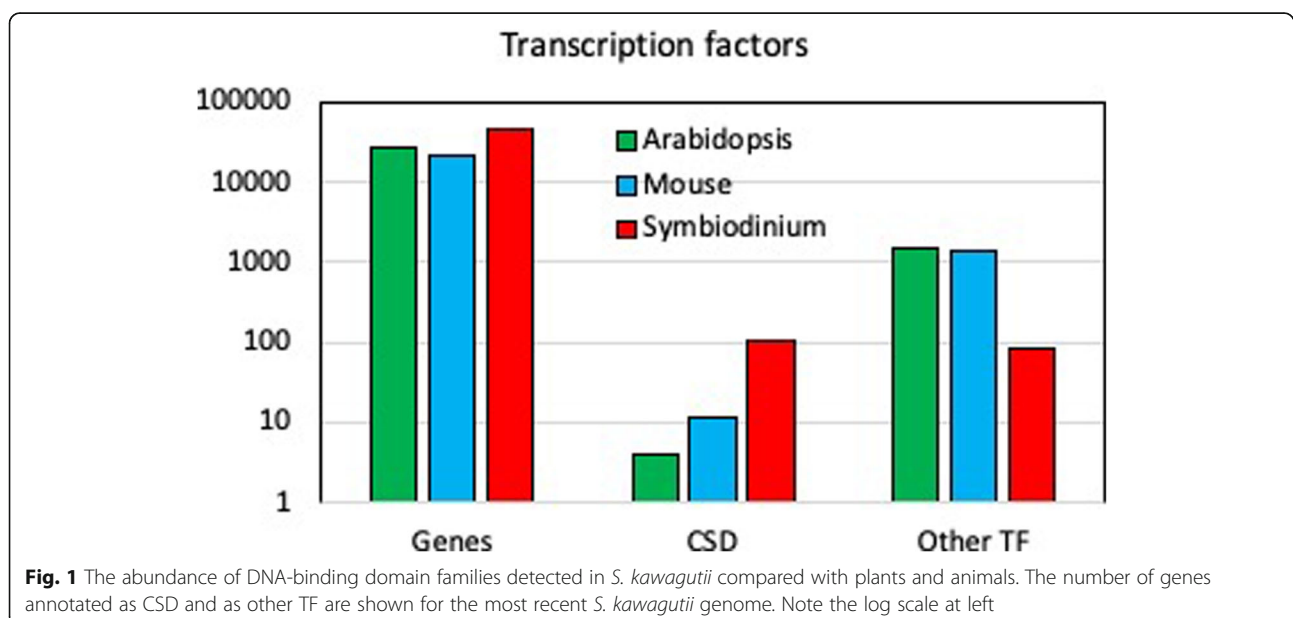
measure if they were active in binding nucleic acids. In a second approach, selection and amplification binding assays (SAAB) was used to determine if these proteins could bind a specific sequence on DNA. All these CSPs were able to bind to DNA and RNA, and no sequence specific binding activity toward DNA was observed.

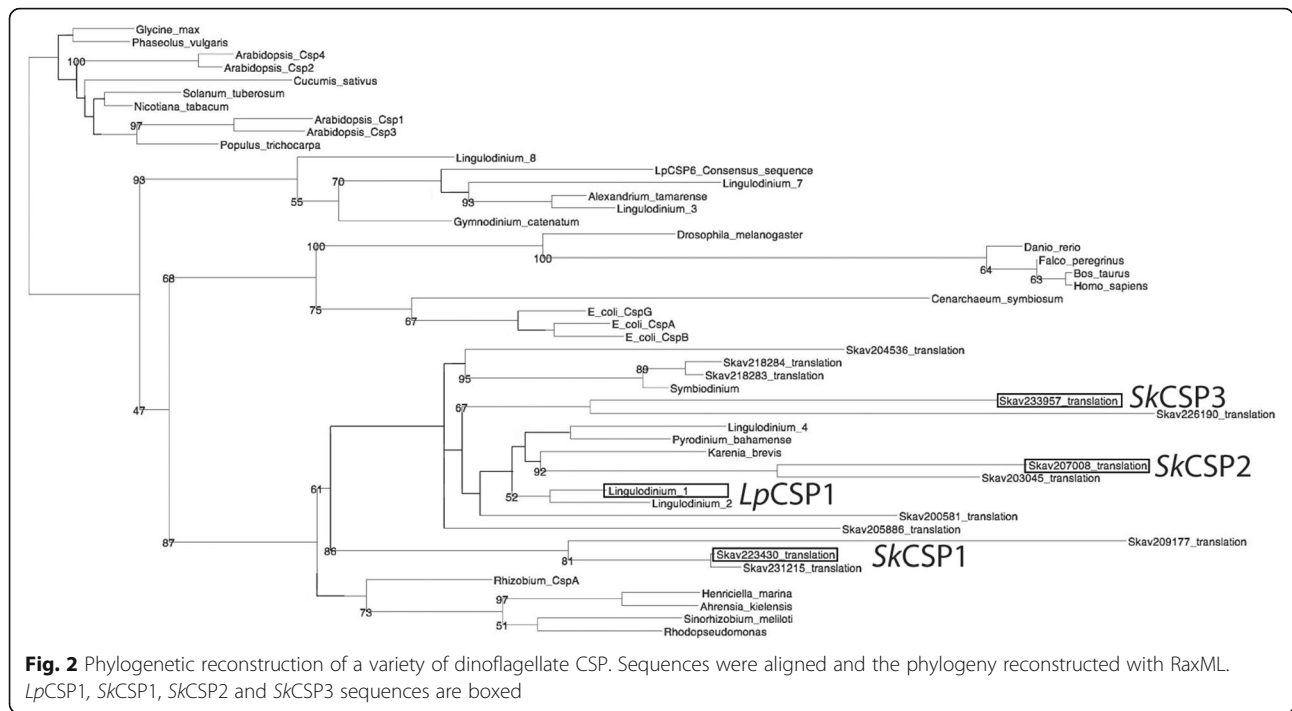
Results

SkCSP1, *SkCSP2* and *SkCSP3* belong to a *Symbiodinium* unique clade

The number of annotated DNA binding proteins in the genome of the *S. kawagutii* genome [10] belonging either to CSD family or other TF (Fig. 1) shows the relative importance of CSDs in dinoflagellates compared to plants and animals. All CSDs contain the two RNA recognition motifs (KGFGFI and VFVHF) shared with bacteria and plants [3, 4]. Phylogenetic analysis of CSDs from 12 predicted *Symbiodinium kawagutii* protein sequences was performed using RaxML, and all were found to cluster together within a single well defined clade together with some bacterial sequences (Fig. 2, Table S2). This is slightly different from the situation in *Lingulodinium* where sequences are distributed among two different clades. The phylogenetic positions of the four CSPs examined here: *LpCSP1* (JO732587), *SkCSP1* (Skav223430), *SkCSP2* (Skav207008) and *SkCSP3* (Skav233957) are boxed.

LpCSP1 with a size of 113 amino acids has been previously cloned [5]. For this study, *SkCSP1*, *SkCSP2* and *SkCSP3* were also cloned and have sizes of 128, 120 and 182 residues, respectively. All four CSPs were expressed as GST-tagged proteins and used for EMSA after removal of the GST tag (Fig. S1). Two of the *S. kawagutii* proteins contain an N-terminal extension (Fig. 3).





Lingulodinium and Symbiodinium CSPs bind to DNA and RNA

EMSA experiments were conducted on *LpCSP1*, *SKCSP1*, *SKCSP2* and *SKCSP3* to analyze their binding to radiolabeled double-stranded (dsDNA), single-stranded (ssDNA) and RNA probes (Fig. 4). Fusion proteins still containing the glutathione *S*-transferase (*GST*) tags also bind nucleic acids but migrating slower on the gel, and all EMSA experiments used proteins after removal of the tag by thrombin.

All proteins were able to bind dsDNA, ssDNA and RNA as seen by the presence of a radioactive band of lower mobility. The mobility of probe sequence was reduced to roughly the same extent with all proteins with the exception of *LpCSP1* binding to ssDNA or RNA. The amount of the reduced mobility band seemed to

increase with increasing concentrations of the CSPs, although not precisely proportional to the amount of protein. We conclude that all four CSPs were able to bind to all three types of nucleic acids tested.

Symbiodinium CSP1 shows preferential binding to single-stranded nucleic acids

To assess the specificity of *Symbiodinium* CSPs interactions with different nucleic acid substrates, binding to dsDNA and ssDNA probes was evaluated using *SKCSP1* and unlabeled (cold) competitors (Fig. 5). When dsDNA was used as a probe, the intensity of the slowly migrating bands decreased dramatically when the amount of competing cold ssDNA was increased. In contrast, band intensity using ssDNA probes was mostly stable using increasing amounts



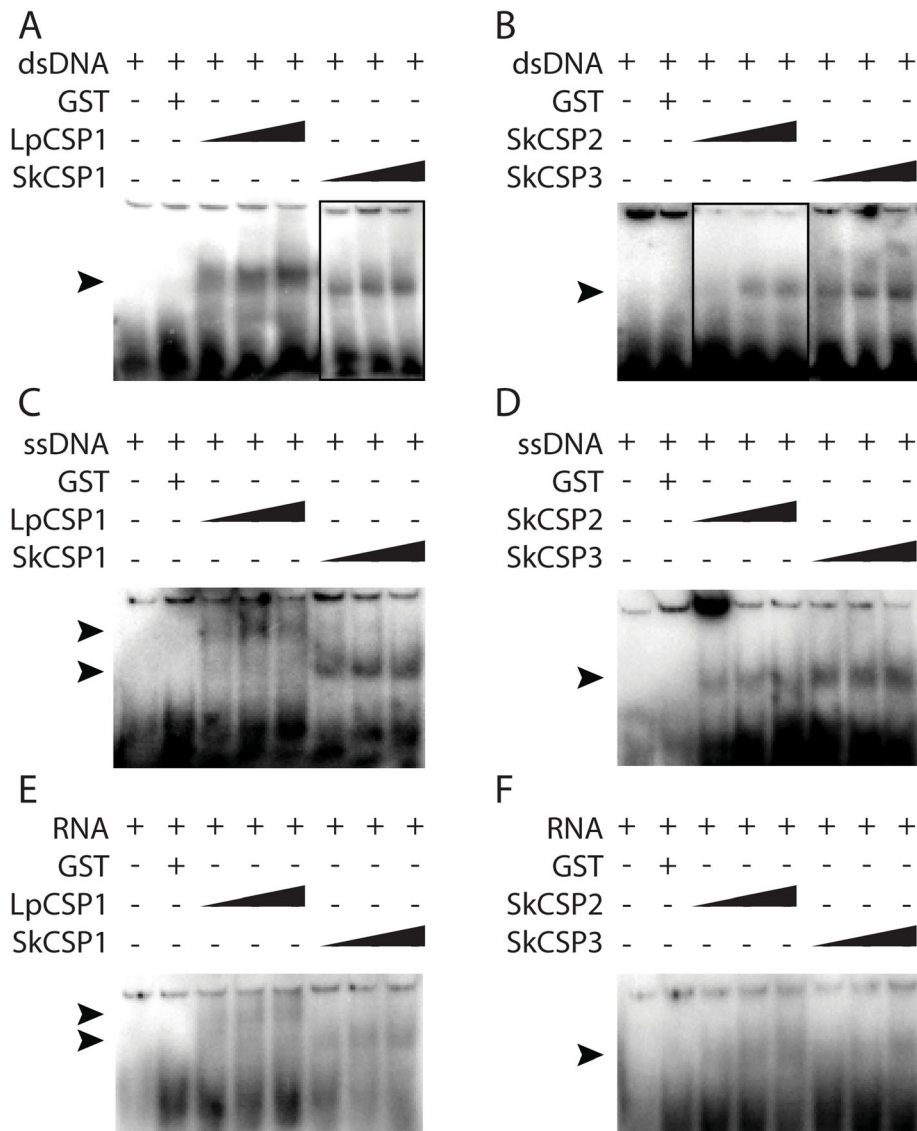


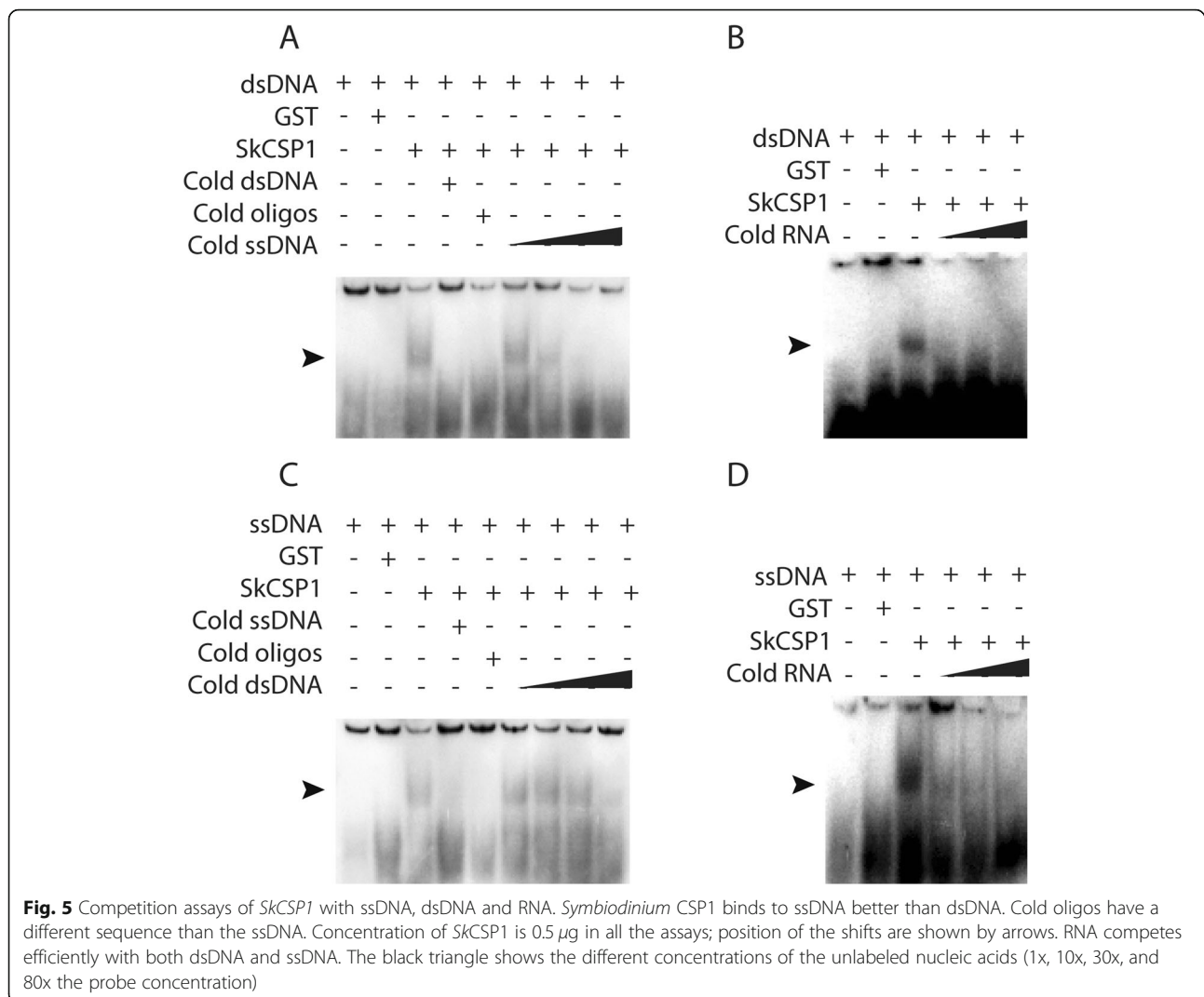
Fig. 4 Nucleic acid binding activity of *L. polyedra* and *S. kawagutii* CSPs in EMSA. ssDNA (**a**, **b**), dsDNA (**c**, **d**) and RNA (**e**, **f**) probes were used. The black triangle shows the different concentrations of the CSPs (0.5, 1 and 3 μ g in all the assays); position of the shifts are shown by arrows

of cold dsDNA. Furthermore, RNA appears to compete efficiently with both dsDNA and ssDNA. These results indicate that *SkCSP1* has a preference for single-stranded nucleic acids, with RNA preferred over DNA. This is consistent with a previous report for *Lingulodinium* CSP1 [5]. While the potential tendency to bind to ssDNA may support a role for these proteins in uncoiling the DNA structure during transcription, preferential binding to RNA suggests this may not be their primary role.

L. polyedra and *S. kawagutii* CSPs bind non-specifically to DNA and RNA

To assess the possibility of sequence specific binding of *Lingulodinium* and *Symbiodinium* CSPs to dsDNA, we

performed a selection and amplification binding enrichment (SAAB) with DNA containing 9 random nucleotides (N9) flanked by PCR primers. These experiments used the fusion proteins directly to facilitate purification of bound DNA sequences, as the presence of the GST tag did not affect DNA binding on EMSA assays. After 3 rounds of SAAB, samples containing double-stranded N9 enriched by binding to *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* were sequenced (Fig. 6). Over 12,000 sequences were obtained for each CSP, but sequence alignments after binding to all four shows no evidence for a consensus motif for any of the CSPs (Fig. 7). We conclude that there is no specific dsDNA which can be enriched by binding to *Lingulodinium* or *Symbiodinium* CSPs.



Discussion

Cold shock domain (CSD) proteins were recognized in *Escherichia coli* during cold shock stress [7, 11, 12]. The conservation of CSD in these proteins was discovered in bacteria, archaea, plants, and animals. In prokaryotes, CSPs containing only a CSD act mainly as RNA chaperones. Some *E. coli* CSPs are cold-inducible and act as RNA chaperones disrupting RNA secondary structures [7, 13]. They are also involved in the transcription regulation by binding specifically to *gyrA* promoter (CspA) [12, 14]. In eukaryotes, CSPs are composed of CSD and additional domains and aid in responding to cold stress, nutrient limitation and growth [7, 13, 15–17]. Plants CSPs are engaged in regulation of translation during cold stress and also complicated physiological processes such as seed and flower germination [7, 18, 19]. In *A. thaliana*, CSP3 interacts with other proteins involved in mRNA processing path [19]. A vertebrate CSP called YB1 (Y-box binding protein) is responsible for the

regulation of transcription by binding to a Y-box specific sequence, and is also involved in regulation of translation and RNA processing [20–24] and DNA repair [7, 12, 25]. YB1 prefers to bind to ssDNA rather than dsDNA, thus disentangling the double helix structure of DNA has been proposed for the activation of transcription [7, 26]. YB1 also prefers RNA over ssDNA [7] with the consensus CA(U/C) C sequence as the RNA-binding site [27, 28]. In dinoflagellates, CSPs are mostly in the form of one conserved CSD either alone or with a C-terminal G-rich domain [5]. Previously, a Y-box sequence (CTGATTGGCT) was used to study the binding specificity of *L. polyedra* CSPs [5]. Here we used different random C-rich sequences to test the possibility of sequence privileged targeting. For the SAAB assay, we synthesized a DNA sequence with 9 random nucleotides (N9) nestled between flanking PCR primers. The goal of this experiment was to see if several cycles of binding, elution and amplification would enrich for a particular sequence

motif that could constitute a potential promoter element. However, no sequence motifs were enriched by binding to any of the four CSPs indicating that these proteins are unlikely to function as conventional sequence-specific transcription factors. It is not possible to rule out a role in DNA unwinding similar to what has been proposed for YB1, in which non-specific binding of CSPs to ssDNA was thought to help stabilize the structure, but it must be noted CSPs have no known helicase activity.

The importance of examining the nucleic acid binding properties of CSPs is due to the finding that the majority of the proteins annotated as transcription factors in the transcriptome of *Lingulodinium* [3], *Symbiodinium* [4] and the genome of *Symbiodinium* [10, 29] (Fig. 1) are CSDs. Our hypothesis was that to act as transcription factors, dinoflagellate CSPs should bind to dsDNA in a sequence specific manner. We assessed nucleic acid binding activity of *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* using two different approaches. In one approach, electrophoretic mobility shifts assays (EMSA) were used to show that all four CSPs could bind both double- and single- stranded DNA as well as RNA (Fig. 4). When tested in competition EMSA experiments, RNA was found to compete with binding to DNA probes better than DNA competed with binding to RNA probes (Fig. 5). These characteristics are not what would be predicted for a transcription factor. In a second approach, selection and amplification binding (SAAB) experiments showed none of the four CSPs tested enriched a specific motif after three cycles of binding and PCR amplification, again inconsistent with a role as a sequence specific transcription factor.

Our results indicate that *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* binding to nucleic acids does not depend on sequence. We infer that the dinoflagellate CSPs in general are unlikely to act as sequence-specific transcription factors. Although only one *S. kawagutii* CSP (*SkCSP1*) was extensively analyzed by competition EMSA, the similarity to the *Lingulodinium* CSP1 suggests the nucleic acid binding properties found may be a consistent lineage-specific feature. The balance of the evidence thus suggests that CSPs do bind nucleic acids, thus explaining why they were annotated as transcription factors. However the details of the binding suggest they are unlikely to play this role in vivo. Additional characterization studies of dinoflagellate CSPs would be essential to recognize more about their function and possible interaction with other partners.

Conclusions

The four CSPs examined here do not bind DNA in a sequence specific manner. Furthermore, *SkCSP1* prefers binding to single-stranded RNA. CSPs are unlikely to function as transcription factors in dinoflagellates.

Methods

Cell cultures

Cultures of *Symbiodinium kawagutii* (strain CCMP2468) and *Lingulodinium polyedra* (strain CCMP1936) were obtained from the National Center for Marine Algae (Boothbay Harbor, Maine). Cells were grown in f/2 sea water medium prepared from Instant Ocean under 12 h cool white fluorescent light and 12 h darkness as described [30] except that the temperature was $25 \pm 1^\circ\text{C}$ for *S. kawagutii*.

Phylogenetic reconstruction and primer design

The CSP sequences for *Lingulodinium* and *Symbiodinium* were obtained from the dinoflagellate transcriptomes deposited at NCBI and from the *Symbiodinium kawagutii* genome at the Symbiodiniaceae and Algal Genomic Resource (SAGER) database [29]. Phylogenetic analysis of CSDs from the predicted protein sequences (Table S2) was performed using a webserver for alignments (http://www.phylo.org/sub_sections/portal/) [31]. The server performs sequence alignments using MUSCLE, and curation using GBLOCKS. Phylogenetic reconstructions were built with RaxML using the CIPRES portal (http://www.phylo.org/sub_sections/portal/). Trees were visualized by TreeDyn. Primers were designed using Geneious software [32] or BLAST integrated into Galaxy [33] for amplification and subsequent cloning of the CSPs. Geneious software [32] was also used for sequence alignments.

Cloning, expression and purifying of CSPs

Symbiodinium cultures were harvested by centrifugation and the pellets frozen in liquid nitrogen. Frozen pellets were crushed into a fine powder using a pre-chilled mortar and pestle, and the powder was added to Trizol (Invitrogen). Primer pairs based on sequences from the *Symbiodinium* transcriptome or genome were used to amplify CSPs from a first strand cDNA reaction product using the total RNA extracted from *Symbiodinium* cells as described [5]. For cDNA amplification, the reverse transcription reaction was performed with ProtoScript II first strand cDNA synthesis kit (New England BioLabs). The sequences were cloned into the pGEM-T vector (Promega) and sequenced. A second PCR was performed on the insert in the pGEM-T plasmid using primers containing restriction sites required for directional cloning into the bacterial expression vectors pGEX-4 T2 (GE Healthcare) [34] (Table. S1 in the supplementary data). The reading frame of all clones were confirmed by sequencing and the size of the CSP fusion protein verified by SDS PAGE (Fig. S1). The pGEX4T2 vectors containing CSP sequences were used to transform the chemically competent cells of BL21. Liquid Luria Bertani (LB) medium was used to grow one colony of transformed *E. coli* overnight at 37°C with vigorous shaking in the

presence of ampicillin to maintain selection for the plasmid. Protein expression were induced using Isopropyl β -D-1-thiogalactopyranoside (IPTG). Cells were collected by centrifugation, resuspended in PBS buffer and broken in a French pressure cell (Fisher Scientific). The cell lysates were then centrifuged and the supernatants were incubated with Glutathione Sepharose 4B beads (Promega) for 45 min at room temperature with end-over-end agitation. Beads were washed 4 times in PBS and resuspended in PBS supplemented with thrombin to cleave the GST tag. The size, and purity of the single CSPs were then analyzed by SDS-PAGE on acrylamide gel (Fig. S1) and the Bradford assay (BioRad) was used to assess the protein concentration.

CSP electrophoretic mobility shift assays

[γ - 32 P] ATP (PerkinElmer) was used to 5'-end-label 32 nt ssDNA 5'-TCCGCCCTCCCTCCCCCGCCCTCCCTCCCA-3' and 25 bp dsDNA 5'-GGCGCCCTCCCTCCGCCCTCCCTCA-3' C-rich sequences using a T4 polynucleotide kinase kit (NEB). A QIAquick nucleotide removal kit (Qiagen) was used for removing the unincorporated nucleotides and purifying the probes. Either dsDNA or ssDNA 32P-labelled probes (1 ng) and increasing concentrations of CSPs (0.5–3 μ g) were incubated in 20 μ L of 2x binding buffer (20 mM Tris-Cl [pH 7.0], 20 mM MgCl₂, 50 mM KCl, 10% glycerol and 1 mM DTT) for 30 min at room temperature. The CSP/DNA complexes were run through a 5% native polyacrylamide gels for 45 mins at 80 V in 1x Tris-borate-EDTA (TBE) buffer at room temperature. The gels were dried immediately and exposed overnight at -80 °C with a phosphorimager screen (Amersham). The images were analyzed with a Typhoon Trio+ (Amersham) using ImageQuant 5.2. Competition reactions were prepared by incubation of the CSPs and increasing amounts of cold unlabeled ssDNA, dsDNA or RNA probes (described below) for specific binding and a 40x excess of random 22 nt single-stranded oligonucleotide (TTATTGGGGC ACACCGCATGCT) for non-specific competition in the binding buffer for 15 mins before adding the radiolabeled probes.

Forty nt RNAs were synthesized by T7 RiboMAX RNA production kit (Promega) using dsDNA templates containing the N9 and T7 promoter sequences. Thereafter, RQ1 RNase-free DNase (Promega) was used for degradation of the dsDNA templates. The in vitro transcribed RNAs were quantitated using spectrophotometry (1.2 μ g/ μ L), end-labeled using [γ - 32 P] ATP (PerkinElmer) (see above) and purified using filtration chromatography on a Bio-Gel P10 column (Bio-Rad). One ng labelled probe was incubated with increasing concentrations of CSPs in the binding reactions as described above.

Selection and amplification binding assays

Symbiodinium and *Lingulodinium* CSPs were cloned and expressed as a fusion protein with a C-terminal GST tag as described above. The BL21 cell lysates were centrifuged and the supernatants containing GST tagged CSPs were incubated with Glutathione Sepharose 4B beads (Promega) for 45 min at room temperature with end-over-end agitation. Beads were washed 4 times in PBS. Immobilized *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* were tested for sequence-specific DNA binding activities against a set of degenerate oligonucleotides using a selection and amplification binding assay (SAAB) [35, 36]. A set of single-stranded oligonucleotides with PCR primer sequences flanking nine random nucleotides (N9) were synthesized and used to produce double-stranded DNA by a single PCR cycle using the reverse primer. Fifteen μ g of double-stranded DNA (N9) was allowed to bind to 10 μ L of immobilized CSPs in a 100 μ L total volume solution containing 75 mM NaCl, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride, 0.1% Triton X-100, 10 ng of poly (dI-dC) per μ L, 10 mM Tris-HCl (pH 7), 6% glycerol and 1% BSA. After 1 h of agitation at 4 °C, the supernatant containing unbound oligonucleotides were removed. Following 3 times of washing with binding buffer, DNA was released from the protein by boiling in water [35]. DNA was amplified in a PCR reaction to repeat the protein binding step. Three rounds of SAAB were performed before sending out the PCR products for sequencing (Fig. 6).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12860-021-00368-4>.

Additional file 1: Supplemental Figure S1. Purification of *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3*. A shows recombinant *LpCSP1*-GST, *SkCSP1*-GST, *SkCSP2*-GST and *SkCSP3*-GST analyzed on an 18% acrylamide SDS-PAGE gel after affinity purification. B shows *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* after removal of the GST tag by thrombin digestion and binding to glutathione-Sepharose 4B beads. The sizes of the molecular weight markers (left) are shown in kilodaltons.

Additional file 2: Supplemental Table S1. List of primers used for PCR amplification and cloning of *LpCSP1*, *SkCSP1*, *SkCSP2* and *SkCSP3* sequences in pGEX4T2 plasmid.

Additional file 3: Supplementary Table S2. List of proteins selected for phylogenetic reconstruction.

Authors' contributions

BZ designed and performed experiments, analysed data, and participated in writing the manuscript. DM analysed data, participated in writing the manuscript and obtained funding. The author(s) read and approved the final manuscript.

Funding

Funding was obtained from the Canadian National science and Engineering Research Council (grant number 171382-03 to D.M.). The funding body played no role in the design of the study and collection, analysis, and interpretation of data, or in writing the manuscript and decision to publish.

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 23 March 2021 Accepted: 21 April 2021

Published online: 08 May 2021

References

- Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*. 2000;290(5499):2105–10. <https://doi.org/10.1126/science.290.5499.2105>.
- Zhang HM, Chen H, Liu W, Liu H, Gong J, Wang H, et al. AnimalTFDB: a comprehensive animal transcription factor database. *Nucleic Acids Res*. 2012;40(Database issue):D144–9. <https://doi.org/10.1093/nar/gkr965>.
- Beauchemin M, Roy S, Daoust P, Dagenais-Bellefeuille S, Bertomeu T, Letourneau L, et al. Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic. *Proc Natl Acad Sci U S A*. 2012;109(39):15793–8. <https://doi.org/10.1073/pnas.1206683109>.
- Bayer T, Aranda M, Sunagawa S, Yum LK, Desalvo MK, Lindquist E, et al. Symbiodinium transcriptomes: genome insights into the dinoflagellate symbionts of reef-building corals. *PLoS One*. 2012;7(4):e35269. <https://doi.org/10.1371/journal.pone.0035269>.
- Beauchemin M, Roy S, Pelletier S, Averback A, Morse D. Characterization of two dinoflagellate cold shock domain proteins. *mSphere*. 2016;1:e00034–15.
- Mihailovich M, Militti C, Gabaldon T, Gebauer F. Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. *Bioessays*. 2010;32(2):109–18. <https://doi.org/10.1002/bies.200900122>.
- Budkina KS, Zlobin NE, Kononova SV, Ovchinnikov LP, Babakov AV. Cold shock domain proteins: structure and interaction with nucleic acids. *Biochemistry (Mosc)*. 2020;85(Suppl 1):S1–S19. <https://doi.org/10.1134/S0006297920140011>.
- Bae W, Xia B, Inouye M, Severinov K. Escherichia coli CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci U S A*. 2000;97(14):7784–9. <https://doi.org/10.1073/pnas.97.14.7784>.
- Roy S, Letourneau L, Morse D. Cold-induced cysts of the photosynthetic dinoflagellate *Lingulodinium polyedrum* have an arrested circadian bioluminescence rhythm and lower levels of protein phosphorylation. *Plant Physiol*. 2014;164(2):966–77. <https://doi.org/10.1104/pp.113.229856>.
- Li T, Yu L, Song B, Song Y, Li L, Lin X, et al. Genome improvement and core gene set refinement of fugacium kawagutii. *Microorganisms*. 2020;8(1):102. <https://doi.org/10.3390/microorganisms8010102>.
- Jones PG, VanBogelen RA, Neidhardt FC. Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol*. 1987;169(5):2092–5. <https://doi.org/10.1128/JB.169.5.2092-2095.1987>.
- Heinemann U, Roske Y. Cold-shock domains-abundance, structure, properties, and nucleic-acid binding. *Cancers*. 2021;13(2):190. <https://doi.org/10.3390/cancers13020190>.
- Graumann PL, Marahiel MA. A superfamily of proteins that contain the cold-shock domain. *Trends Biochem Sci*. 1998;23(8):286–90. [https://doi.org/10.1016/S0968-0004\(98\)01255-9](https://doi.org/10.1016/S0968-0004(98)01255-9).
- Jones PG, Krah R, Tafuri SR, Wolffe AP. DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. *J Bacteriol*. 1992;174(18):5798–802. <https://doi.org/10.1128/JB.174.18.5798-5802.1992>.
- Karlson D, Imai R. Conservation of the cold shock domain protein family in plants. *Plant Physiol*. 2003;131(1):12–5. <https://doi.org/10.1104/pp.014472>.
- Nakaminami K, Karlson DT, Imai R. Functional conservation of cold shock domains in bacteria and higher plants. *Proc Natl Acad Sci U S A*. 2006;103(26):10122–7. <https://doi.org/10.1073/pnas.0603168103>.
- Wistow G. Cold shock and DNA binding. *Nature*. 1990;344(6269):823–4. <https://doi.org/10.1038/344823c0>.
- Fusaro AF, Bocca SN, Ramos RL, Barroco RM, Magioli C, Jorge VC, et al. AtGRP2, a cold-induced nucleo-cytoplasmic RNA-binding protein, has a role in flower and seed development. *Planta*. 2007;225(6):1339–51. <https://doi.org/10.1007/s00425-006-0444-4>.
- Kim MH, Sonoda Y, Sasaki K, Kaminaka H, Imai R. Interactome analysis reveals versatile functions of Arabidopsis COLD SHOCK DOMAIN PROTEIN 3 in RNA processing within the nucleus and cytoplasm. *Cell Stress Chaperones*. 2013;18(4):517–25. <https://doi.org/10.1007/s12192-012-0398-3>.
- Izumi H, Imamura T, Nagatani G, Ise T, Murakami T, Uramoto H, et al. Y-box-binding protein-1 binds preferentially to single-stranded nucleic acids and exhibits 3'→5' exonuclease activity. *Nucleic Acids Res*. 2001;29(5):1200–7. <https://doi.org/10.1093/nar/29.5.1200>.
- Lasham A, Moloney S, Hale T, Homer C, Zhang YF, Murison JG, et al. The Y-box-binding protein, YB1, is a potential negative regulator of the p53 tumor suppressor. *J Biol Chem*. 2003;278(37):35516–23. <https://doi.org/10.1074/jbc.M303920200>.
- Sommerville J. Activities of cold-shock domain proteins in translation control. *Bioessays*. 1999;21(4):319–25. [https://doi.org/10.1002/\(SICI\)1521-1878\(199904\)21:4<319::AID-BIES8>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1521-1878(199904)21:4<319::AID-BIES8>3.0.CO;2-3).
- Kleene KC. Y-box proteins combine versatile cold shock domains and arginine-rich motifs (ARMs) for pleiotropic functions in RNA biology. *Biochem J*. 2018;475(17):2769–84. <https://doi.org/10.1042/BCJ20170956>.
- Mordovkina D, Lyabin DN, Smolin EA, Sogorina EM, Ovchinnikov LP, Eliseeva I. Y-box binding proteins in mrnp assembly, translation, and stability control. *Biomolecules*. 2020;10(4):591. <https://doi.org/10.3390/biom10040591>.
- Sangermano F, Delicato A, Calabro V. Y box binding protein 1 (YB-1) oncoprotein at the hub of DNA proliferation, damage and cancer progression. *Biochimie*. 2020;179:205–16. <https://doi.org/10.1016/j.biochi.2020.10.004>.
- MacDonald GH, Itoh-Lindstrom Y, Ting JP. The transcriptional regulatory protein, YB-1, promotes single-stranded regions in the DRB promoter. *J Biol Chem*. 1995;270(8):3527–33. <https://doi.org/10.1074/jbc.270.8.3527>.
- Wei WJ, Mu SR, Heiner M, Fu X, Cao LJ, Gong XF, et al. YB-1 binds to CAUC motifs and stimulates exon inclusion by enhancing the recruitment of U2AF to weak polypyrimidine tracts. *Nucleic Acids Res*. 2012;40(17):8622–36. <https://doi.org/10.1093/nar/gks579>.
- Yang XJ, Zhu H, Mu SR, Wei WJ, Yuan X, Wang M, et al. Crystal structure of a Y-box binding protein 1 (YB-1)-RNA complex reveals key features and residues interacting with RNA. *J Biol Chem*. 2019;294(28):10998–1010. <https://doi.org/10.1074/jbc.RA119.007545>.
- Yu L, Li T, Li L, Lin X, Li H, Liu C, et al. SAGER: a database of Symbiodiniaceae and algal genomic resource. *Database (Oxford)*. 2020;2020. <https://doi.org/10.1093/database/baaa051>.
- Wang Y, Jensen L, Hojrup P, Morse D. Synthesis and degradation of dinoflagellate plastid-encoded psbA proteins are light-regulated, not circadian-regulated. *Proc Natl Acad Sci U S A*. 2005;102(8):2844–9. <https://doi.org/10.1073/pnas.0406522102>.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res*. 2008;36(Web Server issue):W465–9.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12):1647–9. <https://doi.org/10.1093/bioinformatics/bts199>.
- Cock PJ, Chilton JM, Gruning B, Johnson JE, Soranzo N. NCBI BLAST+ integrated into galaxy. *Gigascience*. 2015;4(1):39. <https://doi.org/10.1186/s13742-015-0080-7>.
- Xia B, Ke H, Inouye M. Acquisition of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in *Escherichia coli*. *Mol Microbiol*. 2001;40(1):179–88. <https://doi.org/10.1046/j.1365-2958.2001.02372.x>.
- Chang C, Jacobs Y, Nakamura T, Jenkins NA, Copeland NG, Cleary ML. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric pbx proteins. *Mol Cell Biol*. 1997;17(10):5679–87.
- Magnani E, Sjolander K, Hake S. From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell*. 2004;16(9):2265–77. <https://doi.org/10.1105/tpc.104.023135>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.