

Research Article

Predicted Cold Shock Proteins from the Extremophilic Bacterium *Deinococcus maricopensis* and Related *Deinococcus* Species

Michael J. LaGier

Department of Biology, Grand View University, Des Moines, IA, USA

Correspondence should be addressed to Michael J. LaGier; mlagier@grandview.edu

Received 4 May 2017; Revised 5 July 2017; Accepted 10 July 2017; Published 18 September 2017

Academic Editor: Todd R. Callaway

Copyright © 2017 Michael J. LaGier. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

While many studies have examined the mechanisms by which extremophilic *Deinococci* survive exposure to ionizing radiation, very few publications have characterized the cold shock adaptations of this group, despite many species being found in persistent cold environments and environments prone to significant daily temperature fluctuations. Bacterial cold shock proteins (Csp) are a family of conserved, RNA chaperone proteins that commonly play a role in cold temperature adaptation, including a downward shift in temperature (i.e., cold shock). The primary aim of this study was to test whether a representative, desert-dwelling *Deinococcus*, *Deinococcus maricopensis*, encodes Csp as part of its genome. Bioinformatic approaches were used to identify a Csp from *D. maricopensis* LB-34. The Csp, termed Dm-Csp1, contains sequence features of Csp including a conserved cold shock domain and nucleic acid binding motifs. A tertiary model of Dm-Csp1 revealed an anticipated Csp structure containing five anti-parallel beta-strands, and ligand prediction experiments identified N-terminally located residues capable of binding single-stranded nucleic acids. Putative Csp were identified from 100% of (27 of 27) *Deinococci* species for which genome information is available; and the *Deinococci*-encoded Csp identified contain a C-terminally located region that appears to be limited to members of the class *Deinococci*.

1. Introduction

The Genus *Deinococcus* currently contains 59 recognized species (<http://www.bacterio.net>), including 27 species for which genomic information is available (<https://www.ncbi.nlm.nih.gov/genome/browse/>). Members of *Deinococcus* demonstrate a capacity to resist environmental stress including exposure to significant levels of ionizing radiation and desiccation [1]. Species of *Deinococci* have been isolated from a wide range of environments including air, desert soil, fresh and marine waters, alpine environments, Antarctica, hot springs, and radioactive sites [2]. One species, *Deinococcus maricopensis*, was originally isolated from an arid soil sample collected from the Sonoran Desert in Arizona (USA), and the genome of the type strain of this species (LB-34) has been recently sequenced as part of the Genomic Encyclopedia of Bacteria and Archaea

project [3, 4]. *D. maricopensis* LB-34 cells are mesophilic, Gram-positive, rod-shaped, and nonmotile, with an optimal growth temperature of 40°C and an observed growth range of 10°C to 45°C [3]. The relatively wide growth range of *D. maricopensis* correlates with observed diurnal swings of 15°C or more during a typical day within the Sonoran Desert (<https://science.nps.gov/im/units/sodn/sonoran.cfm>). Hence, *D. maricopensis* has adapted to an environment prone to rapid temperature changes, including a shift to colder temperatures.

Although many studies have examined the mechanisms by which *Deinococci* survive exposure to significant levels of ionizing radiation, few studies have looked at the cold shock adaptations of this group, despite many species being found in persistent cold environments (e.g., Antarctica, cold water oceans) and environments prone to significant daily temperature fluctuations such as deserts [5–7]. Using the model

species of the group, *Deinococcus radiodurans*, data from a previous proteomic study showed changes in global protein expression patterns following a cold shock [8]. Although this study was useful in demonstrating that gene expression changes do occur in response to cold shock, studies are lacking that examine additional *Deinococcus* genomes for putative cold shock proteins.

Bacterial cold shock proteins (Csps) represent a family of small proteins (smaller than 100 amino acids in length) with common structural features and a capacity to bind to single-stranded nucleic acids by way of conserved, N-terminally located nucleic acid-binding motifs [9]. Csp expression levels are typically increased following an ambient temperature decrease, but not always [9, 10]. Within the *Deinococcus* species, the best characterized protein that shares significant sequence homology to known Csps is from *D. radiodurans* [11]. The protein, termed PprM, appears to play a role in the radiative and oxidative stress resistance phenotypes of *D. radiodurans* [11]. Interestingly, while one study showed PprM to be induced [8] after heat shock (and apparently not induced after cold shock), a second study failed to demonstrate induction of PprM after heat shock [12]. Hence, our understanding of the potential role of Csps within the *Deinococcus* species as a group is limited.

Outside the *Deinococcus* species, the best studied Csps come from the mesophile *Escherichia coli*, although putative Csps have been identified from psychrophiles, mesophiles, and thermophiles [13]. In *E. coli*, nine members of the Csp family have been identified: CspA, B, C, D, E, F, G, H, and I. Although a significant number of the *E. coli* Csps are induced by cold (A, B, G, and I), E and C are constitutively expressed at physiological temperature, and D is induced via nutrient stress [9]. Of the nine, CspA is best characterized and appears to play a role in bacterial cold shock by acting as an RNA chaperone, destabilizing secondary structures within mRNAs that are more likely to form as extracellular temperatures decrease, and thus can reduce transcriptional and translational efficiency [9, 13, 14].

The current study identifies and characterizes, using bioinformatics, a potential Csp homolog from *D. maricopenensis*, a desert-dwelling species that is naturally exposed to significant temperature fluctuations including daily temperature downshifts [3]. The *D. maricopenensis* Csp, termed Dm-Csp1, contains expected nucleic acid binding motifs, shares significant sequence identity and similarity to known Csp proteins, contains five conserved anti-parallel beta-strands according to structural modeling, and is predicted to have a capacity to bind single-stranded nucleic acids. In addition, the presence of Csps was identified in all *Deinococcus* genomes queried (27 of 27). Interestingly, the Csps identified all contain a C-terminal region that is unique among members of the class *deinococci*; and for which a potential function is unknown.

2. Materials and Methods

2.1. Sequence Retrieval. Protein data available for *Deinococcus maricopenensis* LB-34 at the National Center for Biotechnology

Information (NCBI) was queried via a text search using “cold shock” as the input. An 86 amino acid long protein, NCBI Accession #WP_013556416.1 was then selected for further study by storing the protein sequence locally as a FASTA file. This protein was termed Dm-Csp1.

2.2. Analysis of Primary Features of Dm-Csp1. A tool within the ExPASy site, ProtParam (<http://web.expasy.org/protparam/>), was used to predict the basic physiological and chemical features of Dm-Csp1. For comparison, a known Csp protein (CspA) sequence from *E. coli* K12 (NCBI Accession #P0A9X9) was subjected to ProtParam analysis in parallel. Default settings were used for ProtParam analyses. ProtParam predicts properties of input protein sequences including molecular weight, isoelectric pH, aliphatic index, and extinction coefficients.

2.3. Sequence Similarity Analyses of Dm-Csp1. The FASTA protein sequence of Dm-Csp1 was used as a query in standard BLASTp searches against the nonredundant database (nr) as well as the PDB protein database, which includes only sequences for which a solved protein structure is available (<http://www.rcsb.org/pdb/home/home.do>). To identify potential Csp homologs in the *Deinococcus* as a group, a standard BLASTp search was run (using Dm-Csp1 as the query) for which the nr database was limited to sequences from members of the Genus *Deinococcus* (27 species total, including *D. maricopenensis*) containing permanent draft or finished genomic data available at the Integrated Microbial Genomes and Microbiomes (IMG/M) site (<https://img.jgi.doe.gov/>). A combination of Pfam and ScanProsite was used to identify conserved protein domains within the sequence of Dm-Csp1, under default settings [15, 16]. To determine the number and type of Csps in *D. maricopenensis*, CspA, B, C, D, E, F, G, H, and I protein sequences from *E. coli* K12 were used via BLASTp to query the genome of *D. maricopenensis* LB-34. Unless stated otherwise, BLASTp was run according to default parameters.

2.4. Multiple Sequence Alignment. BLASTp, with Dm-Csp1 as a query, was used to identify closely related proteins from within the Genus *Deinococcus*. Dm-Csp1, along with 10 putative Csp proteins from 10 representative *Deinococcus*, was retrieved as FASTA files from NCBI and aligned using standard parameters of the BioEdit Sequence Alignment Editor (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Amino acids 100% conserved (Figure 1) or 50% conserved (Figure 4) among the 11 aligned proteins were highlighted in blue, and functional motifs were identified by visual inspection of the output alignment.

2.5. Protein to Protein Interactions. The STRING tool was used to identify proteins that Dm-Csp1 may interact with *in vivo*. The STRING database contains information from experimental data and public text collections to predict protein-protein interactions [17]. The basic interaction unit in STRING is the functional association, likely contributing to a common biological purpose. Predicted protein-protein interactions are derived from multiple sources

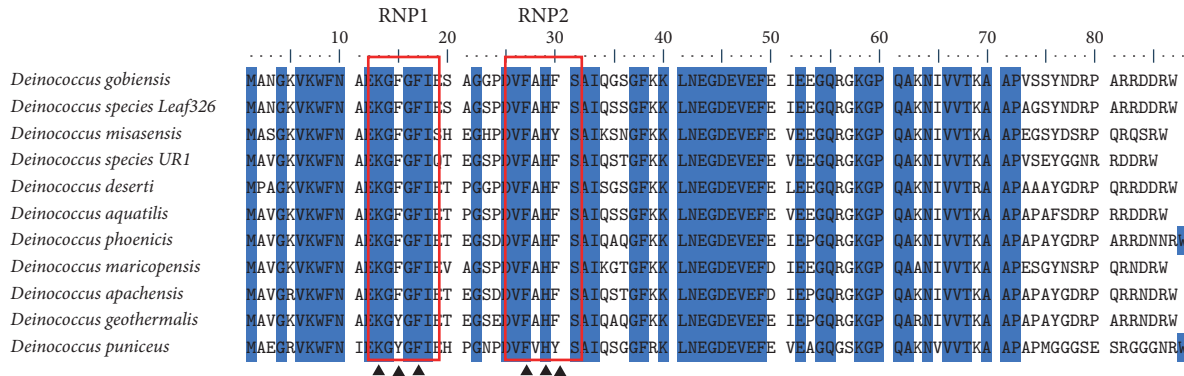


FIGURE 1: Multiple sequence alignment of *Dm-Csp1* and related, representative *Deinococci*-encoded cold shock proteins. The species names are indicated. Areas of 100% conservation are highlighted in blue. The position of conserved RNA-binding motifs (RNP1 and RNP2) is enclosed within red rectangles. Conserved, basic, and aromatic residues within RNP1 and RNP2 that form a nucleic acid binding surface are indicated by arrowheads.

including known experimental interactions (coexpression, experiments), genome functional pathway knowledge from manually curated databases (databases), automated text-mining searches (text-mining), shared gene neighborhood locations across multiple genomes (neighborhood), and interactions observed in one organism and systematically transferred to other organisms, via precomputed orthology relations (cooccurrence and gene fusion). Default parameters of STRING were used.

2.6. *Dm-Csp1* Structure Prediction and Model Quality Assessment. The secondary structure of *Dm-Csp1* was predicted by the PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>), under default settings. The tertiary structure of *Dm-Csp1* was modeled using the (PS)²-v2 server [18]. In (PS)²-v2, autodetect functionality was employed to select a suitable template structure to model against. The output model of *Dm-Csp1* was visualized by Chimera software [19]. The quality of the predicted structure was assessed by PROCHECK [20], QMEAN6 [21], and verify3D [22]. PROCHECK and QMEAN6 are programs within the SWISS-MODEL workspace (<https://swissmodel.expasy.org>), and verify3D is found at the UCLA-DOE Structure Evaluation Server (http://services.mbi.ucla.edu/Verify_3D/).

2.7. *Dm-Csp1* Ligand Binding Site Prediction. The model file generated from (PS)²-v2 was used as an input into COACH [23] to predict potential ligand binding regions within *Dm-Csp1*. COACH is a metasever approach to protein ligand binding site prediction. COACH generates predictions using TM-SITE and S-SITE, which recognize ligand binding templates from the BioLiP protein function database by binding specific substructure and sequence profile comparisons. These predictions are combined with predictions from COFACTOR, FINDSITE, and ConCavity to generate final ligand binding site predictions [23].

3. Results and Discussion

3.1. Identification of *Csps* from *D. maricopenis* and Related *Deinococci*. Initial text queries of *D. maricopenis* protein sequences available at NCBI identified a potential protein of 86 amino acids in length denoted as a cold shock protein (*Csp*). The protein, designated in this study as *Dm-Csp1* (NCBI Accession #WP_013556416.1), was then used as a query sequence in additional BLASTp searches. Specifically, BLASTp analysis against the nonredundant (nr) and PDB protein database showed that *Dm-Csp1* shares significant sequence identity and similarity to bacterial proteins annotated as *Csps*, including those for which a protein structure has been solved (Table 1). In addition, predicted physiological and chemical features of *Dm-Csp1* were similar to those obtained when compared to a known *Csp* from the model bacterium *E. coli*. As shown in Table 2, *Dm-Csp1* is predicted, according to ProtParam, to have a molecular mass of 7.4 Kilodaltons and an isoelectric pH of 5.58.

To better understand the conservation of *Csps* among the Genus *Deinococcus*, the sequence of *Dm-Csp1* was used as a BLASTp query to scan the genomes of related *Deinococci* species, for which genomic sequence information is available (27 total species, including *D. maricopenis*). As summarized in Table 3, at least one protein sharing significant sequence identity and similarity (as well as significant *E*-values), was identified in 100% of the *Deinococci* species queried. Although the number of *Csps* per bacterial genome can vary according to species, the reasoning behind such variation is unknown [10, 13]. The absolute conservation of *Csps* among the *Deinococci* sampled implies that *Csps* may be a significant feature of this groups' capacity in adapting to cold shock conditions and/or adapting or responding to other stress conditions.

While *E. coli* has been shown to contain a total of 9 *Csps* per genome, the *D. maricopenis* LB-34 genome appears to contain 2 *Csps*, according to BLAST testing (Section 2). Specifically, in addition to *Dm-Csp1*, the LB-34 genome

TABLE 1: BLASTp analysis of Dm-Csp1 (NCBI Accession #WP 013556416.1).

Database identifier ^a	BLASTp hit species	Name of hit	GenBank hit number	<i>E</i> -value	Identity (%)	Similarity (%)	Hit query coverage (%)
nr	<i>Deinococcus maricopensis</i>	Cold shock protein	WP 013556476.1 ^b	5.00E – 56	95	96	100
nr	<i>Deinococcus peraridilitoris</i>	Cold shock protein	WP 015234625.1	2.00E – 48	84	90	100
nr	<i>Deinococcus aqualitis</i>	Cold shock protein	WP 019008855.1	7.00E – 48	80	91	100
nr	<i>Deinococcus misasensis</i>	Cold shock protein	WP 034335902.1	1.00E – 47	80	88	100
nr	<i>Deinococcus pimensis</i>	Cold shock protein	WP 027481785.1	2.00E – 47	83	87	100
PDB	<i>Bacillus caldolyticus</i>	Cold shock protein	115F_A ^c	6.00E – 24	64	75	76
PDB	<i>Bacillus caldolyticus</i>	Cold shock protein	1C90_A	6.00E – 24	66	75	76
PDB	<i>Bacillus subtilis</i>	Cold shock protein	2ES2_A	6.00E – 24	64	73	82
PDB	<i>Bacillus caldolyticus</i>	Cold shock protein	1HZB_A	6.00E – 24	64	75	76
PDB	<i>Bacillus subtilis</i>	Cold shock protein	2I5L_X	8.00E – 24	66	73	74

^aNonredundant database (nr) or protein database (PDB), acquired via independent BLASTp searches using Dm-Csp1 as a query. ^bReferred to as Dm-Csp2 within text. ^cProtein database identifier code.

TABLE 2: Physiological and chemical properties of Dm-Csp1 and *E. coli* K12 CspA^a.

Protein identifier	Amino acid length	Molecular weight	Isoelectric point	Aliphatic index	Estimated half-life	Positively charged residues (%)	Negatively charged residues (%)
Dm-Csp1	86	9472.54	5.79	57.91	>10 hours	13.95	15.12
<i>E. coli</i> CspA	70	7403.28	5.58	61.29	>10 hours	10	11.42

^aCalculated using default settings of ProtParam (<http://web.expasy.org/protparam/>).

appears to contain a second Csp, termed in this study as Dm-Csp2 (NCBI Accession #WP 013556476.1), also 86 amino acids in total length. In comparison to *E. coli* Csps (A–I), Dm-Csp1 shares the greatest homology with *E. coli* CspA (BLASTp *E*-value of 9.0E – 23) and Dm-Csp2 shares greatest homology with CspB (BLASTp *E*-value of 2.0E – 21). Dm-Csp1 and Dm-Csp2 share 95% sequence identity and 96% sequence similarity. Dm-Csp1 and Dm-Csp2 differ in a total of 4 amino acids, none of which are found within key, conserved regions of known Csps (data not shown). Hence, Dm-Csp1 and Dm-Csp2 may serve similar functions in *D. maricopensis*. Within a single genome, it is anticipated that bacterial Csps share significant levels of sequence identity and similarity, as is the case with Csps from *E. coli* and *Yersinia pseudotuberculosis* [9]. In *E. coli* and *Y. pseudotuberculosis*, having more than one Csp sharing significant sequence homology is thought to represent functional redundancy, as mutagenesis of one or more Csps can be functionally compensated by remaining, nonmutated homologs [24, 25].

In addition to sequence homology testing, the primary sequence of Dm-Csp1 was compared to a sample of Deinococci-encoded Csps by multiple sequence alignments and was also used as a query to identify conserved protein domains. Pfam searching identified a conserved cold shock DNA-binding domain (Clan #CL0021) within the N-terminal region of Dm-Csp1 with a significant *E*-value score of 2.9E – 25. Consistent with Pfam, ScanProsite results indicated an N-terminally located cold shock domain signature (#PS00352) from amino acid position 15 to 33 with good confidence. A multiple sequence of Dm-Csp1 with 10 related Deinococci-encoded Csps revealed the presence of an additional feature of Csps, two nonspecific RNA-binding sequence motifs [9], ribonucleotide motifs 1 and 2 (Figure 1, RNP1 and RNP2). In bacterial Csps, the canonical sequence for RNP1 is (K/S-G-F/K/Y-G-F/L-I) and for RNP2 is (L/I/V-F/Q-V/A/L-H-X-S/T/R). In the aligned representative Deinococci (Figure 1) Csps, the sequences of RNP1 and RNP2 are very similar to expected, with RNP1 being

TABLE 3: BLASTp analysis of Dm-Csp1 (NCBI Accession #WP 013556416.1) versus deinococci group.

BLASTp hit species	Name of hit	Gene ID number ^a	E-value	Identity (%)	Similarity (%)
<i>Deinococcus maricopensis</i>	Cold shock protein	649901717	5.00E – 56	95	96
<i>D. reticulitermitis</i>	Cold shock protein	2617968483	6.00E – 45	76	87
<i>D. peraridilitoris</i>	Cold shock protein	2509591723	2.00E – 48	84	91
<i>D. deserti</i>	Cold shock protein	643799117	7.00E – 45	78	85
<i>D. xibeiensis</i>	Cold shock protein	2586308013	5.00E – 45	77	86
<i>D. marmoris</i>	Cold shock protein	2562219581	1.00E – 33	72	85
<i>D. wulumuqiensis</i>	Cold shock protein	2553983725	5.00E – 45	77	86
<i>D. species</i> 2009	Cold shock protein	2516938626	1.00E – 44	78	87
<i>D. pimensis</i>	Cold shock protein	2509082093	1.00E – 47	83	87
<i>D. misasensis</i>	Cold shock protein	2571066521	6.00E – 47	79	87
<i>D. phoenicis</i>	Cold shock protein	2579632034	1.00E – 40	78	84
<i>D. geothermalis</i>	Cold shock protein	641229653	2.00E – 45	80	87
<i>D. apachensis</i>	Cold shock protein	2519136834	3.00E – 47	84	87
<i>D. gobiensis</i>	Cold shock protein	25128201432	3.00E – 47	81	88
<i>D. proteolyticus</i>	Cold shock protein	649963862	5.00E – 33	62	77
<i>D. hopiensis</i>	Cold shock protein	2509586483	2.00E – 46	83	86
<i>D. grandis</i>	Cold shock protein	2691382520	1.00E – 45	80	90
<i>D. species</i> RL	Cold shock protein	2611619696	3.00E – 47	81	88
<i>D. radiodurans</i>	Cold shock protein	2558116098	5.00E – 45	77	86
<i>D. murrayi</i>	Cold shock protein	2528309711	3.00E – 47	81	88
<i>D. ficus</i>	Cold shock protein	2526061465	1.00E – 33	78	87
<i>D. fringens</i>	Cold shock protein	2562223450	1.00E – 33	72	85
<i>D. humi</i>	Cold shock protein	2597538219	2.00E – 47	79	92
<i>D. soli</i>	Cold shock protein	2637699313	1.00E – 45	80	90
<i>D. species</i> Leaf 326	Cold shock protein	2645444050	2.00E – 46	82	90
<i>D. aquatilis</i>	Cold shock protein	2520600036	7.00E – 48	80	92
<i>D. species</i> YIM 77859	Cold shock protein	2582523845	8.00E – 47	82	91

^aAccording to the Integrated Microbial Genomes database (<https://img.jgi.doe.gov/>).

(K-G-F/Y-G-F-I-X-X) and RNP2 being (V-F-V/A-H-F/Y-S). Within the known tertiary structures of Csps, the basic and aromatic residues within RNP1 and RNP2, collectively, form a nucleic acid-binding surface (Figure 1).

3.2. Structural Analyses of Dm-Csp1. To better understand the potential biological function of Dm-Csp1, a structural model of Dm-Csp1 was constructed. The secondary structure of Dm-Csp1 was predicted using the PSIPRED server. The result of PSIPRED, as shown in Figure 2, has good confidence and contains 5 beta-strands, which are expected among Csps [9, 26].

The modeling of the likely tertiary structure of Dm-Csp1, as predicted by (PS)²-v2, used a known Csp from *Bacillus caldolyticus* (PDB #1C9O) as a template. The *B. caldolyticus* Csp crystal structure used as a template is 66 amino acids in length and shares 61.76% sequence identity with Dm-Csp1 [27]. The model of Dm-Csp1 was found to contain five antiparallel beta-strands that form a barrel structure (Figure 3). These features are known characteristics of Csps. The quality of the output model structure was assessed using a combination of PROCHECK, QMEAN6, and verify3D (Section 2). PROCHECK reveals that all amino acids residues within the model (68 of 68) are within the

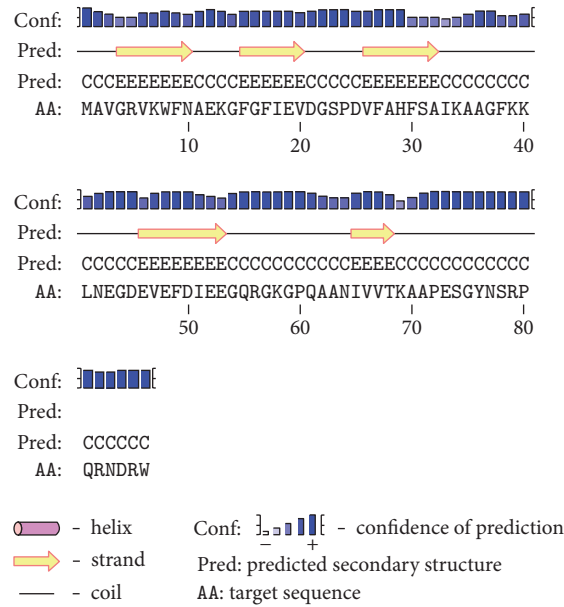


FIGURE 2: Secondary protein structure of *Dm-Csp1* as predicted by PSIPRED. The relative height of blue bars indicates confidence of prediction.

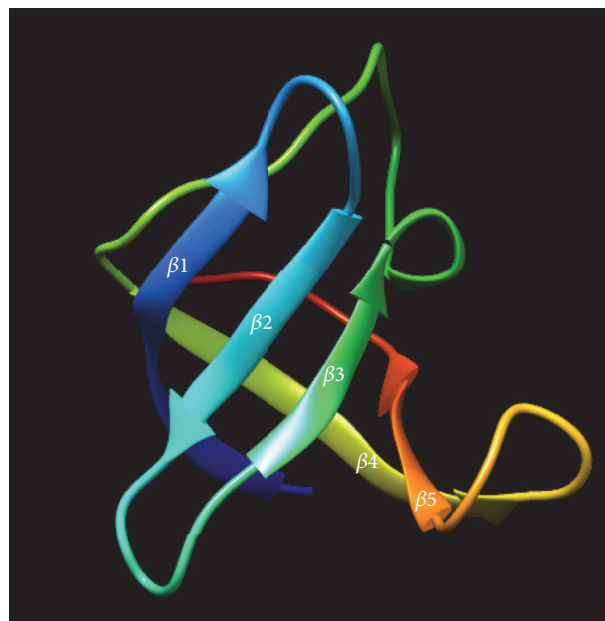


FIGURE 3: Modeled tertiary structure of *Dm-Csp1*. The five anti-parallel strands that collectively form a barrel structure are indicated.

limits of a Ramachandran plot, with 96.4% of residues in the most favored regions and only a single residue (1.8%) in a disallowed region (Figure 4(a)). A good model is expected to have over 90% of residues in most favored regions [20]. QMEAN6, which estimates the global quality of structural models on the basis of 6 structural descriptors, places the model of *Dm-Csp1* in the dark region of the

estimated absolute model quality graph with a significant QMEAN6 score of 0.781 and a significant Z -score of 0.413 (Figure 4(b)). Using verify3D, the environmental profile of the model was acceptable (Figure 4(c)), yielding a high score of 0.57 [22].

The C-terminal region of *Dm-Csp1* could not be modeled, specifically from amino acids 69 to 86. The reason

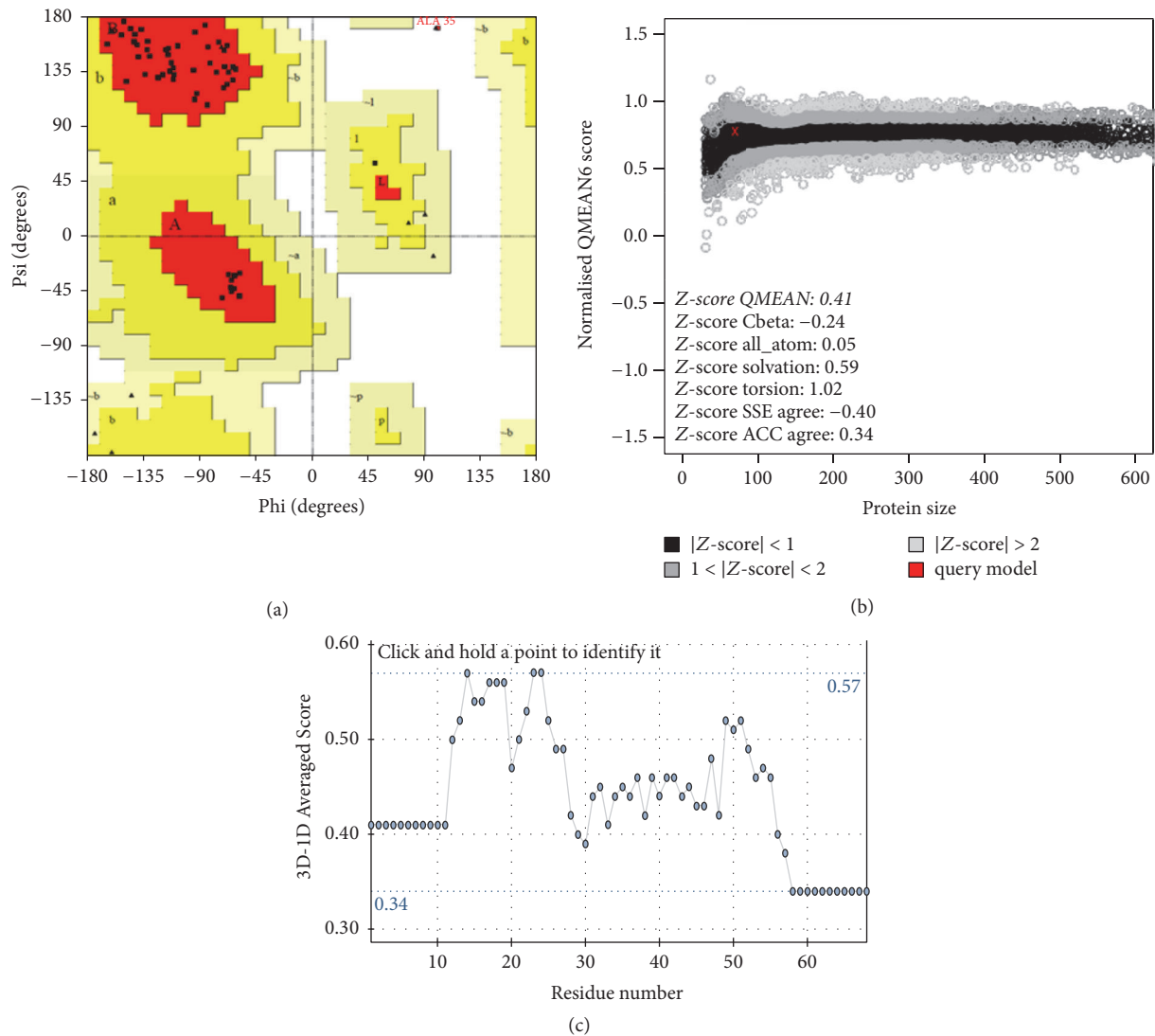


FIGURE 4: Tertiary structural model assessment. (a) Ramachandran plot of protein model generated by PROCHECK. The position of 96.4% amino acids within the generated structural model are within most favored regions. The position of 1 amino acid (ALA-35, red text) is found within a disallowed region. (b) Graphical output of estimation of absolute quality of model via QMEAN6. The position of the Dm-Csp1 tertiary model is indicated by a red X. (c) Verify3D output data for Dm-Csp1 predicted tertiary structure, yielding an acceptably high score of 0.57.

for this inability to model is due to the overall length of Dm-Csp1 being 86 amino acids. Bacterial Csps identified thus far are approximately 70 amino acids in length, with the N-terminal region containing conserved nucleic acid binding regions [9, 13]. The extended C-terminal region of Dm-Csp1, according to BLASTp analysis, appears to be unique to the class Deinococci. For example, in addition to species of Deinococci, some members (for which genomic information is available) of the class Deinococci encode at least one putative Csp with the C-terminal extension including *Truepera radiovictrix* (data not shown, NCBI Accession #WP_013178602). As shown in Figure 5, all the Deinococci Csps included in the Figure 1 alignment contain an extended C-terminal region that appears to be rich in

positively charged arginine residues and end in the combination arginine-tryptophan. The potential function of this conserved region within Deinococci Csps is unknown but warrants further investigation.

3.3. Biological Function Prediction. The synthesized, and quality assessed, structural model of Dm-Csp1 was used as a template to predict the potential of the protein to bind ligands. Using the COACH server, the N-terminal region of Dm-Csp1 appears able to bind single-stranded nucleic acids, with a significant support score (C-score) of 0.2 [23]. A total of 8 residues are predicted to bind nucleic acids, 4 of which are located with the previously identified within

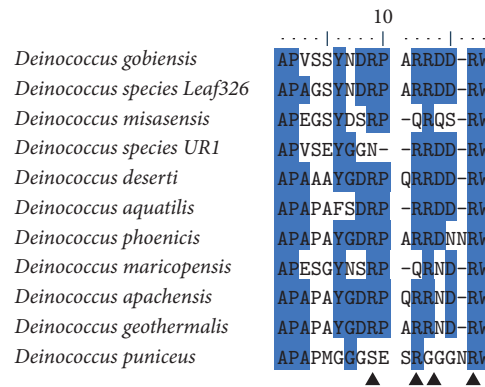


FIGURE 5: Multiple sequence alignment of C-terminal region of Dm-Csp1 and related, representative Deinococci-encoded cold shock proteins. Species names are indicated, regions of 50% or greater conservation are marked blue, and the positions of arginine-rich areas are denoted by arrowheads.

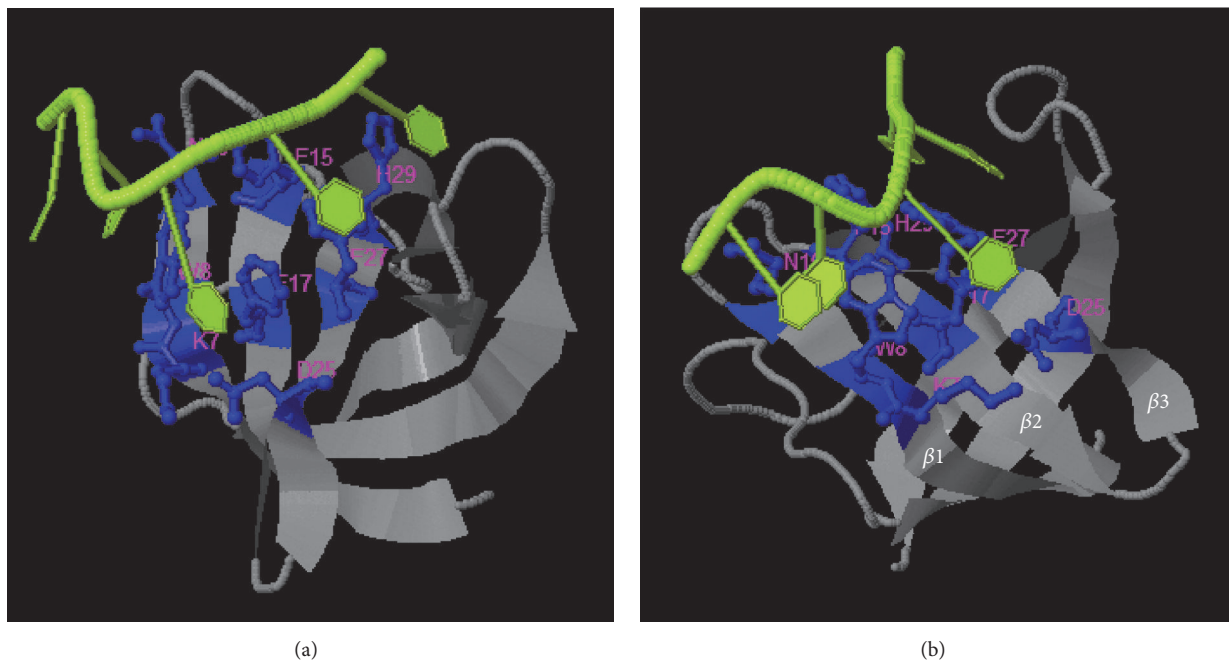


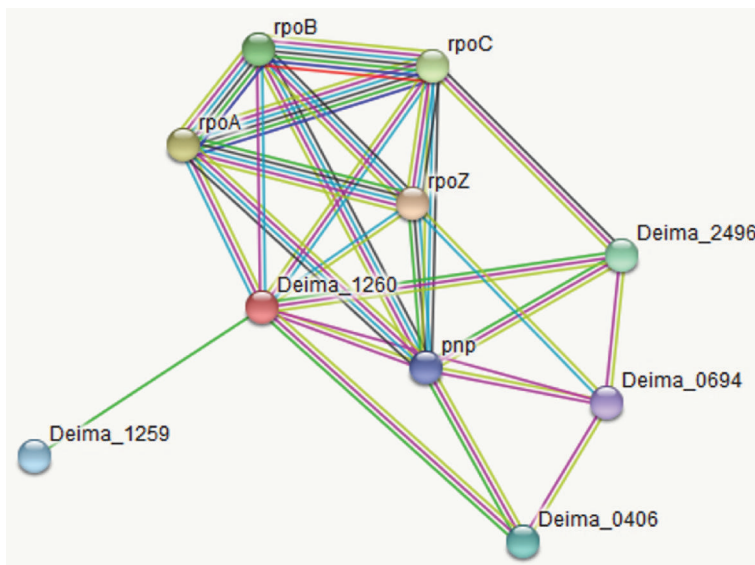
FIGURE 6: Functional site prediction of Dm-Csp1. COACH output identifying N-terminally located residues predicted to bind single-stranded nucleic acids ((a), nucleic acid colored green, binding residues colored blue), and the same output repositioned to show the predicted nucleic acid binding surface is found within three anti-parallel beta-strands that collectively form a nucleic acid binding surface (b).

the RNP1 and RNP2 motifs including F15, F17, F27, and H29 (Figure 6(a)). The binding of nucleic acids using a combination of three sequential beta-strands (Figure 6(b)) is consistent with previously solved structures of Csp1 [9, 13, 14, 26, 28]. According to STRING analysis, Dm-Csp1 is predicted to functionally interact with a variety of proteins including those (rpoB, rpoC, and rpoA, and rpoZ, RNA polymerase subunits) involved in transcription and a polyribonucleotide nucleotidyltransferase (pnp) potentially playing a role in mRNA degradation (Figure 7). The predicted interactions of Dm-Csp1 with transcription-related proteins make sense when considering previous studies demonstrating the role of

Csps in modulating transcription, including acting as a RNA chaperone [9, 13, 29, 30].

4. Conclusions

The findings here suggest that one adaptation by which *D. maricopensis* deals with significant, diurnal temperature shifts is by utilizing genome-encoded cold shock proteins (Csps). Significantly, all additional members of the Genus *Deinococcus*, for which genomic information is available, also encode Csps. Interestingly, the identified *Deinococci* Csps contain a C-terminal extension not observed in bacteria



(a)

Your Input:

Deima_1260 cold-shock DNA-binding domain-containing protein (86 aa)

Predicted Functional Partners:

		Neighborhood	Gene fusion	Cooccurrence	Coexpression	Experiments	Databases	Text-mining	[Homology]	Score
rpoZ	DNA-directed RNA polymerase subunit omega; Promotes RNA polymerase assembly. Latches the N- and C- terminal regio...					●		●		0.770
rpoA	DNA-directed RNA polymerase subunit alpha; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RN...					●		●		0.723
rpoC	DNA-directed RNA polymerase subunit beta ¹ ; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RN...					●		●		0.641
rpoB	DNA-directed RNA polymerase subunit beta; DNA-dependent RNA polymerase catalyzes the transcription of DNA into RN...					●		●		0.625
Deima_2496	DEAD/DEAH box helicase (604 aa)		●			●		●		0.609
Deima_0406	DEAD/DEAH box helicase; The UvrABC repair system catalyzes the recognition and processing of DNA lesions. A damage ...		●			●		●		0.597
Deima_1259	hypothetical protein (299 aa)		●							0.483
pnp	polyribonucleotide nucleotidyltransferase; Involved in mRNA degradation. Hydrolyzes single-stranded polyribonucleotides ...					●		●		0.481
Deima_0694	ATP-dependent helicase HrpB (831 aa)					●				0.406

(b)

FIGURE 7: STRING protein-protein interactions of Dm-Csp1. STRING network map (a) and summary table of predicted interactions (b). The category of interaction is shown in the table. Color saturation of the edges represents the confidence score of a given functional or physical association. The position of Dm-Csp1 is labeled Deima_1260. Shown interactions are considered of high confidence when yielding scores of greater than 0.7.

outside the class Deinococci, and for which a function is undefined. Future studies will seek to further characterize the biological roles of Dm-Csp1, and related Csp1, via *in vitro* approaches.

Conflicts of Interest

The author does not have any conflicts of interest to declare.

Acknowledgments

The author thanks all members of the Biology Department at Grand View University for encouragement and project support.

References

- [1] K. S. Makarova, L. Aravind, Y. I. Wolf et al., “Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics,” *Microbiology and Molecular Biology Reviews*, vol. 65, no. 1, pp. 44–79, 2001.
- [2] E. Gerber, R. Bernard, S. Castang et al., “Deinococcus as new chassis for industrial biotechnology: biology, physiology and tools,” *Journal of Applied Microbiology*, vol. 119, no. 1, pp. 1–10, 2015.
- [3] F. A. Rainey, K. Ray, M. Ferreira et al., “Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus *Deinococcus* obtained from a single soil sample,” *Applied and Environmental Microbiology*, vol. 71, no. 9, pp. 5225–5235, 2005.

- [4] R. Pukall, A. Zeytun, S. Lucas et al., "Complete genome sequence of *Deinococcus maricopensis* type strain (LB-34 T)," *Standards in Genomic Sciences*, vol. 4, no. 2, pp. 163–172, 2011.
- [5] A. de Groot, V. Chapon, P. Servant et al., "*Deinococcus deserti* sp. nov., a gamma-radiation-tolerant bacterium isolated from the Sahara Desert," *International Journal of Systematic and Evolutionary Microbiology*, vol. 55, no. 6, pp. 2441–2446, 2005.
- [6] S. Srinivasan, S. Lim, J.-H. Lim, H.-Y. Jung, and M. K. Kim, "*Deinococcus rubrus* sp. Nov., a bacterium isolated from Antarctic coastal sea water," *Journal of Microbiology and Biotechnology*, vol. 27, no. 3, pp. 535–541, 2017.
- [7] M. Yuan, W. Zhang, S. Dai et al., "*Deinococcus gobiensis* sp. nov., an extremely radiation-resistant bacterium," *International Journal of Systematic and Evolutionary Microbiology*, vol. 59, part 6, pp. 1513–1517, 2009.
- [8] A. Airo, S. L. Chan, Z. Martinez, M. O. Platt, and J. D. Trent, "Heat shock and cold shock in *Deinococcus radiodurans*," *Cell Biochemistry and Biophysics*, vol. 40, no. 3, pp. 277–288, 2004.
- [9] G. Horn, R. Hofweber, W. Kremer, and H. R. Kalbitzer, "Structure and function of bacterial cold shock proteins," *Cellular and Molecular Life Sciences*, vol. 64, no. 12, pp. 1457–1470, 2007.
- [10] D. N. Ermolenko and G. I. Makhatadze, "Bacterial cold-shock proteins," *Cellular and Molecular Life Sciences*, vol. 59, no. 11, pp. 1902–1913, 2002.
- [11] S.-H. Park, H. Singh, D. Appukuttan et al., "PprM, a cold shock domain-containing protein from *Deinococcus radiodurans*, confers oxidative stress tolerance to *Escherichia coli*," *Frontiers in Microbiology*, vol. 7, article no. 2124, 2017.
- [12] A. K. Schmid, M. S. Lipton, H. Mottaz, M. E. Monroe, R. D. Smith, and M. E. Lidstrom, "Global whole-cell FTICR mass spectrometric proteomics analysis of the heat shock response in the radioresistant bacterium *Deinococcus radiodurans*," *Journal of Proteome Research*, vol. 4, no. 3, pp. 709–718, 2005.
- [13] R. Keto-Timonen, N. Hietala, E. Palonen, A. Hakakorpi, M. Lindström, and H. Korkeala, "Cold shock proteins: a minireview with special emphasis on Csp-family of enteropathogenic *Yersinia*," *Frontiers in Microbiology*, vol. 7, p. 1151, 2016.
- [14] K. Yamanaka, "Cold shock response in *Escherichia coli*," *J. Mol. Microbiol. Biotechnol.*, vol. 1, no. 2, pp. 193–202, 1999.
- [15] E. de Castro, C. J. A. Sigrist, A. Gattiker et al., "ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins," *Nucleic Acids Research*, vol. 34, pp. W362–W365, 2006.
- [16] R. Finn, S. Griffiths-Jones, and A. Bateman, "Identifying protein domains with the Pfam database," in *Current Protocols in Bioinformatics*, chapter 2, Unit 2.5, 2003.
- [17] D. Szklarczyk, A. Franceschini, S. Wyder et al., "STRING v10: protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 43, pp. D447–D452, 2015.
- [18] C.-C. Chen, J.-K. Hwang, and J.-M. Yang, "(PS)²-v2: template-based protein structure prediction server," *BMC Bioinformatics*, vol. 10, article 366, 2009.
- [19] E. F. Pettersen, T. D. Goddard, C. C. Huang et al., "UCSF Chimera—a visualization system for exploratory research and analysis," *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.
- [20] A. Bhattacharya, Z. Wunderlich, D. Monleon, R. Tejero, and G. T. Montelione, "Assessing model accuracy using the homology modeling automatically software," *Proteins*, vol. 70, no. 1, pp. 105–118, 2008.
- [21] P. Benkert, S. C. E. Tosatto, and D. Schomburg, "QMEAN: a comprehensive scoring function for model quality assessment," *Proteins*, vol. 71, no. 1, pp. 261–277, 2008.
- [22] R. Luthy, J. U. Bowie, and D. Eisenberg, "Assesment of protein models with three-dimensional profiles," *Nature*, vol. 356, no. 6364, pp. 83–85, 1992.
- [23] J. Yang, A. Roy, and Y. Zhang, "Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment," *Bioinformatics*, vol. 29, no. 20, pp. 2588–2595, 2013.
- [24] K. Neuhaus, S. Rapposch, K. P. Francis, and S. Scherer, "Restart of exponential growth of cold-shocked *Yersinia enterocolitica* occurs after down-regulation of cspA1/A2 mRNA," *Journal of Bacteriology*, vol. 182, no. 11, pp. 3285–3288, 2000.
- [25] B. Xia, H. Ke, and M. Inouye, "Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in *Escherichia coli*," *Molecular Microbiology*, vol. 40, no. 1, pp. 179–188, 2001.
- [26] G. D'Auria, C. Esposito, L. Falcigno et al., "Dynamical properties of cold shock protein A from *Mycobacterium tuberculosis*," *Biochemical and Biophysical Research Communications*, vol. 402, no. 4, pp. 693–698, 2010.
- [27] U. Mueller, D. Perl, F. X. Schmid, and U. Heinemann, "Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein," *Journal of Molecular Biology*, vol. 297, no. 4, pp. 975–988, 2000.
- [28] J. Lee, K.-W. Jeong, B. Jin et al., "Structural and dynamic features of cold-shock proteins of *Listeria monocytogenes*, a psychophilic bacterium," *Biochemistry*, vol. 52, no. 14, pp. 2492–2504, 2013.
- [29] V. Chaikam and D. T. Karlson, "Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins," *BMB Reports*, vol. 43, no. 1, pp. 1–8, 2010.
- [30] S. Phadtare, V. Tadigotla, W.-H. Shin, A. Sengupta, and K. Severinov, "Analysis of *Escherichia coli* global gene expression profiles in response to overexpression and deletion of CspC and CspE," *Journal of Bacteriology*, vol. 188, no. 7, pp. 2521–2527, 2006.