



# Characterization of Two Dinoflagellate Cold Shock Domain Proteins

Mathieu Beauchemin, Sougata Roy,\* Sarah Pelletier, Alexandra Averback, Frederic Lanthier, David Morse

Institut de Recherche en Biologie Végétale and Département de Sciences Biologiques, Université de Montréal, Montreal, Quebec, Canada

\* Present address: Sougata Roy, EMBL-European Bioinformatics Institute and Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom.

M.B. and S.R. contributed equally to this work.

**ABSTRACT** Roughly two-thirds of the proteins annotated as transcription factors in dinoflagellate transcriptomes are cold shock domain-containing proteins (CSPs), an uncommon condition in eukaryotic organisms. However, no functional analysis has ever been reported for a dinoflagellate CSP, and so it is not known if they do in fact act as transcription factors. We describe here some of the properties of two CSPs from the dinoflagellate *Lingulodinium polyedrum*, *LpCSP1* and *LpCSP2*, which contain a glycine-rich C-terminal domain and an N-terminal cold shock domain phylogenetically related to those in bacteria. However, neither of the two *LpCSPs* act like the bacterial CSP, since they do not functionally complement the *Escherichia coli* quadruple cold shock domain protein mutant BX04, and cold shock does not induce *LpCSP1* and *LpCSP2* to detectable levels, based on two-dimensional gel electrophoresis. Both CSPs bind to RNA and single-stranded DNA in a nonspecific manner in electrophoretic mobility shift assays, and both proteins also bind double-stranded DNA nonspecifically, albeit more weakly. These CSPs are thus unlikely to act alone as sequence-specific transcription factors.

**IMPORTANCE** Dinoflagellate transcriptomes contain cold shock domain proteins as the major component of the proteins annotated as transcription factors. We show here that the major family of cold shock domain proteins in the dinoflagellate *Lingulodinium* do not bind specific sequences, suggesting that transcriptional control is not a predominant mechanism for regulating gene expression in this group of protists.

**KEYWORDS:** RNA binding domain, DNA binding domain, cold shock domain, dinoflagellates, cold shock protein, transcription

Cold shock domains (CSD) are an ancient and conserved nucleic acid binding module (1). They are small, roughly 70 amino acids in length, contain 2 amino acid motifs that are shared by RNA recognition motif domains, and have been shown to bind both DNA (2, 3) and RNA (4). They are found in some archaea, in eubacteria, plants, animals and some fungal lineages. However, the roles played by the CSD depend on the organism and on specific domains associated with it. In bacteria, the cold shock response includes a transient and global translation arrest (5), during which time the cells synthesize a small number of cold-inducible proteins and begin reprogramming their translational machinery to accommodate growth at lower temperatures (6). Cold shock proteins (CSPs), which in bacteria consist only of a CSD, are among the few proteins whose synthesis increases after a 15°C cold shock (7). Many of these bacterial CSPs appear to be functionally redundant, as it is necessary to delete four of the nine *Escherichia coli* *csp* genes before growth at cold temperatures is impaired (8).

Received 28 October 2015 Accepted 5 December 2015 Published 13 January 2016

**Citation** Beauchemin M, Roy S, Pelletier S, Averback A, Lanthier F, Morse D. 2016. Characterization of two dinoflagellate cold shock domain proteins. *mSphere* 1(1):e00034-15. doi: 10.1128/mSphere.00034-15.

**Editor** Aaron P. Mitchell, Carnegie Mellon University

**Copyright** © 2016 Beauchemin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to David Morse, david.morse@umontreal.ca.

The nucleic acid binding properties of CSPs are an integral part of the bacterial cold shock response. For example, *E. coli* CspA stimulates transcription of the DNA gyrase *gyrA* (9), an effect attributable to binding of specific DNA sequences in the *gyrA* promoter (3). In addition, CspA can aid transcription by acting as an antiterminator, an effect suggested to rely on CspA binding to single-stranded regions of the newly synthesized RNA (10). CSP binding to RNA has also been proposed to melt RNA secondary structures that form due to reduced temperature (11). This may improve translation by eliminating inhibitory secondary structures and may also relieve transcriptional stalling due to the formation of stem-loop secondary structures in the nascent RNA. Lastly, CspA can bind its own transcript and protect it from degradation, as the half-life of the message increases roughly 100-fold at 15°C (12). Thus, in *E. coli*, most of the effects produced by the CSP appear to rely on a capacity for binding RNA.

Plant CSPs also contain a CSD, but in addition include a C-terminal glycine-rich domain interspersed with a variable number of CCHC-type zinc finger DNA binding domains (13, 14). The model plant *Arabidopsis thaliana* has four CSD-containing proteins, some of which are upregulated by exposure to cold (13). *A. thaliana* overexpressing selected CSPs showed increased tolerance to freezing (15) and drought (16) through an mRNA chaperone activity (17). Some (but not all) *Arabidopsis* CSPs also complement the cold-sensitive *E. coli* quadruple *csp* mutant (18). However, these CSPs seem to play a broader role as regulators of embryo development, seed germination, and flowering (19, 20). A similar function may be performed by two CSPs from *Oriza sativa*, as the highest expression is found in flowers and seeds (21).

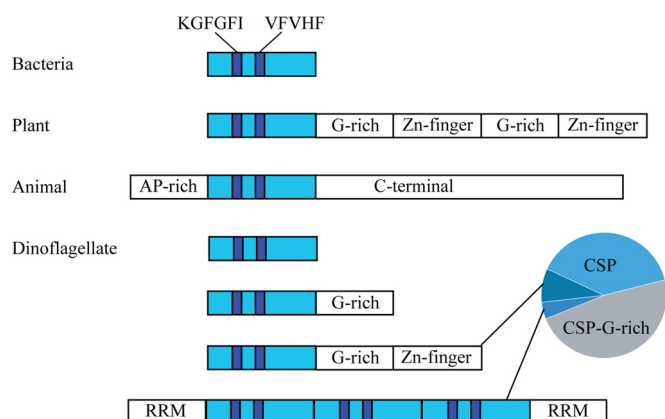
The core CSD in animal CSPs is surrounded by an N-terminal alanine-proline-rich domain and a C-terminal domain with alternating acidic and basic regions. The first example, a protein called YB-1, was identified due to its ability to bind the Y-box (CTGATTGGCCAA) in promoters of major histocompatibility locus genes (2), yet was subsequently revealed to have potent mRNA-stabilizing activity (22). Furthermore, YB-1 has also been shown to regulate translation of specific mRNAs (23), while the frog Y-box protein (FRGY2) was found in ribonucleoproteins particles (RNPs) that sequester maternal mRNAs in oocytes (24).

Recent transcriptomic analyses of dinoflagellates have shown an abundance of CSD proteins (25, 26). These domains are classified as DNA binding by gene ontology (GO) categories, and in dinoflagellate transcriptomes they constitute roughly two-thirds of all the potential transcription factors identified. However, given the many examples of RNA binding by CSPs, it is not clear if the transcription factor classification of dinoflagellate CSPs is accurate. To assess the function of these dinoflagellate proteins, we have expressed and purified two dinoflagellate CSPs, which we named *LpCSP1* and *LpCSP2*, and compared their nucleic acid binding properties with those of the bacterial CSPs to which they are most closely related. Curiously, while able to bind both RNA and DNA in electrophoretic mobility assays (EMSAs), the dinoflagellate proteins do not complement a bacterial CSP mutant and are not induced under cold conditions.

## RESULTS

### ***LpCSP1* and *LpCSP2* are part of a distinct clade in the eukaryotic CSP family.**

Roughly two-thirds of dinoflagellate proteins classified as transcription factors by gene ontology (GO) are cold shock domain (CSD)-containing proteins (CSPs) (25, 26). The overrepresentation of this class of protein thus suggests they might play an important role in the regulation of gene expression. To begin to address the role of these CSPs, we first determined the types of CSPs expressed by dinoflagellates. We used the CSD in BLAST searches of the transcriptomes of *Lingulodinium* and *Symbiodinium* as well as ESTs from *Alexandrium tamarense* and *Karenia brevis*. We recovered a total of 23 full-length sequences as defined by the presence of a single ORF with in-frame start and stop codons encompassing the CSD. All dinoflagellate CSDs contain the two characteristic RNA binding motifs (KGFGLI and VFVHF) (Fig. 1). However, the CSD itself is found in at least four different domain architectures. The vast majority of the sequences recovered contained a CSD either alone or with a C-terminal G-rich domain (Fig. 1).



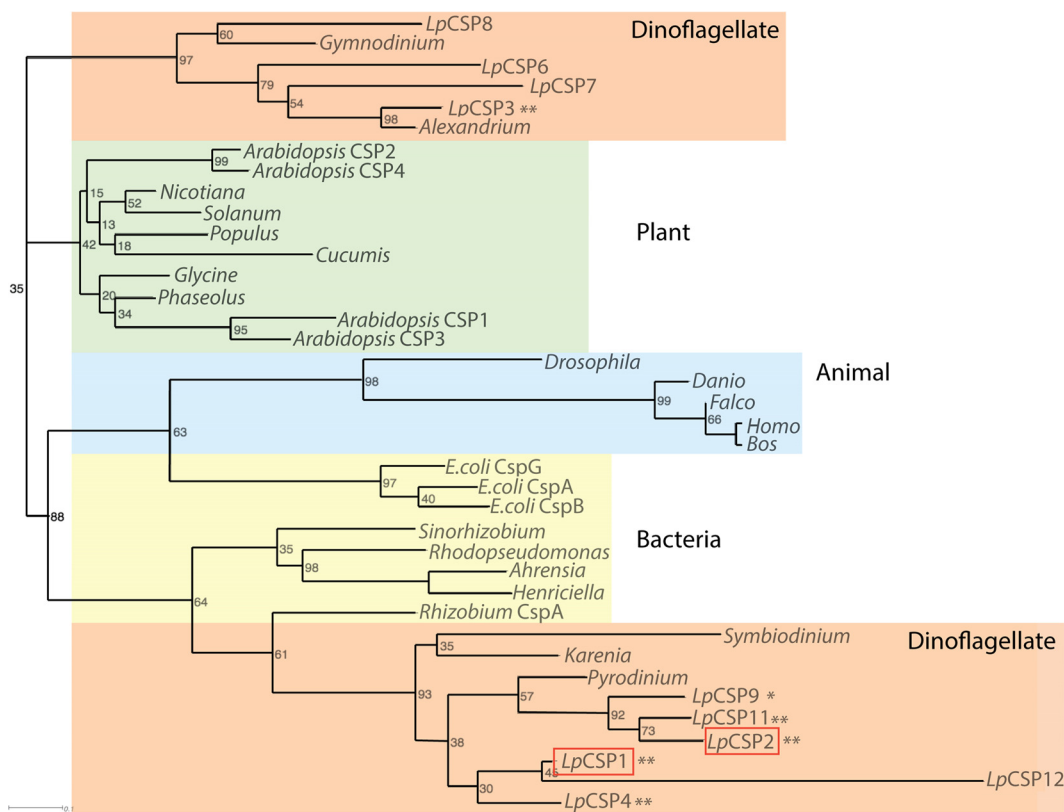
**FIG 1** CSPs contain a conserved CSD. The ~70-amino-acid CSD (light blue) is highly conserved in bacteria, plants, animals, and dinoflagellates and contains two RNA recognition motifs (dark blue; KGFGFI and VFVHF). Bacterial CSPs consist only of a CSD, while the other three classes contain C-terminal extensions. The C terminal of plant CSPs typically contains two to seven repeats of a glycine-rich (G-rich) Zn-finger region, a pattern observed in a few dinoflagellate proteins. More typically, when present, the C-terminal extensions of dinoflagellates are G-rich only. An atypical CSD-containing protein found in *Lingulodinium* has three CSD repeats flanked by an RNA recognition module (RRM). Finally, animal CSDs are unique in that they contain an AP-rich N-terminal extension. The pie chart shows the relative abundance of the four different dinoflagellate CSP architectures in the transcriptomes of *Lingulodinium* and *Symbiodinium* species and in the expressed sequence tags of *Alexandrium tamarense* and *Karenia brevis*.

Smaller numbers of representatives contained a Zn-finger domain following the G-rich domain, and even fewer examples were found of sequences with multiple CSDs and one or more RNA recognition motifs (RRM) (Fig. 1; the pie chart shows distribution of sequence numbers). The domain structure of most dinoflagellate CSPs is thus closest to that found in bacteria or in plants, although there is no sequence similarity between the G-rich domains of plant (27) and dinoflagellate CSPs.

As a complement to the domain structure analysis, a phylogenetic analysis of the CSDs from 10 *Lingulodinium* sequences was also carried out using sequences from a wide array of organisms. In general, support for the different clades was poor, as the CSD is short and the sequence quite conserved. However, there is strong support for grouping the dinoflagellate sequences into two different clades (Fig. 2). One of these two clades also contains both animal and bacterial CSPs, and since the bacterial CSPs are among the best characterized, we elected to begin our analysis with *Lingulodinium* representatives from this clade.

**LpCSP1 and LpCSP2 are not functional complements of *E. coli* CSPs.** Two dinoflagellate CSPs, which we have termed *LpCSP1* and *LpCSP2*, were cloned and expressed. Both are small proteins (113 and 110 residues, respectively) whose domain structure is similar to that found in the largest class of CSD proteins (an N-terminal CSD followed by a glycine-rich domain). In both, the CSD contains the two expected RNA binding motifs (Fig. 1; see also Fig. S1 in the supplemental material). Bacterial CSPs are required for cell growth at low temperature (28), and a bacterial strain harboring a mutation in four different CSP genes is unable to grow at 17°C (8). However, while overexpression of the bacterial *CspA* gene into the quadruple mutant allowed for growth at lower temperatures (Fig. 3), neither an empty vector (pINIII) nor the *Lingulodinium* CSPs in pINIII were able to fulfill this role. These dinoflagellate CSPs thus differ from the bacterial CSPs.

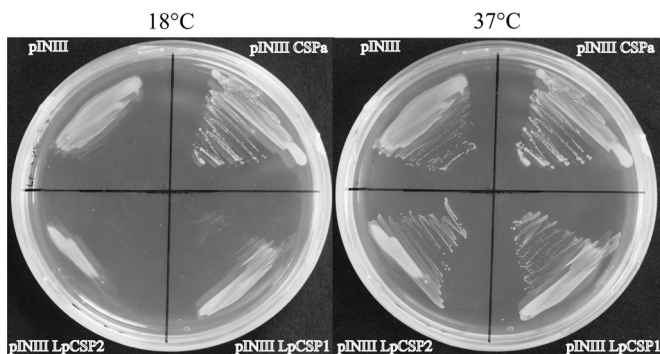
**LpCSP1 and LpCSP2 expression are not detectable after cold shock.** Since bacterial CSPs are strongly induced by cold temperature, we also tested if *LpCSP1* and *LpCSP2* induction could be detected by two-dimensional (2D) PAGE, as seen in *E. coli* (29). The encystment of *Lingulodinium* as a cellular response to cold temperatures (8°C) has previously been reported (30), and we therefore compared the protein profile of cells grown at normal culture room temperatures (Fig. 4A) with those of cysts at 8°C



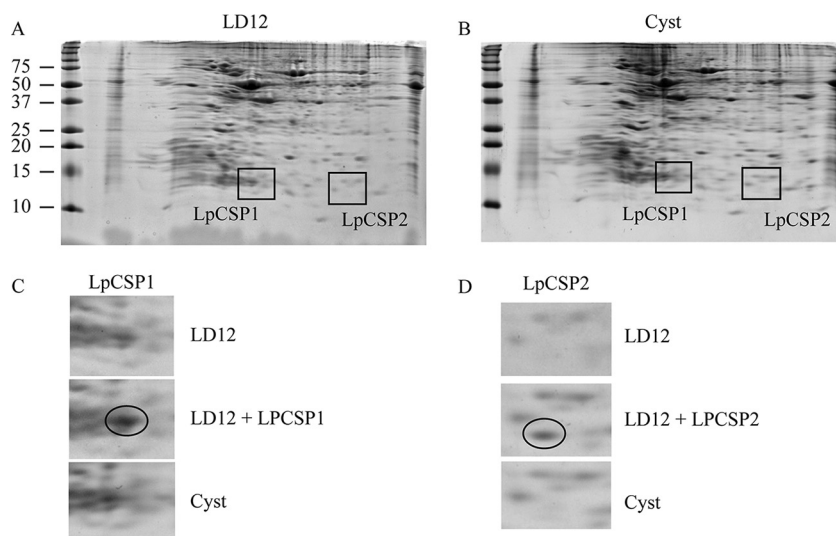
**FIG 2** Phylogenetic reconstructions place dinoflagellate CSPs in two distinct clades, one clustered with bacterial and animal sequences and the second in a separate clade unique to dinoflagellates. The *Lingulodinium* CSPs analyzed here are boxed. Sequences marked with a single asterisk contain only a CSD, while those marked with two asterisks contain a CSD and a G-rich domain. The accession numbers for the sequences used are listed in Materials and Methods.

(Fig. 4B). To determine the expected position of *LpCSP1* and *LpCSP2*, 150 µg of *Lingulodinium* protein sample was spiked with 150 ng of purified *LpCSP1* or *LpCSP2* before electrophoresis (Fig. 4C and D). No detectable induction of *Lingulodinium* CSPs was observed after 24 h at 8°C.

**LpCSPs bind nucleic acids in vitro.** Binding of DNA and RNA was evaluated for *LpCSP1* and *LpCSP2* via EMSA. Both proteins were expressed as glutathione *S*-transferase (GST) fusions and purified, and the GST domain was removed by thrombin digestion prior to use. However, since some residual GST remained in the purified protein fraction (see Fig. S2 in the supplemental material), GST alone was also tested for



**FIG 3** Dinoflagellate CSPs do not complement a quadruple bacterial CSP mutant. Plates were streaked with a quadruple CSP mutant strain BX04 transformed with an empty vector (pINIII), an *E. coli* CspA, *LpCSP1*, or *LpCSP2* and grown first at 18°C for 5 days (left) and then transferred to 37°C overnight (right).



**FIG 4** Cold treatment does not induce detectable levels of *LpCSP1* and *LpCSP2*. 2D-PAGE analysis results are shown for *Lingulodinium* total proteins from cells grown at 18°C and collected at LD12 (A) and from cysts harvested after 24 h at 8°C (B). (C and D) The boxed areas in panels A and B correspond to the position of *LpCSP1* (C) and *LpCSP2* (D) are shown from both gels and from a gel (middle panel) in which 150 ng of recombinant protein was added before isoelectric focusing (circled area). Molecular masses in kilodaltons were estimated using Precision Plus protein standards (Bio-Rad).

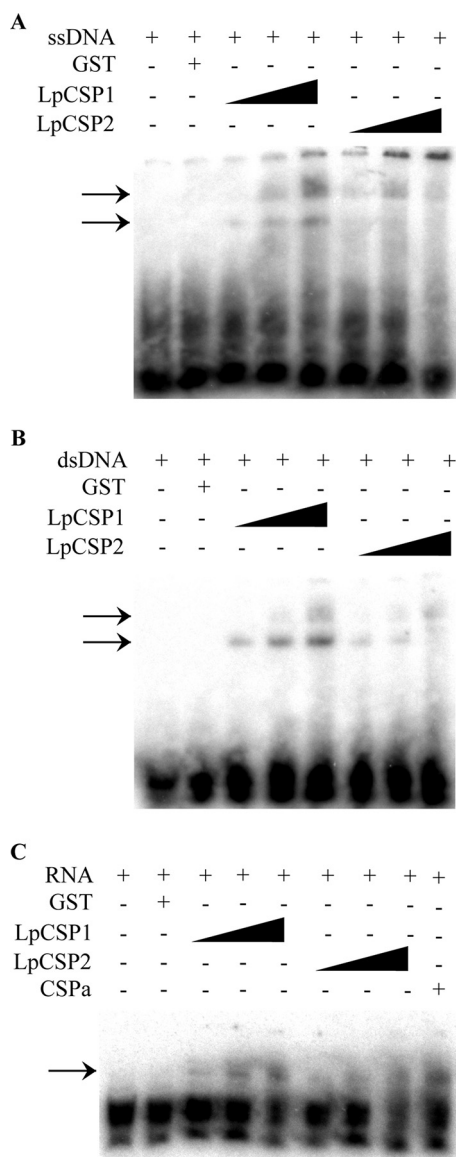
binding under the same conditions. When single-stranded DNA (ssDNA) was used as a probe, both *LpCSP1* and *LpCSP2* showed the concentration-dependent slower-migrating bands on EMSA, indicative of nucleic acid binding (Fig. 5A). However, a considerable amount of radiolabeled DNA remained in the loading wells, suggestive of binding by a multimeric CSP to several probe molecules. Both proteins were also able to bind double-stranded DNA (dsDNA), although in this case all bound radiolabel was observed to migrate into the gel (Fig. 5B). Lastly, binding to radiolabeled RNA was also detected for both proteins, with binding similar to that observed for an authentic *E. coli* CSP (Fig. 5C). Both the two *Lingulodinium* proteins thus displayed a broad nucleic acid binding capacity.

To assess the possibility that *LpCSP1* might display a capacity for sequence-specific binding, different competitors were added to *LpCSP1* in the presence of an ssDNA probe (see Fig. S3A in the supplemental material). These competition experiments indicated that dsDNA competes poorly for binding to the ssDNA and that altering the sequence of the competing ssDNA does not affect its ability to compete. In contrast, ssDNA is an effective competitor of the binding between *LpCSP1* and a dsDNA probe (see Fig. S3B). *LpCSP1* thus prefers ssDNA substrates.

## DISCUSSION

In this report, we have assessed the role of two dinoflagellate CSPs. CSPs are potentially important in dinoflagellates, since high-throughput transcriptome studies of *Lingulodinium* (25) and *Symbiodinium* (26) have shown that the majority of the proteins annotated as transcription factors have CSDs. Despite this annotation, however, it is not clear whether any of the dinoflagellate CSPs actually play a role in transcription. To date, with the sole exception of a TATA binding protein (TBP)-like protein, which replaces the TBP usually found in eukaryotic TFIID complexes (31), no transcription factor has been described and characterized experimentally in dinoflagellates.

Two types of CSPs are likely predominant in dinoflagellates, based on the protein domain structure. One form contains only the CSD, while the other contains the CSD and a C-terminal G-rich domain (Fig. 1). Although this analysis is based on the frequency of finding different types of sequence within a transcriptome, it seems likely



**FIG 5** *LpCSP1* and *LpCSP2* show similar nucleic acid binding activity in EMSA. The EMSAs were carried out using an ssDNA probe (A), a dsDNA probe (B), and an RNA probe (C). The black triangle above the autoradiograms denotes the different concentrations of the CSP recombinant protein used for the assays (100, 300, and 1,000 ng in all the assays; 1,000 ng of CSPa in panel C). Arrows show the bands corresponding to shifted nucleic acids.

that the more-frequently found forms would result in a greater amount of protein. The cellular *Lingulodinium* CSP pool may thus be dominated by these two forms, although we cannot rule out the possibility that a protein with a different domain structure could be highly expressed.

Molecular phylogeny of the *L. polyedrum* CSD sequences indicates they are found in two well-supported clades (Fig. 2), perhaps indicative of a functional diversity among dinoflagellate CSPs. We have examined the properties of two members of the group most closely related to bacterial CSPs, *LpCSP1* and *LpCSP2* (Fig. 2). Despite this relationship with bacterial sequences, however, neither protein appears functionally equivalent to those in bacteria. This was most clearly shown through its inability to complement the growth of the mutant *E. coli* BX04 strain at low temperatures (Fig. 3). The nonspecific mRNA binding activity of bacterial CSPs is the key to their chaperone activity during cold stress (1), so the lack of complementation by *LpCSP1* and *LpCSP2*

is puzzling, given that both dinoflagellate proteins are able to bind an RNA probe *in vitro* (Fig. 5C). However, comparison of two similar *Arabidopsis* CSPs, each containing an N-terminal CSD and a C-terminal G-rich domain, showed that only one was able to complement the BX04 mutant (18). It is thus possible that small differences between sequences are sufficient to disrupt the RNA chaperone activity, which in turn suggests that other dinoflagellate CSD-containing proteins could show RNA chaperone activity and potentially rescue the BX04 strain. However, it is also possible that the lack of complementation in the bacterial assay may reflect the amount of the *Lingulodinium* CSPs actually expressed in the bacteria or an element of sequence specificity that remained undetected in our *in vitro* assays.

We find no evidence that *LpCSP1* and *LpCSP2* are induced by exposure to cold temperatures. Transcription of both *Arabidopsis* and *E. coli* CSPs can be induced during cold shock (18, 32), but a recent comparison of the transcriptomes of normally growing and cold-shocked *L. polyedrum* cells showed no induction of CSP transcripts by the cold treatment (30). In addition, exposure to cold induces CSP at the protein level in both prokaryotes and eukaryotes, in keeping with their ability to prevent cells from freezing, but a global analysis of the proteome of cold-shocked *L. polyedrum* showed no difference, compared to normally growing cells (30). This agrees with the 2D gel analysis results shown here, where we specifically looked for *LpCSP1* and *LpCSP2* induction at low temperatures (Fig. 4). Thus, unlike the bacterial CSPs (7, 29), there is no induction of *LpCSP1* or *LpCSP2* mRNAs, nor are more proteins synthesized, although we cannot rule out a potential posttranslational modification that could modify the position of CSPs spots on a 2D gel. Taken together, a role in cold tolerance seems unlikely for these dinoflagellate CSPs.

Unlike the bacterial CSPs, whose cellular role seems primarily to block formation of extensive RNA secondary structure at cold temperatures (10, 33), CSPs in vertebrates and plants display a number of different roles. The mammalian Y-box binding protein YB1, which shares 40% amino acid identity with the bacterial CSPs (34, 35), has diverse physiological roles apart from the cold stress response (14). YB1 is known to regulate transcription (36, 37) by binding to duplex DNA containing a Y-box (CTGATTGGCT) (2). Similarly, the frog Y-box protein FRGY1 has been shown to stimulate transcription from a promoter containing a Y-box (38). YB-1 has been shown to bind ssDNA with greater affinity than dsDNA (39), and the resulting destabilization of the DNA double helix has been proposed as the transcriptional activation mechanism. However, *LpCSP1* and *LpCSP2* bind a Y-box dsDNA at very high protein concentrations (300 and 1,000 ng) (Fig. 5B) compared to binding of Y-box DNA sequence by YB1, for which only 30 ng is sufficient (40). The physiological levels of the *LpCSP1* and *LpCSP2* proteins appear much lower than this, as indicated by our 2D gel profiles of normally growing cells (Fig. 4). This binding also does not appear to be sequence specific (see Fig. S3 in the supplemental material), indicating that specific targeting of the Y-box sequence is doubtful. However, it must be noted that we have not yet tested if other DNA sequences might show preferential binding enabling lower concentrations of *LpCSPs* to be effective. The stronger binding to ssDNA leaves the door open to a potential role for *LpCSPs* as destabilizers of the DNA double helix, potentially in cooperation with other, more specific transcription factors.

Plant CSD-containing proteins also have a broader role than a simple response to cold shock. The precise molecular mechanism that enables plant CSPs to regulate processes other than acclimation to abiotic stress, such as seed and flower development, is still unknown. However, the interactions of the *A. thaliana* CSP3 with diverse proteins, such as poly(A) binding proteins, ribosomal proteins, and mRNA-decapping protein, suggest an involvement in multiple RNA processing steps (41). A similar role might be envisioned for *LpCSP1* and *LpCSP2*, as they are able to bind RNA (Fig. 5C). Furthermore, the nonspecific RNA binding by *LpCSPs* is consistent with a role in mRNA packaging and stability (42), an intriguing prospect in dinoflagellates, where particularly long mRNA half-lives have been documented (43).

Taken together, our results are most consistent with the idea that *LpCSP1* and

*LpCSP2* are not sequence-specific transcription factors, contrary to the impression left by the Gene Ontology assignment. This further reduces the already-scarce number of potential transcription factors in dinoflagellates, a fact that fits well with the limited scale of transcriptional variation seen during the circadian cycle (44), as well as in response to nutrient limitation (45, 46) or abiotic stress (30). We suggest that *LpCSPs* may still be able to participate in the transcription process, potentially by unwinding the DNA helix due to their capacity to bind ssDNA. However, further studies, such as protein localization and identification of binding partners, will be required to define more precisely the role of these proteins in dinoflagellates. It will also be of interest to examine members of the second dinoflagellate clade, in case these CSPs have substantially different properties.

## MATERIALS AND METHODS

**Cell culture.** Unialgal but not axenic cultures of *Lingulodinium polyedrum* (formerly *Gonyaulax polyedra*; strain CCMP1936) were obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (Boothbay Harbor, ME) and grown in modified seawater medium (f/2) (47) at constant temperature ( $18 \pm 1^\circ\text{C}$ ). The culture room's light cycle was 12 h with cool white fluorescent light at an intensity of  $50 \mu\text{mol photons m}^{-2} \text{s}^{-2}$ , followed by 12 h dark. Under these conditions, the beginning of the light period is termed LD 0 and the beginning of the dark period LD 12. Cultures were typically grown to a cell density of 12,000 to 14,000 cells/ml before cell collection by filtration on Whatman 541 paper supported by a Buchner funnel. Cysts were obtained by placing the cultures at  $8^\circ\text{C}$  for 24 hours as described previously (30). All cells were either used immediately or frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  until further use.

**Sequence alignment and phylogenetic analyses.** The CSP sequences used for sequence alignment and phylogenetic analyses were obtained from the *Lingulodinium* transcriptome deposited at NCBI. A search of the *Lingulodinium* transcriptome deposited at The Marine Microbial Eukaryote Transcriptome Sequencing Project (48) was also done to retrieve complete sequences when needed. Trees were constructed by using an online tool available at <http://www.phylogeny.fr> (49). In the workflow, multiple sequences were aligned using MUSCLE and curated using Gblocks. Phylogenetic reconstructions were made using PhyML and visualized using TreeDyn. Sequences used for phylogenetic reconstructions were from the genera (or proteins) *Ahrensia* (WP\_018687722), *Alexandrium* (GAIT01073544), *Arabidopsis* CSP1 (AEE86603), *Arabidopsis* CSP2 (NP\_195580), *Arabidopsis* CSP3 (NP\_565427), *Arabidopsis* CSP4 (Q38896), *Bos* (DAA26237), *Cucumis* (XP\_004140332), *Danio* (XP\_001340141), *Drosophila* (NP\_647983), *E. coli* CspA (AAN82813), *E. coli* CspB (AAB61739), *E. coli* CspG (NP\_309172), *Falco* (XP\_005244100), *Glycine* (XP\_003540832), *Gymnodinium* (GAIL01018775), *Henricella* (WP\_018146825), *Homo* (EAW48433), *Karenia* (FK848095), *Lingulodinium* 1 (JO733348), *Lingulodinium* 2 (JO729000 and CAMNT\_0033837443), *Lingulodinium* 3 (JO730956 and CAMNT\_0033776061), *Lingulodinium* 4 (JO734870 and CAMNT\_0033828465), *Lingulodinium* 6 (JO720996 and CAMNT\_0033712139), *Lingulodinium* 7 (JO766444 and CAMNT\_0033724295), *Lingulodinium* 8 (JO761018 and CAMNT\_0033635737), *Lingulodinium* 9 (JO730992 and CAMNT\_0033829387), *Lingulodinium* 11 (JO736519), *Lingulodinium* 12 (JO732587), *Nicotiana* (P27484), *Pyrodictum* (GAI001020278), *Rhizobium* (YP\_770349), *Phaseolus* (ESW08176), *Populus* (XP\_002313723), *Rhodospseudomonas* (NP\_948738), *Sinorhizobium* (AAC64672), *Solanum* (XP\_006359670), and *Symbiodinium* 1 (GAF001002801).

**CSP cloning, expression, and purification.** Primers designed from the *Lingulodinium* transcriptome sequences JO733348 and JO729000 (see Table S1 in the supplemental material) (25) were used to amplify what we termed *LpCSP1* and *LpCSP2*, respectively, from a first-strand cDNA reaction product prepared from total RNA extracted from *L. polyedrum* cells by using Trizol (Invitrogen). The reverse transcription reaction was performed with an Moloney murine leukemia virus reverse transcriptase (Clontech) and the 5' CDS primer A of the SMARTer RACE (random amplification of cDNA ends) kit (Clontech). The sequences were cloned in the pGEM-T vector (Promega) and verified by sequencing. To allow directional cloning into the multiple-cloning site of the bacterial expression vectors pGEX-4T2 (GE Healthcare) and pNIII, a second PCR was performed on the pGEM-T plasmid containing *LpCSP1* or *LpCSP2* to add proper restriction sites (see Table S1). *E. coli* CSpA was amplified from a pNIII-CspA plasmid (8) and cloned in pGEX4T2 using similar procedures. The correct frame was verified by sequencing, and the sizes of GST-*LpCSP1*, GST-*LpCSP2*, and GST-CspA fusion proteins were verified by SDS-PAGE. The pGEX4T2-*LpCSP1*, pGEX4T2-*LpCSP2*, and the pGEX4T2-CspA vector were used to transform chemically competent BL21 host cells (Life Technologies).

A single colony of BL21 *E. coli* containing either pGEX4T2-*LpCSP1*, pGEX4T2-*LpCSP2*, or pGEX4T2-CspA was inoculated in 5 ml of Luria-Bertani (LB) broth and cultured overnight at  $37^\circ\text{C}$  with vigorous shaking. The overnight cultures were transferred to 250 ml of LB medium supplemented with ampicillin ( $100 \mu\text{g/ml}$ ) and grown at  $37^\circ\text{C}$  with vigorous shaking to an optical density at 600 nm of 0.5. At this point, protein expression was induced for 2 h by addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM. After harvesting by centrifugation, the bacterial pellets were resuspended in cold phosphate-buffered saline (PBS) ( $137 \text{ mM NaCl}$ ,  $2.7 \text{ mM KCl}$ ,  $4.3 \text{ mM Na}_2\text{HPO}_4$ ,  $1.47 \text{ mM KH}_2\text{PO}_4$ ) containing  $1 \text{ mM}$  phenylmethylsulfonyl fluoride (PMSF),  $1 \text{ mM}$  dithiothreitol (DTT),  $1 \text{ mM}$  ethylenediaminetetraacetic acid (EDTA), and  $0.25\%$  Triton X-100 and broken with a French pressure cell press (Fisher Scientific). The cell lysates were then centrifuged at  $10,000 \times g$  for 10 min, and the supernatants were



incubated with 100  $\mu$ l of 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 300  $\mu$ l of PBS supplemented with 2 units of thrombin at room temperature for 2 h to cleave the GST tag. Supernatants containing the cleaved CSPs were then electrophoresed by SDS-PAGE on an 18% acrylamide gel to assess purity, and protein concentration was estimated using the Bradford assay (BioRad). Aliquots of purified protein were frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until further use.

**Bacterial complementation assay.** BX04, a quadruple deletion *E. coli* mutant lacking four CSPs, was used to assess the role of *LpCSP1* and *LpCSP2* in allowing growth at low temperature ( $18^{\circ}\text{C}$ ) (8). The *E. coli* CspA in pINIII and the empty pINIII vector were used as positive and negative controls, respectively. The plasmids were transformed into chemo-competent BX04 cells. A single transformed BX04 colony with the respective plasmids was inoculated in 5 ml of LB medium and then cultured overnight at  $37^{\circ}\text{C}$  with vigorous shaking. The cultures were streaked on a single LB plate with ampicillin (100  $\mu\text{g}/\text{ml}$ ) and IPTG (0.2 mM) and grown at  $18^{\circ}\text{C}$  for 120 h. Growth was monitored every 24 h, and after 120 h the same plate was incubated overnight at  $37^{\circ}\text{C}$ .

**2D-PAGE of proteins from cysts, LD12 cells, and thrombin-cleaved *LpCSP1* and *LpCSP2*.** Total protein was extracted from cysts and LD12 cells by using Trizol (Invitrogen) as described elsewhere (30). A total of 150  $\mu\text{g}$  of LD12 or cyst protein, or 150  $\mu\text{g}$  of LD12 protein spiked with 150 ng of either *LpCSP1* or *LpCSP2*, was loaded on separate 7-cm immobilized pH gradient strips (pH 3 to 10; Bio-Rad) as the first dimension. SDS-PAGE gels containing 15% acrylamide were used for the second dimension and were stained overnight with Coomassie brilliant blue G-250 (50).

**Electrophoretic mobility shift assays.** Double-stranded and single-stranded oligonucleotides were designed as described elsewhere (40) and used after high-performance liquid chromatography (HPLC) purification. The ssDNA was a 32-mer, 5'-TCGATCGGGGCGGGGCGATCGGGGCGGGGCGA-3', and the dsDNA was prepared by mixing equimolar amounts of the 25-mer 5'-GGTGAGGCTGATTGGCTGGGCAGGA-3' (the Y-box is shown in italics) and its reverse complement. All oligonucleotides were end-labeled with [ $\gamma$ - $^{32}\text{P}$ ]ATP (PerkinElmer) by using polynucleotide kinase (NEB) and the conditions described by the manufacturer. Labeled probes were purified using the QIAquick nucleotide removal kit (Qiagen). Typically, 1 ng radiolabeled ssDNA or dsDNA probe was incubated with the purified proteins for 30 min in  $1\times$  binding buffer ( $2\times$  binding buffer contained 20 mM Tris-Cl [pH 7.0], 50 mM KCl, 1 mM DTT, 20 mM  $\text{MgCl}_2$ , and 10% glycerol) at room temperature. Binding was assessed by electrophoresis on 5% native polyacrylamide gels in  $0.5\times$  Tris-borate-EDTA (TBE) buffer at 70 V for 1 h at room temperature, followed by overnight autoradiography on a phosphorimager screen (Amersham) and subsequent imaging with a Typhoon Trio+ (Amersham). For competition assays, the binding buffer was supplemented with a  $50\times$  excess of unlabeled ssDNA or dsDNA probe for specific competition and a  $50\times$  excess of random single-stranded oligonucleotides or salmon sperm DNA for nonspecific competition.

RNA probes were prepared by *in vitro* transcription using the T7 RiboMAX RNA production system (Promega) from a 120-bp dsDNA template that included the 5'-untranslated region of the peridinin chlorophyll *a* binding protein (PCP; GenBank accession number U93077) preceded by the spliced leader sequence (51) and a T7 promoter. The dsDNA templates were degraded after completion of the reaction by using RQ1 RNase-free DNase (Promega). The purified RNAs were end-labeled by the same procedure used for ssDNA and dsDNA probes, and unincorporated nucleotides were removed by chromatography on a Bio-Gel P10 column (Bio-Rad). Binding assays were performed as described above.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/mSphere.00034-15>.

Table S1, DOCX file, 0.02 MB.

Figure S1, DOCX file, 0.4 MB.

Figure S2, DOCX file, 0.3 MB.

Figure S3, DOCX file, 0.2 MB.

## ACKNOWLEDGMENTS

We gratefully acknowledge the generous gift of the BX04 mutant *E. coli* strain, the CspA in pINIII, and the empty pINIII vectors from M. Inouye and S. Phadtare.

M.B. is the recipient of a doctoral scholarship and S.P. and F.L. are recipients of undergraduate student research awards from the Natural Sciences and Engineering Research Council (NSERC). Funding for this work is provided through an NSERC research grant to D.M.

## FUNDING INFORMATION

NSERC Canada provided funding to David Morse under grant number 171382-03.

S.P. and F.L. were recipients of an NSERC summer studentship. M.B. was a recipient of a studentship from NSERC.

## REFERENCES

- Graumann PL, Marahiel MA. 1998. A superfamily of proteins that contain the cold-shock domain. *Trends Biochem Sci* **23**:286–290. [http://dx.doi.org/10.1016/S0968-0004\(98\)01255-9](http://dx.doi.org/10.1016/S0968-0004(98)01255-9).
- Didier DK, Schifffenbauer J, Woulfe SL, Zacheis M, Schwartz BD. 1988. Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box. *Proc Natl Acad Sci U S A* **85**:7322–7326. <http://dx.doi.org/10.1073/pnas.85.19.7322>.
- Brandi A, Pon CL, Gualerzi CO. 1994. Interaction of the main cold shock protein CS7.4 (CspA) of *Escherichia coli* with the promoter region of *hns*. *Biochimie* **76**:1090–1098. [http://dx.doi.org/10.1016/0300-9084\(94\)90035-3](http://dx.doi.org/10.1016/0300-9084(94)90035-3).
- Phadtare S, Inouye M. 1999. Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of *Escherichia coli*. *Mol Microbiol* **33**:1004–1014. <http://dx.doi.org/10.1046/j.1365-2958.1999.01541.x>.
- Jones PG, VanBogelen RA, Neidhardt FC. 1987. Induction of proteins in response to low temperature in *Escherichia coli*. *J Bacteriol* **169**:2092–2095.
- Gualerzi CO, Giuliodori AM, Pon CL. 2003. Transcriptional and post-transcriptional control of cold-shock genes. *J Mol Biol* **331**:527–539. [http://dx.doi.org/10.1016/S0022-2836\(03\)00732-0](http://dx.doi.org/10.1016/S0022-2836(03)00732-0).
- Goldstein J, Pollitt NS, Inouye M. 1990. Major cold shock protein of *Escherichia coli*. *Proc Natl Acad Sci U S A* **87**:283–287. <http://dx.doi.org/10.1073/pnas.87.1.283>.
- Xia B, Ke H, Inouye M. 2001. Acquisition of cold sensitivity by quadruple deletion of the *cspA* family and its suppression by PNPase S1 domain in *Escherichia coli*. *Mol Microbiol* **40**:179–188. <http://dx.doi.org/10.1046/j.1365-2958.2001.02372.x>.
- Jones PG, Krah R, Tafuri SR, Wolffe AP. 1992. DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. *J Bacteriol* **174**:5798–5802.
- Bae W, Xia B, Inouye M, Severinov K. 2000. *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci U S A* **97**:7784–7789. <http://dx.doi.org/10.1073/pnas.97.14.7784>.
- Jiang W, Hou Y, Inouye M. 1997. CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. *J Biol Chem* **272**:196–202. <http://dx.doi.org/10.1074/jbc.272.1.196>.
- Goldenberg D, Azar I, Oppenheim AB. 1996. Differential mRNA stability of the *cspA* gene in the cold-shock response of *Escherichia coli*. *Mol Microbiol* **19**:241–248. <http://dx.doi.org/10.1046/j.1365-2958.1996.363898.x>.
- Karlson D, Imai R. 2003. Conservation of the cold shock domain protein family in plants. *Plant Physiol* **131**:12–15. <http://dx.doi.org/10.1104/pp.014472>.
- Sasaki K, Imai R. 2011. Pleiotropic roles of cold shock domain proteins in plants. *Front Plant Sci* **2**:116. <http://dx.doi.org/10.3389/fpls.2011.00116>.
- Kim MH, Sasaki K, Imai R. 2009. Cold shock domain protein 3 regulates freezing tolerance in *Arabidopsis thaliana*. *J Biol Chem* **284**:23454–23460. <http://dx.doi.org/10.1074/jbc.M109.025791>.
- Park SJ, Kwak KJ, Oh TR, Kim YO, Kang H. 2009. Cold shock domain proteins affect seed germination and growth of *Arabidopsis thaliana* under abiotic stress conditions. *Plant Cell Physiol* **50**:869–878. <http://dx.doi.org/10.1093/pcp/pcp037>.
- Juntawong P, Sorenson R, Bailey-Serres J. 2013. Cold shock protein 1 chaperones mRNAs during translation in *Arabidopsis thaliana*. *Plant J* **74**:1016–1028. <http://dx.doi.org/10.1111/tpj.12187>.
- Kim JS, Park SJ, Kwak KJ, Kim YO, Kim JY, Song J, Jang B, Jung CH, Kang H. 2007. Cold shock domain proteins and glycine-rich RNA-binding proteins from *Arabidopsis thaliana* can promote the cold adaptation process in *Escherichia coli*. *Nucleic Acids Res* **35**:506–516. <http://dx.doi.org/10.1093/nar/gkl1076>.
- Yang Y, Karlson DT. 2011. Overexpression of AtCSP4 affects late stages of embryo development in *Arabidopsis*. *J Exp Bot* **62**:2079–2091. <http://dx.doi.org/10.1093/jxb/erq400>.
- Fusaro AF, Bocca SN, Ramos RL, Barrôco RM, Magioli C, Jorge VC, Coutinho TC, Rangel-Lima CM, De Rycke R, Inzé D, Engler G, Sachetto-Martins G. 2007. AtGRP2, a cold-induced nucleocytoplasmic RNA-binding protein, has a role in flower and seed development. *Planta* **225**:1339–1351. <http://dx.doi.org/10.1007/s00425-006-0444-4>.
- Chaikam V, Karlson D. 2008. Functional characterization of two cold shock domain proteins from *Oryza sativa*. *Plant Cell Environ* **31**:995–1006. <http://dx.doi.org/10.1111/j.1365-3040.2008.01811.x>.
- Evdokimova V, Ruzanov P, Imataka H, Raught B, Svitkin Y, Ovchinnikov LP, Sonenberg N. 2001. The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. *EMBO J* **20**:5491–5502. <http://dx.doi.org/10.1093/emboj/20.19.5491>.
- Evdokimova V, Tognon C, Ng T, Ruzanov P, Melnyk N, Fink D, Sorokin A, Ovchinnikov LP, Davicioni E, Triche TJ, Sorensen PH. 2009. Translational activation of *snail1* and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. *Cancer Cell* **15**:402–415. <http://dx.doi.org/10.1016/j.ccr.2009.03.017>.
- Bouvet P, Wolffe AP. 1994. A role for transcription and FRGY2 in masking maternal mRNA within *Xenopus* oocytes. *Cell* **77**:931–941. [http://dx.doi.org/10.1016/0092-8674\(94\)90141-4](http://dx.doi.org/10.1016/0092-8674(94)90141-4).
- Beauchemin M, Roy S, Daoust P, Dagenais-Bellefeuille S, Bertomeu T, Letourneau L, Lang BF, Morse D. 2012. Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic. *Proc Natl Acad Sci U S A* **109**:15793–15798. <http://dx.doi.org/10.1073/pnas.1206683109>.
- Bayer T, Aranda M, Sunagawa S, Yum LK, Desalvo MK, Lindquist E, Coffroth MA, Voolstra CR, Medina M. 2012. Symbiodinium transcriptomes: genome insights into the dinoflagellate symbionts of reef-building corals. *PLoS One* **7**:e35269. <http://dx.doi.org/10.1371/journal.pone.0035269>.
- Karlson D, Nakaminami K, Toyomasu T, Imai R. 2002. A cold-regulated nucleic acid-binding protein of winter wheat shares a domain with bacterial cold shock proteins. *J Biol Chem* **277**:35248–35256. <http://dx.doi.org/10.1074/jbc.M205774200>.
- Phadtare S. 2004. Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* **6**:125–136.
- Etchegaray JP, Inouye M. 1999. CspA, CspB, and CspG, major cold shock proteins of *Escherichia coli*, are induced at low temperature under conditions that completely block protein synthesis. *J Bacteriol* **181**:1827–1830.
- Roy S, Letourneau L, Morse D. 2014. Cold-induced cysts of the photosynthetic dinoflagellate *Lingulodinium polyedrum* have an arrested circadian bioluminescence rhythm and lower levels of protein phosphorylation. *Plant Physiol* **164**:966–977. <http://dx.doi.org/10.1104/pp.113.229856>.
- Guillebault D, Sasorith S, Derelle E, Wurtz JM, Lozano JC, Bingham S, Tora L, Moreau H. 2002. A new class of transcription initiation factor, intermediate between TBP and TLFs is present in the marine unicellular organism: the dinoflagellate *Cryptocodinium cohnii*. *J Biol Chem* **277**:40881–40886. <http://dx.doi.org/10.1074/jbc.M205624200>.
- Phadtare S, Inouye M. 2004. Genome-wide transcriptional analysis of the cold shock response in wild-type and cold-sensitive, quadruple-csp-deletion strains of *Escherichia coli*. *J Bacteriol* **186**:7007–7014. <http://dx.doi.org/10.1128/JB.186.20.7007-7014.2004>.
- Sachs R, Max KE, Heinemann U, Balbach J. 2012. RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* **18**:65–76. <http://dx.doi.org/10.1261/ma.02809212>.
- Wistow G. 1990. Cold shock and DNA binding. *Nature* **344**:823–824. <http://dx.doi.org/10.1038/344823c0>.
- Wolffe AP, Tafuri S, Ranjan M, Familari M. 1992. The Y-box factors: a family of nucleic acid binding proteins conserved from *Escherichia coli* to man. *New Biol* **4**:290–298.
- Wang S, Hirschberg R. 2011. Y-box protein-1 is a transcriptional regulator of BMP7. *J Cell Biochem* **112**:1130–1137. <http://dx.doi.org/10.1002/jcb.23027>.
- Lasham A, Moloney S, Hale T, Homer C, Zhang YF, Murison JG, Braithwaite AW, Watson J. 2003. The Y-box-binding protein, YB1, is a potential negative regulator of the p53 tumor suppressor. *J Biol Chem* **278**:35516–35523. <http://dx.doi.org/10.1074/jbc.M303920200>.
- Tafuri SR, Wolffe AP. 1990. *Xenopus* Y-box transcription factors: molecular cloning, functional analysis and developmental regulation. *Proc Natl Acad Sci U S A* **87**:9028–9032. <http://dx.doi.org/10.1073/pnas.87.22.9028>.
- MacDonald GH, Itoh-Lindstrom Y, Ting JP. 1995. The transcriptional regulatory protein, YB-1, promotes single-stranded regions in the DNA

- promoter. *J Biol Chem* **270**:3527–3533. <http://dx.doi.org/10.1074/jbc.270.8.3527>.
40. **Izumi H, Imamura T, Nagatani G, Ise T, Murakami T, Uramoto H, Torigoe T, Ishiguchi H, Yoshida Y, Nomoto M, Okamoto T, Uchiumi T, Kuwano M, Funa K, Kohno K.** 2001. Y box-binding protein-1 binds preferentially to single-stranded nucleic acids and exhibits 3'→5' exonuclease activity. *Nucleic Acids Res* **29**:1200–1207. <http://dx.doi.org/10.1093/nar/29.5.1200>.
  41. **Kim MH, Sonoda Y, Sasaki K, Kaminaka H, Imai R.** 2013. Interactome analysis reveals versatile functions of Arabidopsis cold shock domain protein 3 in RNA processing within the nucleus and cytoplasm. *Cell Stress Chaperones* **18**:517–525. <http://dx.doi.org/10.1007/s12192-012-0398-3>.
  42. **Sommerville J.** 1999. Activities of cold-shock domain proteins in translation control. *BioEssays* **21**:319–325. [http://dx.doi.org/10.1002/\(SICI\)1521-1878\(199904\)21:4<319::AID-BIES8>3.0.CO;2-3](http://dx.doi.org/10.1002/(SICI)1521-1878(199904)21:4<319::AID-BIES8>3.0.CO;2-3).
  43. **Morey JS, Van Dolah FM.** 2013. Global analysis of mRNA half-lives and de novo transcription in a dinoflagellate, *Karenia brevis*. *PLoS One* **8**:e66347. <http://dx.doi.org/10.1371/journal.pone.0066347>.
  44. **Roy S, Beauchemin M, Dagenais-Bellefeuille S, Letourneau L, Capadocia M, Morse D.** 2014. The *Lingulodinium* circadian system lacks rhythmic changes in transcript abundance. *BMC Biol* **12**:107. <http://dx.doi.org/10.1186/s12915-014-0107-z>.
  45. **Morey JS, Monroe EA, Kinney AL, Beal M, Johnson JG, Hitchcock GL, Van Dolah FM.** 2011. Transcriptomic response of the red tide dinoflagellate, *Karenia brevis*, to nitrogen and phosphorus depletion and addition. *BMC Genomics* **12**:346. <http://dx.doi.org/10.1186/1471-2164-12-346>.
  46. **Moustafa A, Evans AN, Kulis DM, Hackett JD, Erdner DL, Anderson DM, Bhattacharya D.** 2010. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PLoS One* **5**:e9688. <http://dx.doi.org/10.1371/journal.pone.0009688>.
  47. **Guillard RRL, Ryther JH.** 1962. Studies on marine planktonic diatoms: *Cyclotella nana* Hufstedt and *Denotula confervacea* (Cleve) Gran. *Can J Microbiol* **8**:229–239. <http://dx.doi.org/10.1139/m62-029>.
  48. **Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ, Beszteri B, Bidle KD, Cameron CT, Campbell L, Caron DA, Cattolico RA, Collier JL, Coyne K, Davy SK, Deschamps P, Dyrman ST, Edvardsen B, Gates RD, Gobler CJ, Greenwood SJ, Guida SM, Jacobi JL, Jakobsen KS, James ER, Jenkins B, John U, Johnson MD, Juhl AR, Kamp A, Katz LA, Kiene R, Kudryavtsev A, Leander BS, Lin S, Lovejoy C, Lynn D, Marchetti A, McManus G, Nedelcu AM, Menden-Deuer S, Miceli C, Mock T, Montresor M, Moran MA, Murray S, Nadathur G, Nagai S, Ngam PB, Panenik B, Pawlowski J, Petroni G, Piganeau G, Posewitz MC, Rengefors K, Romano G, Rumpho ME, Rynearson T, Schilling KB, Schroeder DC, Simpson AGB, Slamovits CH, Smith DR, Smith GJ, Smith SR, Sosik HM, Stief P, Theriot E, Twary SN, Umale PE, Vault D, Wawrik B, Wheeler GL, Wilson WH, Xu Y, Zingone A, Worden AZ.** 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* **12**:e1001889. <http://dx.doi.org/10.1371/journal.pbio.1001889>.
  49. **Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**:W465–W469. <http://dx.doi.org/10.1093/nar/gkn180>.
  50. **Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L, Righetti PG.** 2004. Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**:1327–1333. <http://dx.doi.org/10.1002/elps.200305844>.
  51. **Zhang H, Hou Y, Miranda L, Campbell DA, Sturm NR, Gaasterland T, Lin S.** 2007. Spliced leader RNA trans-splicing in dinoflagellates. *Proc Natl Acad Sci U S A* **104**:4618–4623. <http://dx.doi.org/10.1073/pnas.0700258104>.