Regulation of a cold inducible antifreeze protein in soil bacterium *Pseudomonas putida* GR12-2

by

Janet Lorv

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2014

Author’s Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In cold inhabiting bacteria, a commonly used strategy to combat freezing stress is the production of antifreeze proteins. These proteins assist in survival below 0°C by minimizing extracellular ice crystal growth. With a temperature dependent activity, these proteins are regulated by changes in ambient temperature. In the soil bacterium Pseudomonas putida GR12-2, expression of antifreeze protein AfpA was previously observed to be regulated by cold induction at 5°C. To better understand how this antifreeze protein helps with low temperature survival, I investigated AfpA regulation at multiple levels. Transcriptional regulation of this protein was investigated by studying the DNA upstream of the antifreeze gene. Within this region, an RpoD promoter with a CspA regulation element was found. Using an indirect luciferase reporter gene assay, the afpA promoter was found to be most active during exponential growth under optimal conditions. From these results, temperature dependent regulation was determined to be at the posttranscriptional level. At this level, I determined that AfpA expression was likely regulated by formation of different RNA structures. From a phylogenetic analysis, I also determined that *afpA* may be regulated as an operon with multiple promoters that can individually regulate each gene. Aside from temperature dependent regulation, I further characterized AfpA. This protein was identified to have ice recrystallization inhibition activity, a putative ice binding site and is likely secreted by a type I secretion system.

Acknowledgements

After two years, I could not have written this thesis alone without a great group of people to support me. First and foremost I would like to express my sincerest gratitude to both my supervisors, Dr. Bernie Glick and Dr. David Rose, for their support throughout the years. I would like to thank them for giving me their patience, encouragement and perspectives during the rough moments of research. Thank you to Dr. Barbara Butler for your encouragement during committee meetings. I would also like to thank Dr. Brendan McConkey and Dr. Andrew Doxey for their help and guidance with Bioinformatics. A special thank you goes to Dr. Stephen Graether at the University of Guelph for giving me access to equipment and taking the time to train me.

During these years, I am very lucky to have a great lab and lab members. I am very grateful to former lab members Jin Duan and Shimaila Ali for mentoring me when I was new to the world of research. Your patience and suggestions were invaluable to me. A special thank you to my current lab mate, Daiana Duca. You made this journey fun and bearable. I am grateful that we both managed to keep each other sane. Thank you to the Rose lab for their help when I had trouble with my proteins. I would also like to thank the Bols lab for giving me access to their equipment at night.

Finally, I would like to thank all my friends and family for all their love and support. I would like to thank my friends for being supportive and understanding about my schedule. Although I was not grateful at the time, thank you guys for dragging me out to have fun instead of working. I am immensely grateful to my parents and siblings for encouraging me and believing in me despite being clueless about my work.

Table of Contents

List of Figures

List of Tables

1.0 Introduction

1.1 Microbial Cold Adaptation

On Earth, cold environments are one of the most common conditions that microorganisms encounter. With approximately 80 to 85% of Earth's biosphere exposed to temperatures 15°C or below, microorganisms indigenous to cold habitats experience many challenges arising from cold stress (Margesin and Miteva 2011). One such challenge is cellular desiccation, where low temperatures hinder cellular processes by decreasing enzymatic activity and aggregating endogenous proteins (Bouvet and Ben 2003).

In addition, these microbes are in danger of cell death by freezing either through osmotic pressure changes or physical rupture by ice crystals (Sun et al. 1995). Cryodamage incurred from these scenarios is dependent on the bacterial freezing rate and location. Internal frost damage is minimized with a high freezing rate (-100°C/min) and becomes increasingly detrimental as the freezing rate decreases (Bouvet and Ben 2003). For intracellular freezing, a lower freezing rate leads to formation of large ice crystals that can physically rupture cells. On the other hand, external freezing can mimic high saline environments by concentrating extracellular solutes through removal of available water during ice formation (Kawahara 2008). This change in osmotic pressure draws out internal water required for cellular processes leading to cell death.

To cope with freezing stress, many organisms have developed one of two main survival strategies, freeze avoidance or freeze tolerance. Freeze avoidant organisms rely

on the ability to prevent ambient temperature from reaching below the environment's freezing point (Gilbert et al. 2005; Middleton et al. 2012). This strategy is optimal for mobile organisms, such as fish and insects, and habitats with low temperature fluctuations. However, some organisms will encounter inevitable freezing. These organisms, such as many immobile microbes and plants, employ a freeze tolerance strategy, which aims to minimize frost damage (Griffith and Ewart 1995; Sidebottom et al. 2000). This strategy usually occurs in environments with high fluctuations in temperature or consistently low subzero temperatures (Middleton et al. 2012; Wilson et al. 2006). Despite the different objectives, both survival strategies use ice crystal controlling proteins, such as antifreeze and ice nucleation proteins.

Ice crystal controlling proteins combat freezing stress by altering ice crystal formation and growth (Kawahara 2008). Using very similar mechanisms of action, each ice crystal controlling protein targets one of the two stages of ice crystal formation (Figure 1). Ice nucleation proteins facilitate the formation of embryonic ice crystals while antifreeze proteins interact and inhibit subsequent growth. However, not all organisms produce ice crystal controlling proteins. In organisms that do not, freeze avoidance or tolerance is achieved through alternative methods, which include cryoprotectant production and sporulation (Wilson et al. 2006; Wilson and Walker 2010; Xu et al. 1998). Since cold adaptation is a complex process, microbes may use more than one strategy or none of the ones mentioned (Wilson et al. 2006).

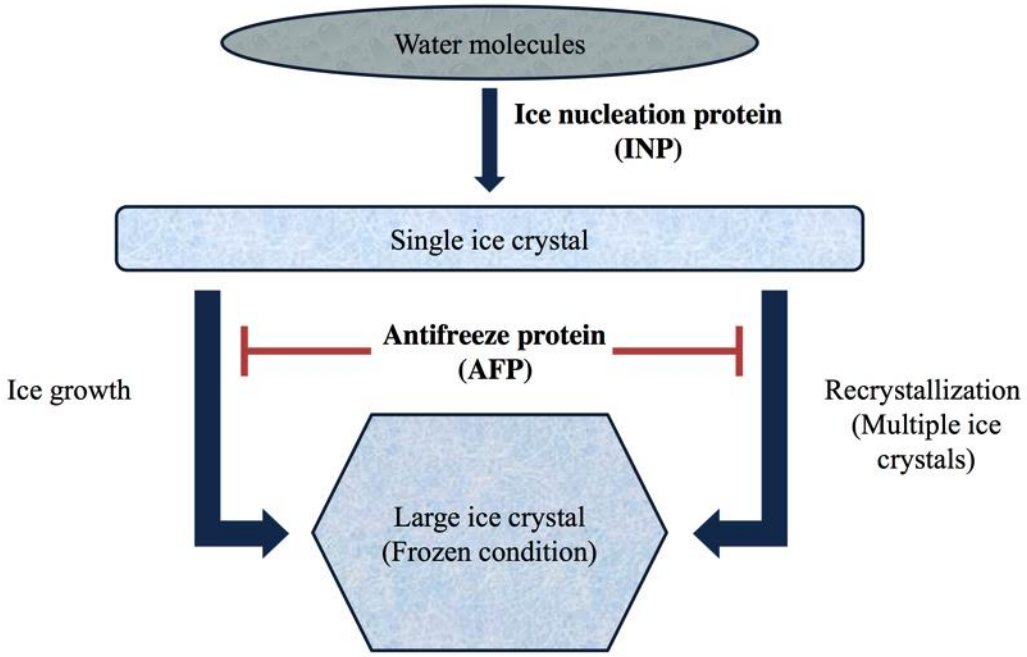


Figure 1. The two stages of ice crystal formation. Ice nucleation proteins facilitate formation of a single embryonic ice crystal while antifreeze proteins inhibit further growth into larger ice crystals (figure taken from Lorv et al., 2014)

1.1.1 Ice Crystals

The mechanism used by ice crystal controlling proteins involves direct interaction of these proteins with ice crystal surfaces. Ice crystals, as a three-dimensional organized structure of water molecules, contain many different surfaces, known as ice planes, and are sites of protein-ice interaction (Griffith and Ewart 1995). These ice planes differ by the spacing of oxygen atoms due to the asymmetry of water molecules. The two main ice crystal planes are the primary prism plane and the basal plane (Figure 2). The primary prism plane lies perpendicular to six α-axes and parallel to the c-axis while the basal plane is the inverse, lying perpendicular to the c-axis and parallel to the α-axes (Kawahara 2008; Raymond and DeVries 1977). These two ice planes also differ by oxygen spacing on the ice surface. Oxygen atoms are spaced in regions of 7.35 by 4.5 Ångstrom in the primary prism plane, but are spaced in regions of 7.8 by 4.5 Ångstrom in the basal plane. An ice crystal usually grows by adding an additional layer of water molecules atop these ice planes. Consequently, protein-ice interactions at these planes can affect these extra layers and alter ice crystal growth.

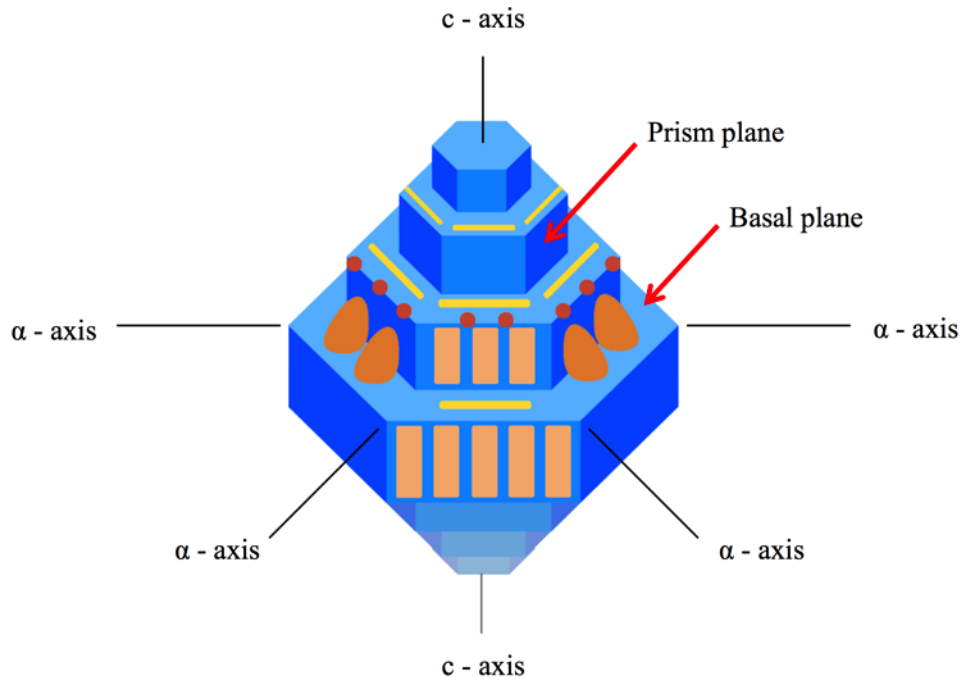


Figure 2. Morphology of an ice crystal. Antifreeze proteins can bind to the two main ice crystal planes, the (primary) prism plane and the basal plane. This binding shapes ice crystal morphology by inhibiting ice growth along the axes. Blue – ice crystal; non-blue objects – antifreeze proteins (figure taken from Lorv et al., 2014)

1.1.2 Antifreeze Proteins

Antifreeze proteins are the most well studied ice crystal controlling proteins at the protein level. These proteins are characterized by their ability to control ice crystal growth and shape through an adsorption inhibition mechanism upon binding to ice crystal planes (Kawahara 2008; Xu et al. 1998). This protein-ice interaction gives these proteins two distinct ice crystal controlling activities: thermal hysteresis activity and ice recrystallization inhibition (Raymond and DeVries 1977; Scotter et al. 2006).

Thermal hysteresis is the temperature gap that is created between the melting and freezing temperature of an ice crystal (Jia and Davies 2002). This gap is created when antifreeze protein non-colligatively depresses the freezing temperature, known as freezing hysteresis, as well as slightly elevates the melting temperature, known as melting hysteresis (Celik et al. 2010). When ambient temperature falls in this thermal hysteresis, ice crystals observed by a nanoliter osmometer are reported to neither grow nor melt (Gilbert et al. 2005) (Figure 3). However, if the ambient temperature falls below the depressed freezing temperature, known as the gap endpoint, ice growth resumes in distinctive explosive burst patterns (Scotter et al. 2006). The severity of this uncontrollable ice growth is correlated with the thermal hysteresis activity level where larger temperatures gaps lead to increasingly dangerous explosive ice growth (Yu et al. 2010). Due to the dangers of this uncontrollable ice growth, an antifreeze protein with high thermal hysteresis is a defining feature of freeze-avoidant organisms (Gilbert et al. 2005; Middleton et al. 2012). With inevitable freezing, freeze-tolerant organisms generally produce antifreeze proteins with low thermal hysteresis activity to minimize explosive ice growth damage (Garnham et al. 2008; Scotter et al. 2006).

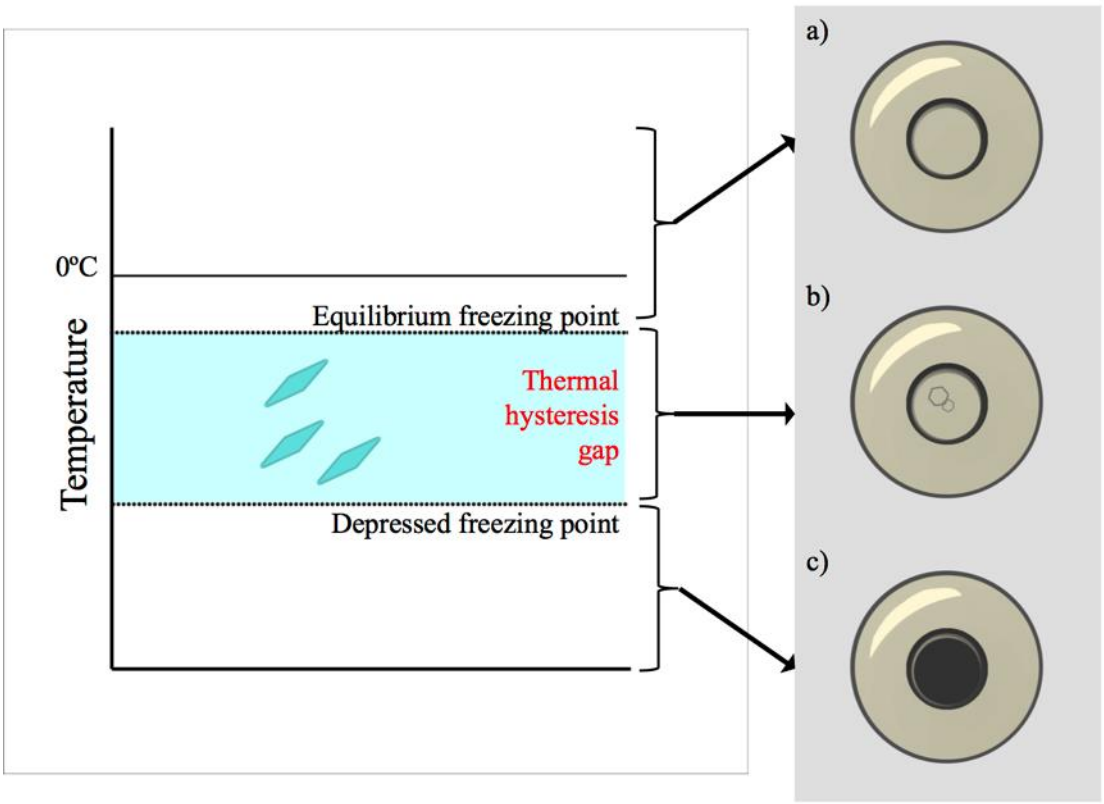


Figure 3. Antifreeze protein induced freezing (thermal) hysteresis. During freezing hysteresis, the freezing point becomes depressed. This creates a temperature gap between the original equilibrium freezing point (also known as the equilibrium melting point) and the depressed freezing point. In this gap, ice crystals neither grow nor melt. (a) No ice crystals are observed above the equilibrium freezing point. (b) Individual ice crystals and its morphology can be observed in the thermal hysteresis gap. (c) Complete freezing (opaque) occurs beyond the depressed freezing point (Figure taken from Lorv et al. 2014)

Due to the non-colligative nature of thermal hysteresis activity, thermal hysteresis values are inconsistent and vary greatly between organisms making comparisons difficult (Sun et al. 1995). Furthermore, due to the non-linear relationship between concentration and activity, thermal hysteresis activity is reported as maximal thermal hysteresis values. For instance, insect antifreeze proteins have up to 20-fold greater activity than fish equivalents. These thermal hysteresis values range from 3 to 6°C and 0.4 to 2.0°C respectively (Doucet et al. 2000). Most plant and bacterial antifreeze proteins are consistently lower than either insect or fish proteins with a range of 0.15 to 0.7°C and 0.1 to 0.3°C respectively (Doucet et al. 2000; Hobbs et al. 2011; Lin et al. 2011; Mok et al. 2010; Sidebottom et al. 2000; Sun et al. 1995).

The second antifreeze activity is ice recrystallization inhibition, which is characterized as the prevention of smaller ice crystals from recrystallizing into larger ice crystals at high subzero temperatures or from freeze thaw cycles (Walker et al. 2006; Wilson et al. 2006). At high subzero temperatures, antifreeze proteins have been speculated to inhibit recrystallization by preventing water movement between ice crystals and ultimately destabilizing small ice crystal grains (Yu et al. 2010). Antifreeze protein-ice interaction stabilizes these small ice crystals making them more energetically favourable and stable than larger ice crystals. In order to minimize frost damage, freeze-tolerant organisms usually produce antifreeze proteins focusing on ice recrystallization inhibition (Griffith and Ewart 1995; Wilson et al. 2006).

Both antifreeze activities are hypothesized to use the same mechanism, adsorption inhibition (Kawahara 2008; Raymond and DeVries 1977; Yu et al. 2010). For these

protein-ice interactions to occur, antifreeze proteins use a flat, hydrophobic surface that is complementary to the oxygen spacing of ice crystal planes. This surface complementary is determined by the spacing of solvent exposed amino acid residues. Variations in these residues can significantly affect the affinity and location of antifreeze protein binding, as well as the active state of the protein (Garnham et al. 2012; Middleton et al. 2012; Modig et al. 2010). These differences are especially important as they differentiate between moderate and hyperactive antifreeze proteins (Celik et al. 2010; Garnham et al. 2008; Garnham et al. 2011a; Gilbert et al. 2005; Middleton et al. 2012).

Upon adsorption of the antifreeze protein on to an ice crystal surface, further ice growth is inhibited via the Kelvin effect (Raymond and DeVries 1977). According to the Gibbs-Thompson equation, after protein adsorption, it is energetically unfavourable for additional water molecules to bind to the ice surface. One proposed model, the mattress model, indicates that free water molecules are continuously added in the limited space between bound antifreeze proteins until a maximally curved surface is formed (Bouvet and Ben 2003; Pertaya et al. 2007). Upon reaching this point, it becomes energetically unfavourable to add additional water molecules, inhibiting the formation of an extra water layer and, consequently, further ice growth.

However, adsorption inhibition makes one major assumption, i.e. that protein adsorption is irreversible (Pertaya et al. 2007). Reversible protein adsorption is problematic since transient desorption can present new sites for water molecule binding leading to ice growth, especially when water molecules are usually in a very large molar excess (Pertaya et al. 2007). Furthermore, empirical evidence of the non-colligative nature of antifreeze proteins supports irreversible binding since a colligative thermal

hysteresis of 0.5°C would require a local concentration of 300 mg/ml of antifreeze proteins. However, one research group has argued that antifreeze proteins reversibly bind to ice crystals prior to reaching maximal thermal hysteresis (Pertaya et al. 2007). Extrapolating their data, using a docking simulation and assuming that a slow exchange of antifreeze proteins is occurring, this group proposed that antifreeze protein binding is quasi-permanent taking more than a week to be exchanged.

Ice crystal shaping and structuring, although a minor activity, is another defining feature of antifreeze proteins arising as a byproduct of adsorption inhibition. Ice crystal morphology is altered based on the location of antifreeze adsorption on the ice crystal surface (Cheng et al. 2007; Kawahara 2008; Sidebottom et al. 2000; Sun et al. 1995). Binding to the primary prism plane yields a hexagonal or hexagonal bipyramid. For example, binding to both the primary prism and basal plane can shape ice morphology to a hexagonal plate or a lemon shape (Lin et al. 2012; Scotter et al. 2006; Wilson et al. 2006). However, this activity is not present in all antifreeze proteins. Some bacterial antifreeze proteins, such as the ones from *Marinomonas primoryensis* and *Flavobacterium xanthum*, lack this ice structuring capability (Gilbert et al. 2005; Kawahara et al. 2007).

1.1.2.1 Known Bacterial Antifreeze Proteins

Over the last two decades, a variety of antifreeze activities have been found in bacteria, especially those indigenous to cold habitats (Gilbert et al. 2004; Kawahara et al. 2004; Kawahara et al. 2007; Sun et al. 1995; Wilson and Walker 2010; Yamashita et al.

2002). Unfortunately, most of the reported studies have focused only on basic characterization of observed antifreeze activity. Since bacteria are usually freeze-tolerant microorganisms, these studies have focused on the presence of ice recrystallization inhibition and ice structuring activity.

The first bacterial antifreeze activity ever documented belonged to the psychrophile *Micrococcus cryophilus* and the soil bacterium *Rhodococcus erythropolis* with thermal hysteresis values of 0.29°C and 0.35°C respectively (Duman and Olsen 1993). Since then, a few bacterial antifreeze proteins have been documented and characterized at the protein level.

The first isolated and second most well characterized bacterial antifreeze protein at the protein level was from *Pseudomonas putida* GR12-2 (Sun et al. 1995). This plant growth-promoting soil bacterium is capable of low temperature growth and survival at freezing temperatures up to -50°C without exogenous cryoprotectants. This freezing tolerance is attributed to the bacterium's secreted antifreeze protein, AfpA, which was found to have a moderately low thermal hysteresis value of 0.13°C. When ice crystal morphology was observed, AfpA was documented to shape ice into hexagonal bipyramids.

Following isolation, AfpA from *Pseudomonas putida* GR12-2, was characterized as a 164 kDa protein containing both sugar and lipid moieties (Xu et al. 1998). Based on its primary protein sequence, this lipoglycoprotein was predicted to have 7 N-glycosylation, 2 O-glycosylation and 20 myristolation sites (Muryoi et al. 2004). In addition, Xu and colleagues (1998) reported that at least 72 kDa of the original 164 kDa

antifreeze protein was composed of glycans. Furthermore, the amount of lipidation has yet to be determined. Interestingly, when the antifreeze gene *afp*A was expressed in *Escherichia coli*, the recombinant protein retained sugar and lipid moieties (Muryoi et al. 2004). This high degree of modification is unusual for a bacterial protein and requires further work to elucidate the functional role of these moieties.

At the protein level, AfpA lacks a conserved canonical N-terminal signal peptide (Muryoi et al. 2004). However, this antifreeze protein also features five Gly-X-Gly-X-Asp calcium binding motifs flanked by autotransporter domains. Another motif, Asp-U-U-U (where U represents any hydrophobic residue), was found near the C-terminus of AfpA. Due to these motifs, Muryoi et al. (2004) speculate that AfpA may be secreted by either a hemolysin-like or type V autotransporter secretion system. Based on the antifreeze gene sequence, this research group was also able to predict a Shine Dalgarno sequence as well as the transcription start site.

Another bacterial antifreeze protein was found in *Moraxella* sp. isolates from Antarctica (Yamashita et al. 2002). This 52 kDa lipoprotein was measured to have a thermal hysteresis value of approximately 0.2°C at a concentration of 1.5 mg/ml and produces hexagonal ice crystals. Another bacterial antifreeze protein (*fl*AFP) was recently isolated from *Flavobacterium xanthum* (Kawahara et al. 2007). This 59 kDa cytoplasmic antifreeze protein was found to have a thermal hysteresis activity of 0.04°C when present in a cell extract with a protein concentration of 0.7 mg/ml, and lacks ice structuring capabilities. When purified by size exclusion chromatography and supplemented with 0.5 M malate, *fl*AFP produced a thermal hysteresis value of 5.2°C. This antifreeze activity is

substantially greater than a typical bacterial antifreeze protein and is closer to insect antifreeze proteins indicating its hyperactivity.

To date, the most well characterized bacterial antifreeze protein, at the protein level, was isolated from *Marinomonas primoryensis* formerly classified as *Marinomonas protea* (Gilbert et al. 2005). This Antarctic bacterium expresses a large hyperactive protein (>1000 kDa), *Mp*AFP. Similar to *fl*AFP, *Mp*AFP has high thermal hysteresis activity, 2°C at 0.1 mg/ml, while also lacking ice-structuring capabilities (Garnham et al. 2008). This protein is very similar to insect antifreeze proteins with its other hyperactivity characteristics such as deadly ice burst patterns upon reaching the gap endpoint and binding to both primary prism as well as basal ice planes (Garnham et al. 2011a; Gilbert et al. 2005). Unlike other bacterial antifreeze proteins, *Mp*AFP hyperactivity suggests a freeze avoidance strategy that is optimal in its habitat, a thermally buffered Antarctic lake ranging between -1° to 1°C (Gilbert et al. 2004).

The distinguishing feature of *Mp*AFP is its calcium ion dependency. From antifreeze assays and structural studies, *Mp*AFP requires calcium ions to fold into its active state (Garnham et al. 2011a; Gilbert et al. 2005; Guo et al. 2013). Removal of calcium ions or addition of EDTA eliminates antifreeze activity. Out of the five distinct domains in *Mp*AFP, two (Region II and IV) have conserved tandem repeats that are characteristic of antifreeze proteins (Garnham et al. 2008; Lin et al. 2011; Middleton et al. 2012). Region II has 120 tandem repeats with 104 conserved amino acids per repeat while region IV has 13 tandem repeats consisting of 19 amino acids per repeat. Both

regions have been shown to rely on calcium ions for protein folding and stabilization. The latter region has been identified as the antifreeze active domain (Guo et al. 2012).

The consensus sequence of each tandem repeat in region IV is X-Gly-Thr-Gly-Asn-Asp- X-U-X-U-Gly-Gly-X-U-X-Gly-X-U-X where X and U residues represent hydrophilic and hydrophobic residues, respectively (Garnham et al. 2008; Garnham et al. 2011a). This region was modeled as a left handed β-roll with one repeat representing a β-loop. The X-Gly-Thr-Gly-Asn-Asp motif of each loop in this model creates calcium-binding turns along one side of region IV (Figure 4). The remaining residues in each repeat are necessary to maintain the structure and rigidity of the protein by sustaining the hydrophobic core. Calcium ions stabilize the structure by binding to inward facing residues of calcium-binding turns, and are locked into place by a Glu residue in the last loop (Garnham et al. 2008).

The alignment of this calcium-binding motif creates a flat, hydrophobic surface (Garnham et al. 2011a). At the calcium binding turns, Thr and Asn residues are solvent exposed and responsible for antifreeze activity, similar to the Thr-X-Thr motifs in insect antifreeze proteins (Garnham et al. 2008; Lin et al. 2011). With pitch changes introduced by calcium binding, these Thr and Asn residues form a flat surface where oxygen atoms are ordered in long arrays spaced 7.4 Å by 4.6 Å apart (Garnham et al. 2011a). This surface is the site of ice crystal interaction. Any disruptions to the flatness or ordering, such as by site-directed mutagenesis of Thr to a Tyr residue, can lead to a significant activity loss.

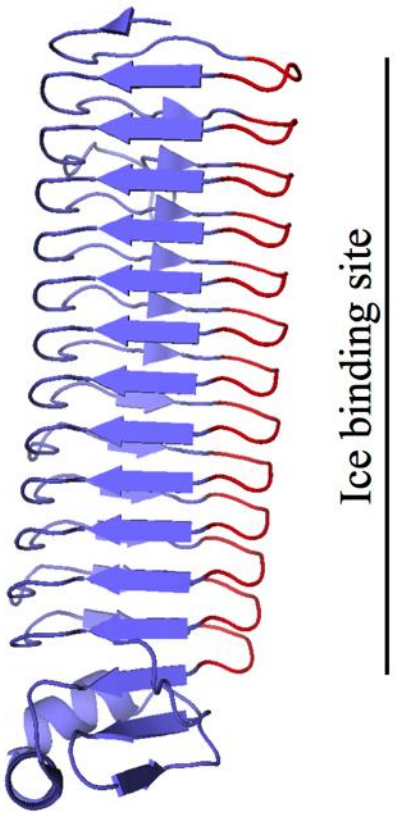


Figure 4. Ice binding site of *Mp*AFP region IV. Alignment of 13 calcium binding motifs (X-Gly-X-Gly-X-Asp) along one side of domain creates a flat, hydrophobic ice binding surface (Figure taken from Lorv et al., 2014)

The arrangement of inward facing Gly residues between Thr, Asn, and X residues in these turns create troughs on the surface (Garnham et al. 2011a). Within these troughs, *Mp*AFP maintains anchored clathrate waters (i.e., water molecules arrangement in an ice-like lattice) via the relative hydrophobicity of the surface and hydrogen bonding to solvent exposed residues. Through molecular docking simulations, these anchored clathrate waters have been implicated to be involved in direct ice interactions for binding (Hakim et al. 2013; Modig et al. 2010; Smolin and Daggett 2008).

Interestingly, region IV accounts for only 2% of the full *Mp*AFP while the remaining regions are non-ice binding (Garnham et al. 2008). Two recent studies of region II suggest that *Mp*AFP may not primarily be an antifreeze protein, but an ice binding adhesion protein (Guo et al. 2012; Guo et al. 2013). Strong ice binding by region IV, represented by high thermal hysteresis, allows the bacterium to bind to ice covering a lake while maintaining a safe distance with region II. This allows the bacterium to remain in the more nutritious and oxygen rich zone of the lake. Consequently, the bacterial antifreeze protein anomaly, *Mp*AFP, has been reclassified as an ice binding adhesion protein (Vance et al. 2014).

1.1.3 Ice Nucleation Proteins

Ice nucleation proteins are characterized by their ability to initiate heterogeneous ice formation at high subzero temperatures by providing a template for free water molecules to bind in an orderly manner to ice crystal nuclei (Kawahara 2008; Xu et al. 1998). To achieve this activity, ice nucleation proteins use ice binding site(s), very

similar to antifreeze proteins, which are complementary to ice crystal planes (Garnham et al. 2011b; Graether and Jia 2001).

Ice nucleation activities have been documented in many bacteria commonly found to be gram negative, epiphytic and pathogenic (Kawahara 2008). They may be either psychrophiles or mesophiles. Although counter-intuitive, ice nucleation activity has been speculated to assist in freeze tolerance survival (Xu et al. 1998). When transported to the outer membrane, ice nucleation proteins are suggested to direct ice formation externally into the extracellular space (Kajava and Lindow 1993; Xu et al. 1998). This localization of ice formation allows the bacterium sufficient time to physiologically adapt to inevitable freezing.

Depending on the level of ice nucleation activity, these proteins may be divided into three different classes (Turner et al. 1991). Ice nucleation activity is usually reported as T_{50} values, which is the temperature at which 50% of solution droplets are observed to freeze. Class A (type I) ice nucleation proteins are the most active and thought to be composed of lipoglycoprotein aggregates anchored to bacterial outer membranes with phosphatidylinositol (Kozloff et al. 1991; Turner et al. 1991). This group has the highest activity with T_{50} values up to -2°C. Two strong, native class A ice nucleators belong to *Pseudomonas syringae* and *Pseudomonas borealis* DL7 with an activity of -2 and -3.7°C, respectively (Maki et al. 1974; Wilson et al. 2006). Although *Pseudomonas syringae* has many variants of ice nucleation proteins (InaK, InaQ, InaV, and InaZ), the activities of these variants are similar to each other (Kajava and Lindow 1993; Kawahara 2008; Li et al. 2012a). Class B (type II) ice nucleation proteins are speculated to be composed of

glycoprotein aggregates with moderate ice nucleation activity and T_{50} values around -4.5°C (Turner et al. 1991). Class C (type III) ice nucleation proteins are composed of protein aggregates that are up to 1000 kDa in size. This class has the weakest ice nucleation activity where T_{50} is less than -8.0°C. Ice nucleation proteins from *Pseudomonas fluorescens* KUAF-68 and *Flavobacterium* sp. GL7 belong in this class with T_{50} values of -10.6 and -8°C respectively (Kawahara et al. 2004; Wilson et al. 2006).

However, not all ice nucleation active bacteria fall distinctly in these classes. For some bacteria, ice nucleation activity is exerted by secreted liposomes known as extracellular ice nucleating material (Kawahara et al. 1993; Muryoi et al. 2003; Obata et al. 1993; Phelps et al. 1986). These extracellular ice nucleating materials were speculated to have ice nucleation proteins embedded into liposomes rather than the outer bacterial membrane. When observed, ice nucleation patterns of these extracellular materials matched all three protein classes. Therefore, ice nucleation proteins are not limited to anchoring to cell membranes and can be secreted.

The typical ice nucleation protein is usually large, multimeric, hydrophilic and anchored to the outer cell membrane (Kawahara 2008). Monomeric subunits can range from 120 to 150 kDa in size. As a structurally homologous class of protein, ice nucleation proteins are composed of three domains: an N-terminal domain, a central repeating domain (CRD) and a C-terminal domain. Variations between ice nucleation proteins usually occur in the CRD with differences in amino acid sequence and amount of sequence repetition.

The N-terminal domain is relatively hydrophobic and globular (Kajava and Lindow 1993). Comprising 15% of the protein, this domain is hypothesized to be involved with binding to sugars, lipids, phospholipids and other ice nucleation proteins (Kajava and Lindow 1993; Kozloff et al. 1991). Anchoring to cell membranes has been documented to greatly influence ice nucleation activity because it allows the ice nucleation protein to aggregate and organize into assemblies with greater activity (Govindarajan and Lindow 1988). Targeted destabilization of cell membrane by ozone treatment was shown to decrease ice nucleation activity in *Pseudomonas syringae* from -2.8 to -7.3°C (Sarron et al. 2013). However, this activity could be recovered by phospholipid supplementation supporting the importance of lipids on ice nucleation protein (Govindarajan and Lindow 1988; Yu et al. 2013). Recently, Li and colleagues (2012a) confirmed the anchoring of ice nucleation protein InaQ from *Pseudomonas syringae* to the cell membrane using its N-terminal domain. Truncation of the transmembrane sequences in this domain resulted in a loss of 93% surface anchoring while elongation increased it. This domain is speculated to anchor via a mannose-phosphatidylinositol either through N or O linkage with Asp, Ser and Thr residues (Kawahara 2008).

The largest portion of most ice nucleation proteins resides in the central repeating domain, which contains the proposed ice-binding site for protein-ice interaction (Kawahara 2008). In this domain there are three levels of repeat fragments (Kajava and Lindow 1993). Comprised of 48 residues, the largest fragment usually has the highest sequence conservation. In InaZ from *Pseudomonas syringae*, the 48 residue fragment has a consensus sequence of Ala-Gln-Glu-Gly-Ser-Asn-Leu-Thr-Ala-Gly-Tyr-Gly-Ser-Thr-

Gly-Thr-Ala-Gly-Ala-Asp-Ser-Ser-Leu-Ile-Ala-Gly-Tyr-Gly-Ser-Thr-Gln-Thr-Ser-Gly-Ser-Glu-Ser-Ser-Leu-Thr-Ala-Gly-Tyr-Gly-Ser-Thr-Gln-Thr. This fragment can be further broken down into three slightly conserved fragments comprised of 16 residues. In the ice nucleation protein from *Pseudomonas borealis* the consensus sequence of these fragments is Gly-Tyr-Gly-Ser-Thr-X-Thr-Ala-X-X-X- Ser-X-Leu-X-Ala (Garnham et al. 2011b). It is speculated that these residues are responsible for forming a circular loop with β turns (Kumaki et al. 2008). Once again these 16 residue fragments can be broken down into two unconserved octapeptide fragments speculated to dictate protein bending during folding (Kajava and Lindow 1993).

From a recent molecular simulation study, an ice nucleation protein from *Pseudomonas borealis* was modeled as a right-handed β-helical structure (Garnham et al. 2011b). Stabilization of this protein occurs using internal Gly and Ser ladders as well as dimerization via desolvation of an exposed Tyr ladder. This dimerization is essential as alignment of two tetrapeptides, i.e. side one from monomer A and side two from monomer B, creates a potential flat ice-binding surface (Figure 5). Due to the repetitive nature of the CRD, this combined octapeptide surface (i.e., Ser- Leu-Thr-Ala and Thr-Gln-Thr-Ala) is aligned along the length of the β-helix creating Thr-X-Thr motifs. These motifs are then speculated to create an ice-binding site similar to *Mp*AFP and insect antifreeze proteins with Thr and Ser as the ice binding residues (Garnham et al. 2011b; Graether and Jia 2001; Kajava and Lindow 1993). Although ice nucleation proteins are hydrophilic overall, these flat surfaces are considered sufficiently hydrophobic for ice crystal interactions.

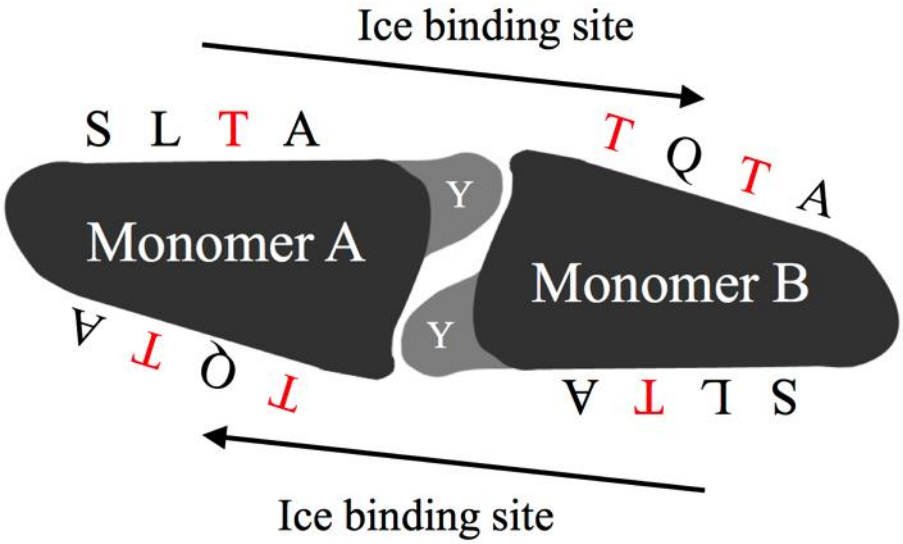


Figure 5. Ice binding site of an ice nucleation protein from Pseudomonas borealis. Dimerization via tyrosine ladders creates a relatively flat, hydrophobic ice binding surface with Thr-X-Thr motifs (Figure taken from Lorv et al., 2014)

The last 15% of an ice nucleation protein is the C-terminal domain. A brief test by Li and colleagues (2012a) revealed that the C-terminal domain is not involved in protein anchoring. Its function remains elusive and requires further study (Kajava and Lindow 1993; Kawahara 2008).

1.1.4 Simultaneous Antifreeze and Ice Nucleation Activity

Some ice active bacteria can have both antifreeze and ice nucleation activity. Cooperatively expressing both activities has been speculated to enhance freeze tolerance (Kawahara et al. 2004; Xu et al. 1998). Hypothetically, this increased tolerance occurs because ice nucleation proteins direct ice formation externally while antifreeze proteins maintain the ice as small ice crystal grains through ice recrystallization inhibition (Xu et al. 1998).

Interestingly, these activities are expressed as either separate antifreeze and ice nucleation proteins, or one protein with both ice crystal controlling activities. For bacteria such as *Pseudomonas fluorescens* KUAF-68 and *Pseudomonas borealis* DL7, separate proteins have been found to be responsible for these activities (Kawahara et al. 2004; Wilson et al. 2006). However, ice crystal controlling proteins with both activities have been found in at least two bacteria. The protein AfpA from *Pseudomonas putida* GR12-2 has both a weak thermal hysteresis activity of 0.11°C and a class C ice nucleation activity of -11°C (Muryoi et al. 2004; Sun et al. 1995). Furthermore, AfpA has low sequence identity and similarity to other ice nucleation proteins. In *Xanthomonas campestris*, an ice nucleation protein was discovered to have ice recrystallization inhibition activity (Nada et

al. 2010). Although this protein could not suppress ice growth completely, it was able to decrease ice growth rate along the c-axis.

Despite different activities, the latest hypothesis suggests that both ice crystal-controlling proteins are very similar to each other. Both protein classes use the same ice binding mechanisms with the molecular size of the protein complex (including protein aggregation) dictating the protein activity (Garnham et al. 2011b; Graether and Jia 2001; Xu et al. 1998). Small proteins, 50 kDa or lower, usually have antifreeze activity (Garnham et al. 2011b). Larger protein complexes, such as each ice nucleation class, usually exhibit ice nucleation activity (Garnham et al. 2011b). Another factor hypothesized to influence ice crystal controlling activity is protein concentration. When type III fish antifreeze protein exceeded a critical aggregation concentration (CAC), the protein complex showed ice nucleation activity (Du et al. 2006). At this CAC, antifreeze proteins become optimally adsorbed to ice crystals. However, beyond this concentration, excess antifreeze protein monomers can oligomerize creating larger protein complexes with exposed ice binding sites. These protein complexes then become ice nucleators, supporting aggregation as a driving factor for ice nucleation protein activity.

1.2 Ecology of Ice Crystal Controlling Protein Producing Bacteria

With up to 85% of Earth's biosphere experiencing temperatures below 15°C, cold and freezing stress is one of the most common difficulties to overcome (Margesin and Miteva 2011; Wilson and Walker 2010). In these cold habitats, a variety of ice active bacteria, some with ice crystal controlling activities, have been identified. Due to the

prevalence of cold stress, the ability to express ice crystal controlling activity is potentially scattered across the globe (Margesin and Miteva 2011; Wilson et al. 2006).

1.2.1 Terrestrial Environments

One common cold habitat for ice active bacteria is the soil. Unlike aquatic environments, surface soil can present harsher conditions with higher temperature fluctuations and lower subzero temperatures (Gilbert et al. 2005; Margesin and Miteva 2011; Middleton et al. 2012; Walker et al. 2006). Several ice active bacteria have been found within these conditions. Although most are poorly characterized, these bacteria have been found to express ice crystal controlling activities.

From Antarctic soil in McMurdo Dry Valley, one research group found that eleven bacterial isolates were positive in ice nucleation activity as well as six strains with antifreeze activity (Kawahara et al. 2004). They classified the bacterial isolate with the greatest ice nucleation activity as *Pseudomonas antarctica*. In this study, the greatest antifreeze active bacterial strain was identified as *Moraxella* sp. with a thermal hysteresis activity of 0.08°C. Another antifreeze active strain is *Pseudomonas fluorescens* KUAF-68 with an antifreeze activity of 0.03°C. When compared to antifreeze activity from Antarctic lake bacterium *Marinomonas primoryensis*, these soil isolates with low thermal hysteresis activity are indicative of a freeze tolerance strategy by expressing antifreeze proteins with low thermal hysteresis values (Gilbert et al. 2004; Kawahara et al. 2004).

In warmer Arctic soils, with their greater water and nutrient bioavailability, pseudomonads have been found to represent 60% of all culturable isolates (Margesin and

Miteva 2011). Of these bacterial isolates, one soil bacterium *Pseudomonas putida* GR12-2 was found in the Canadian High Arctic where temperatures are around 5°C during spring. These temperatures are known to induce antifreeze protein expression; not surprisingly arctic soils contain ice active bacteria.

Microbial communities found in permafrost, defined by subzero temperature conditions persisting for two or more years, were found to be resistant to freeze thaw cycles (Margesin and Miteva 2011; Wilson and Walker 2010). Whether this resistance is primarily from ice crystal controlling protein expression is unclear, but it suggests that some isolates may be ice active. For example two isolates from the genera *Exiguobacterium* and *Psychrobacter* were shown with ice nucleation activity (Margesin and Miteva 2011).

At high mountain altitudes, environmental conditions consist of high temperature fluctuations, high precipitation and freeze-thaw stress (Margesin and Miteva 2011). While at lower altitudes, very quick freeze-thaw cycles can be the norm; in Alberta, Canada, Chinook winds from the Rockies can create freeze thaw cycles ranging from -35°C to 35°C (Walker et al. 2006). This fluctuation in temperature can occur quickly with an increase of up to 20°C within an hour. Not surprisingly, freeze-thaw resistant bacterial isolates were found in this area.

After exposing these isolates to 48 freeze-thaw cycles, cycling between -18 and 5°C, the most freeze resistant isolates were gram negative bacteria, either *Buttiauxella* or *Chryseobacterium* (Walker et al. 2006). Other less freeze resistant strains were either *Acinetobacter* or *Enterococcus*. However, freeze-thaw resistance does not equate to

expression of ice crystal controlling proteins. In one study, biofilm production was also shown to assist in freeze-thaw resistance (Wu et al. 2012). To investigate an ice crystal controlling protein's role in freeze-thaw resistance, a follow up study isolated different bacteria strains in the same soil samples via ice affinity binding (Wilson et al. 2006). In this case, the dominant bacterial isolate belonged to the genus *Paenibacillus*. The difference in isolates from both studies suggests that ice crystal controlling protein assists, but is not the dominant factor in freeze-thaw stress survival.

In temperate lakeside soils, the microbial community endures warmer average temperatures, but seasonal temperature fluctuations can still promote the expression of ice crystal controlling proteins (Mok et al. 2010). After investigating two Canadian lakes, Wilson and colleagues (2012) were able to find bacterial isolates expressing antifreeze activity. The two lakes differed in climate. The first lake was Daring Lake, a lake in Northwest Territories that endures harsh, long winters with great temperature fluctuations in the spring. The second lake was Gould Lake near Kingston, Ontario with lower temperature fluctuations and a milder average temperature.

Using freeze-thaw cycles to isolate resistant bacterial strains, this group of researchers was able to identify bacterial isolates with ice crystal controlling activities. Unlike the previous freeze-thaw study, the dominant genera, *Bacillus* and *Pseudomonas*, showed ice crystal controlling activity (Wilson et al. 2006; Wilson et al. 2012). Daring Lake isolates *Chryseobacterium piscium* and *Pseudomonas* sp. DL13 displayed ice recrystallization inhibition (Wilson et al. 2012). *Bacillus* sp. G1a1 had ice nucleation activity while *Buttiauxella* sp. G2b1 had ice structuring capabilities. However, not all soil bacteria showed these activities. When bacteria in these soil samples isolated with ice

affinity binding, the recovered strains belonged to the genera *Pseudomonas*, *Stenotrophomonas, Chryseobacterium, Flavobacterium* and *Acinetobacter* (Wilson et al. 2006).

The inconsistency between antifreeze activity, ice nucleation activity and freeze-thaw survival suggests that ice crystal controlling protein expression is not the dominant strategy for survival against freezing stress. These bacteria may employ a variety of complex mechanisms and, as a community, sharing of resources between strains can aid with cold and freeze stress (Wilson et al. 2006; Wu et al. 2012).

1.2.1.1 Plants

One common habitat for ice nucleation active bacteria is plant surfaces (Maki et al. 1974). Whether these plants surfaces are home to antifreeze active bacteria is unclear. However, the ice active phenotype can change depending on environmental stimuli such as light intensity and humidity (Cambours et al. 2005; O'Brien and Lindow 1988). Also, ice nucleation activity can vary depending on the bacterial strain and its host plant genotype (O'Brien and Lindow 1988). Another bacterium, the extensively studied ice nucleation active *Pseudomonas syringae* was also isolated from plant leaves (Li et al. 2012a). This bacterium is speculated to use ice nucleation to facilitate ice formation to injure the host plant for subsequent infection (Hirano and Upper 2000). Such plant frost injury is closely correlated to the presence of ice active bacteria. Decreases in ice active bacteria populations have been found to reduce plant injury especially to frost sensitive plants.

Besides plant surfaces, the soil surrounding cold-inhabiting plant, known as the rhizosphere, is also a home to ice active bacteria. Similar, to other soil habitats, these rhizospheric bacterial communities experience seasonal fluctuations and frost damage. Not surprisingly, an antifreeze active bacteria such as *Pseudomonas* sp. UW4 was isolated from the rhizosphere of common reeds in Waterloo, Ontario (Cheng et al. 2007).

1.2.2 Aquatic Environments

Despite representing 71% of the Earth's biosphere with many psychrophilic and piezophilic bacteria, the deep sea has not been thoroughly investigated for ice active bacteria (Margesin and Miteva 2011). Closer to the surface, ice nucleation active bacteria, although with weak activity, have been found entrapped in either Antarctic or Arctic sea-ice (Junge and Swanson 2008). Similar to soil bacteria exposed to Chinook winds, deep sea bacteria may use other cold adaptations strategies such as optimized cell membrane composition, cryoprotectant production and sporulation.

Most ice active bacteria from aquatic environments have been found in and around lakes (Gilbert et al. 2004; Wilson et al. 2012). Unlike arctic and alpine lakes, ice active bacteria were identified in some Antarctic saltwater lakes. These bacterial isolates include *Marinomonas primoryensis*, an antifreeze protein producing bacterium, a *Pseudoalteromonas* isolate and a *Pseudomonas fluorescens* isolate (Garnham et al. 2008; Gilbert et al. 2004). However, these isolates were only 19 isolates out of 186 tested isolates reinforcing the idea that dominant freezing stress strategy may not include ice crystal controlling protein expression.

1.3 Gene Evolution

The relationship between environmental conditions and the expression of ice crystal controlling proteins is unclear. In areas that experience high cold or freezing stress, such as the deep sea, there has been few reports of ice active bacteria. However, this may be due to a lack of studies specifically looking for ice crystal controlling proteins. Even in areas with identified ice active bacteria, the dominant and most freeze resistant isolates may not be ice active bacteria (Gilbert et al. 2004; Walker et al. 2006; Wilson et al. 2006; Wilson et al. 2012). This is not to say that ice crystal controlling proteins are not used for freeze resistance. Environmental factors do influence the development of ice crystal controlling protein expression. Freeze-thaw cycling can select for ice active bacterial isolates indicating some protection from freezing (Wilson et al. 2006). With Antarctic lake bacteria, antifreeze activity correlated with lake salinity (Gilbert et al. 2004). Ice active bacteria were found in saline but not freshwater lakes. This salinity was speculated to be the driving force behind antifreeze protein expression since saline lakes usually have higher temperature fluctuations. The structural diversity of antifreeze proteins across many organisms suggests convergent evolution of these proteins that arose as numerous independent events through environmental selection (Lin et al. 2011). This frequent evolution of antifreeze proteins emphasizes the importance of these proteins for survival in cold habitats.

The low abundance of ice active bacteria may be due to community effects where one ice active microbial member is sufficient for overall freezing tolerance. In one case, a *Chryseobacterium* sp. with ice recrystallization inhibition activity was able to increase the freeze-thaw survivability of *Enterococcus* sp. when the two were inoculated together

(Wilson et al. 2006). Therefore, the development of ice crystal controlling proteins allows some ice inactive bacteria to expand and fill different niches by providing some freeze resistance to the microbial community (Near et al. 2012).

1.4 Gene Regulation

Regulation of cold-based proteins is generally poorly understood, especially for ice crystal controlling proteins. Usually, these cold regulated proteins are seasonally induced by the onset of cold stress following a decrease in temperature (Chen et al. 2003; Hobbs et al. 2011; Nemecek-Marshall et al. 1993; Shivaji and Prakash 2010; Sun et al. 1995; Xu et al. 1998). For this cold induction to occur, the microbe needs to able to detect downshifts in temperature. One general hypothesis for cold detection involves a two-component system (Shivaji and Prakash 2010). As the ambient temperature decreases the cell membrane becomes rigid, triggering changes in a sensor protein. Following subsequent phosphorylation events, a response regulator indirectly upregulates transcription factors that stimulate expression of cold regulated genes. However, this scenario may not be the case for all cold-based proteins. Some cold shock protein homologues belonging to the cold shock protein A (CspA) family in *E. coli* are not cold inducible but are indirectly upregulated at low temperatures (Etchegaray et al. 1996).

Cold induction of ice crystal controlling proteins may be tied to cellular growth. AfpA from *Pseudomonas putida* GR12-2 was expressed following entry into the stationary phase, a growth phase associated with stress and cell maintenance (Kawahara et al. 2001). Similarly, *Pseudomonas syringae* had the highest detectable ice nucleation

activity after growth at low temperature with nutrient starvation (Nemecek-Marshall et al. 1993). This growth condition combination usually initiates entry into the stationary phase. However, nutrient starvation is not always required. When the trancriptome of *Pseudomonas putida* KT2440 was examined, Kim and colleagues (2013) found that growth with glycerol or fructose as the carbon source could stimulate the transcription of *rpoS*, a stationary phase sigma factor of RNA polymerase, as well as cold inducible cold shock proteins. This suggests that specific carbon sources may mimic cold temperature conditions. For *Pseudomonas fluorescens* MACK-4, optimal ice nucleation activity occurred following cold temperature growth with sorbitol, mannose or starch as the carbon source and peptone as the nitrogen source (Chen et al. 2003). In these cases, cold temperature incubation is the common theme with nutrition factors affecting activity levels of cold regulated proteins. However, whether the variations in ice nucleation activity are due to increased expression of the respective protein is unclear.

Although cold induction of ice crystal controlling proteins has yet to be studied, studies on cold inducible CspA from the mesophile *E. coli* has provided some insight into the cold inducible nature of a cold regulated protein. Not surprisingly, transcription of CspA is highly induced after a temperature decrease to 15°C with the highest protein expression levels occurring after 45 to 75 minutes (Fang et al. 1998; Tanabe et al. 1992). Upstream of the transcriptional start site, there are promoter elements differing by a few bases from the canonical ones (-10 and -35 regions) as well as an A-T rich upstream promoter (UP) element found to enhance transcription (Mitta et al. 1997; Tanabe et al. 1992). With these elements, the *cspA* promoter has been found to be very efficient at low temperatures and becomes stronger than the strong *E. coli lpp* promoter, is at 37°C (Mitta

et al. 1997). Despite high transcription rates at low temperatures, the *cspA* gene is also transcribed during growth in non-cold conditions suggesting cold induction is not solely regulated at the transcriptional level.

1.4.1 Post-Transcription Regulation

Protein regulation can occur beyond the transcriptional level. For some bacterial proteins, post-transcriptional regulation occurs via transcriptional or translational attenuation (Merino and Yanofsky 2005). Transcriptional termination involves a controllable transcriptional terminator within the target gene that prematurely terminates transcription. Translational termination occurs when an attenuator within the transcript prevents translation of the primary amino acid sequence. One pioneering example of this type of regulation is the *trp* operon (Merino and Yanofsky 2005; Yanofsky 1981). Amino acid starvation controls the translational attenuation of the *trp* operon genes (Yanofsky 1981). Within the polycistronic transcript and beyond the initial peptide leader sequence, there are four regions (1 to 4) that are either terminator or anti-terminator sites. During translation of the peptide leader sequence, amino acid starvation stalls the ribosome at various points near the four regions in the mRNA transcript. The stalled position influences the interaction of each of these regions with one another, which alters the secondary structure of the mRNA. Starvation of Trp and Arg residues can stall ribosomes at region 1 encompassing it. This causes region 2 and 3 to form a hairpin loop (Yanofsky 1981). Consequently region 4 is left unbound permitting ribosome access to the Shine Dalgarno sequence for translation of the remaining *trp* operon genes. In other conditions, region 3 and 4 remain bound together preventing ribosome access for further translation.

The *trp* operon regulation, known as a ribosome-mediated transcriptional attenuation is only one of many post-transcriptional attenuation systems (Merino and Yanofsky 2005). Other variations include protein-mediated, tRNA-mediated and small-molecule-mediated transcriptional attenuation. All these systems use mRNA to "sense" the respective molecule by binding to it, ultimately shifting the mRNA secondary structure to either be more or less susceptible to further transcription or translation. Since temperature regulation does not involve a tangible molecule, temperature dependent genes use a distinct subset of mRNA known as RNA thermometers (Kortmann and Narberhaus 2012). The basis for this type of regulation is that at higher temperatures mRNA secondary structures are denatured removing any physical repression of transcription or translation. This type of regulation has been seen in heat shock protein regulation (Kortmann and Narberhaus 2012). One problem with RNA thermometers is that at lower temperatures RNA secondary structures become more stable.

CspA from *E. coli* manages to bypass this problem and is regulated by a RNA thermometer. The key component is its unusually long 5' untranslated region (UTR) in the *cspA* mRNA (Fang et al. 1998; Fang et al. 1999; Jiang et al. 1996; Mitta et al. 1997). Interestingly, this 5'UTR is not limited to self-regulation of CspA but affects expression of other cold-based proteins as well. Overexpression of only 5' UTR from the *cspA* mRNA upregulates expression of other proteins such as cold shock protein homologues CspB, CspD, CspE as well as cold shock RNA helicase CsdA (Jiang et al. 1996).

Recently, Giuliodori and colleagues (2010) deduced that the secondary structure of the *cspA* transcript changes at the 5' UTR based on temperature. At 15°C, the 5' UTR forms a pseudoknot (Kortmann and Narberhaus 2012). This pseudoknot prevents the 5'

UTR from interacting with the *cspA* coding sequence, which reveals and allows ribosome access to either the Shine Dalgarno or downstream box sequence. The downstream box is an alternative translation initiator (Yamanaka and Inouye 2001). When the temperature increases to 37°C, the pseudoknot becomes unstable (Giuliodori et al. 2010). In this condition, the free 5'UTR interacts with a portion of the coding sequence, forming a structure that hides the Shine Dalgarno and downstream box sequence. Without either sequence, translation cannot occur (Mitta et al. 1997; Yamanaka and Inouye 2001). Interestingly, once *cspA* mRNA forms the 37°C structure it cannot revert back to the15°C structure (Giuliodori et al. 2010). For protein expression to occur, new mRNA transcripts need to be made at 15°C. The formation of these secondary structures is critical for regulation. Mutation or deletion of this 5' UTR can lead to constant high expression of CspA at 37°C (Fang et al. 1998; Fang et al. 1999). Similarly, deletion of this region can increase the transcription of the *cspA* gene (Fang et al. 1999).

As a protein used for cold acclimation, expression of CspA is only transient via regulation by a cold box sequence (Fang et al. 1998). This cold box sequence is known to bind to a repressor protein and is not directly related to cold induction (Yamanaka et al. 1999). When this cold box sequence is removed, an increase of CspA expression occurs even beyond the cold acclimation period (Fang et al. 1998). This cold box sequence is hypothesized to bind overexpressed CspA after the cold acclimation period (Yamanaka et al. 1999). When bound, it alters the mRNA structure from the cold induced state to the inactive 37°C state and plays a small part in cold regulation (Yamanaka et al. 1999). The lengthy 5' UTR may be an important factor for cold induction and has been found in two other cold regulated proteins: cold shock protein CspL from *Lactobacillus plantarum* and

TrmE from *Pseudomonas syringae* (Derzelle et al. 2002; Singh et al. 2009). Both these proteins are speculated to be post-transcriptionally regulated by temperature via their 5' UTR.

Although translational attenuation can regulate protein expression, attenuation may not be sufficient for complete repression of cold inducible proteins (Yanofsky 1981). Other regulation methods when combined with translational attenuation can be suitable for temperature dependent repression of cold inducible proteins. For CspA, mRNA stability can play a role in cold shock induction. At 37°C, the 5' UTR destabilizes the *cspA* transcript, quickly degrading it and preventing potential ribosome binding (Fang et al. 1998). While at 15°C, there is an increase in mRNA stability leading to ribosome binding. Overall, CspA is tightly regulated for cold induction at 15°C by increased transcription, increased mRNA stability and an mRNA secondary structure favoring Shine Dalgarno and downstream box accessibility (Giuliodori et al. 2010; Mitta et al. 1997). Whether other cold inducible proteins are regulated in a similar manner remains to be determined.

1.5 Research Goals

Around 15 years ago, a bacterial antifreeze protein, AfpA, was discovered and characterized in the plant growth promoting bacterium, *Pseudomonas putida* GR12-2 (Sun et al. 1995; Xu et al. 1998). This protein was found to assist the bacterium's survival at low subzero temperatures as a result of the protein's antifreeze and ice nucleation activity (Muryoi et al. 2004; Xu et al. 1998). As a cold and freeze associated protein, AfpA is known to be expressed at low temperature conditions (Xu et al. 1998). Notwithstanding extensive research, since that time, on cold shock proteins, the

mechanism for temperature regulation of the expression of these types of proteins is still unknown. The main goal of this research is to investigate how the cold-inducible antifreeze protein AfpA is regulated in *P. putida* GR12-2. In this study, the effect of different growth conditions on *afpA* transcription was tested using the *lux*AB indirect reporter gene system (Farinha and Kropinski 1990). Regulation at the posttranscriptional level has also been considered and explored using a bioinformatics approach.

In other organisms, the activities of ice crystal controlling proteins were attributed to an ice binding mechanism using a flat, hydrophobic ice binding site on the protein surface (Garnham et al. 2011a; Garnham et al. 2011b). Unlike bacterial *Mp*AFP and eukaryotic ice crystal controlling proteins, the antifreeze protein AfpA from *P. putida* GR12-2 has low antifreeze activities (Garnham et al. 2008; Sun et al. 1995). Since protein activity is highly correlated with the regularity in spacing of solvent accessible residues on the ice binding surface, a secondary aim in this study is to identify a putative ice binding site. This ice binding site of AfpA presumably plays an important role in the low temperature survival of the bacterium (Garnham et al. 2008; Garnham et al. 2011a; Lin et al. 2011). Thus, characterizing both the protein regulation and structural components should provide a better understanding of how a bacterium can adapt to low temperatures.

2.0 Material and Methods

2.1 Bacterial Strains, Plasmids and Growth Conditions

All bacterial strains, vectors, and plasmids used in this study are described in Table 1.

Escherichia coli DH5α (Invitrogen, Life Technologies, Carlsbad, CA) was used as a cloning host during the construction and maintenance of recombinant plasmids. *Escherichia coli* BL21 DE3 (Novagen, EMD Millipore, Billerica, MA) was used as a recombinant protein expression host designed for pET vector expression. Both *E. coli* strains were grown in Luria-Bertani (LB) broth (ThermoFisher Scientific, Waltham, MA). When appropriate, recombinant vectors were maintained with the appropriate antibiotic, either ampicillin (100 μg/ml) or kanamycin (50 μg/ml).

Wild-type *Pseudomonas putida* GR12-2 was the experimental bacterium used in this study (Sun et al. 1995). This bacterium was kindly provided by Dr. G. Brown, Agrium, Inc. (Saskatoon, SK). It was grown aerobically in tryptic soy broth (TSB) (ThermoFisher Scientific, Waltham, MA) at either 24°C or at 4°C. Growth at 4°C was used to induce antifreeze protein expression.

P. putida GR12-2 and *Pseudomonas* sp. UW4 (Duan et al. 2013) were used as the hosts for promoter activity experiments. Wild-type *Pseudomonas* sp. UW4 was grown at 30°C in TSB supplemented with ampicillin (100 μg/ml). These strains were cultivated in TSB. The constructs were selectively maintained with the addition of 50 μg/ml of kanamycin.

Table 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Description	Reference
Strain		
Escherichia coli DH5α	F^- φ80*lac*ZΔM15 Δ(*lac*ZYA$^-$ *arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk^-, mk^+) *gal*$^-$ *pho*A *sup*E44 λ$^-$ *thi*$^-$ 1 *gyr*A96 *rel*A1	(Grant et al. 1990)
Escherichia coli BL21 (DE3)	F^- ompT hsdS_B ($r_B^-m_B^-$) gal dcm (DE3)	(Studier and Moffatt 1986)
Pseudomonas putida GR12-2	Wild-type strain	(Sun et al. 1995)
Pseudomonas sp. UW4	Wild-type strain	(Duan et al. 2013)
Vectors		
pET30b (+)	N-terminal His•Tag® /thrombin/ S•TagTM /enterokinase configuration plus an optional C-terminal His•Tag	Duca (2013)
pQF70	Broad host range vector with a promoterless *luxAB* reporter gene downstream of a multiple cloning site; Amp^r	(Farinha and Kropinski 1990)
pQF70K	Kanamycin resistance gene from pET30b(+) was inserted at NdeI site in pQF70; Amp^r, Km^r	This study
Plasmid constructs		
pETAFrag	pET30b (+) vector with a fragment of *afpA* gene from *P. putida* GR12-2 inserted in NdeI/XhoI site; uses C-terminal His tag	This study
pQKiaaM	pQF70K vector with the promoter region (850 bp) of *iaaM* gene from *P*. sp. UW4 cloned in SalI/NcoI site	This study
pQKAfp431	pQF70K vector with promoter region (431 bp) of *afpA* gene from *P. putida* GR12-2 cloned in BamHI/XbaI site	This study
pQKAfp131	pQF70K vector with promoter region fragment (131 bp) of *afpA* gene from *P. putida* GR12-2 cloned in BamHI/XbaI site	This study

2.2 General DNA Manipulation

2.2.1 Isolation of Genomic DNA

Genomic DNA from *P. putida* GR12-2 was isolated using a Wizard® Genomic DNA purification kit prepared by Promega Corporation (Madison, WI). DNA was isolated from one milliliter of overnight bacterial culture according to the manufacturer's instructions. The DNA pellet was rehydrated overnight at 4°C in 100 µl of rehydration solution (10 mM Tris-HCl, pH 7.4; 1 mM EDTA, pH 8.0) and later stored at -20ºC.

2.2.2 Isolation of Plasmid DNA

Plasmid DNA was routinely isolated by an alkaline lysis method available as a kit from BioBasic Inc. (Markham, ON). Between 5 and 10 ml of overnight culture was collected by centrifugation at 4500 *x g* for 10 minutes. Soluble plasmid DNA was separated from the genomic DNA and protein precipitate when centrifuged at 15,300 *x g* for two minutes. The purified plasmid DNA was then transferred to a new 1.5 ml microcentrifuge tube and stored at -20ºC.

2.2.3 DNA Amplification by PCR

Both the promoter region and the *afpA* gene itself was amplified from the genomic DNA of *P. putida* GR12-2. Most amplification reactions used different sets of primers, although some sets shared a reverse primer. Primer sets used in this study are listed in Table 2. EMD Millipore's (Billerica, MA) KOD Hot Start polymerase was used for each reaction. A typical 10 µl reaction mixture contained 10X PCR buffer, 0.2 mM dNTPs, 1 mM $MgSO_4$, 100 ng of genomic DNA as template, 300 nM of each primer and

one unit of DNA polymerase. The thermocycler was programed to run one cycle of 95ºC for two minutes then 25 cycles of a 20 second step at 95ºC, a 10 second step at the lowest primer Tm, and 70ºC set for 10s/kb (for 500 bp products) or 20s/kb (for 1.5 kb products). After all cycles, there was a final extension step at 70ºC for five minutes. PCR products were visualized on a 1% agarose gel for quantity and quality. These products were recovered by DNA extraction from an agarose gel.

2.2.5 DNA Extraction from an Agarose Gel

To visualize DNA quality as well as to separate DNA products from proteins and primers, DNA samples were run on agarose gels. After electrophoresis, the DNA band of interest was excised and placed in a 1.5 ml microcentrifuge tube. DNA was extracted and purified using a Wizard® SV gel and PCR clean-up system prepared by Promega Corporation (Madison, WI). DNA extraction was carried out according to the manufacturer's instructions. The concentration of purified DNA was measured using a NanoDrop 2000 UV spectrophotometer (ThermoFisher Scientific, Waltham, MA).

Table 2. Primers used to amplify kanamycin resistance gene, *afpA* gene and promoter region of the *afpA* gene

Primer Description	**Sequence (5' to 3')**
Kan^r forward primer (NdeI site)	TTTTTCATATGGTATGTAGGCGGTGCTACAG
Kan^r reverse primer (NdeI site)	TTTTACATATGATTTCAGGTGGCACTTTTCG
afpA gene forward primer; C-terminal His-tag (NdeI site)	AAAAACATATGCAATACGACAGCCCAATC
afpA gene fragment forward primer; C-terminal His-tag (NdeI site)	AAAAACATATGACCGCTGGCATCGAGTTC
afpA gene reverse primer; C-terminal His-tag (XhoI site)	AAAAACTCGAGCACTGCGCCGTGAAGAAC
afpA 431 bp promoter region forward primer (BamHI site)	TTTTTGGATCCTGGAAGGCTTTGCCATCGT
afpA 131 bp promoter region forward primer (BamHI site)	AAAAAGGATCCCACCCGTTACGAAGATGAAC
afpA promoter region reverse primer (XbaI site)	AAAAATCTAGATTGGGCTGTCGTATTGCATG

2.2.4 Restriction Endonuclease Digestion and Ligation

Fermentas' FastDigest enzymes (ThermoFisher Scientific, Waltham, MA) were used for all single and double restriction enzyme digests. Each digest was carried out in a 10 to 20 µl reaction mixture. These mixtures usually contained 2 µl of 10X FastDigest buffer, 0.1 to 2 µg of DNA sample and one to two units of each restriction enzyme. Depending on the amount of DNA sample, these mixtures were incubated at 37ºC for one to two hours. After the reaction, the digested DNA was visualized by agarose gel electrophoresis and purified using DNA gel extraction.

Purified, digested DNA fragments were ligated with T4 DNA ligase as specified by Fermentas (ThermoFisher Scientific, Waltham, MA). An optimal molar ratio of three inserts to one vector (3:1) was usually used. A typical 20 µl ligation mixture consisted of 2 µl of 10X ligation buffer (400 mM Tris-HCl, 100 mM $MgCl_2$, 100 mM DTT, 5 mM ATP), 1 µl of T4 DNA ligase (five units) and respective amounts of digested insert and vector. The ligation mixtures were then incubated overnight at 4ºC.

2.3 Transformation

2.3.1 Heat Shock Transformation

DNA transformation into *E. coli* strains was done with heat shock transformation. To prepare competent cells, a 500 ml flask containing 100 ml of LB broth was inoculated with 150 µl of an overnight culture and grown at 37°C to an optical density (OD) at 600 nm of approximately 0.4. This culture was equally distributed into two sterile, pre-chilled 50 ml polypropylene tubes. To cool the culture, the tubes were chilled on ice for 10

minutes. The bacterial cells were recovered by centrifugation at 2700 *x g* at 4°C for 10 minutes. After decanting the supernatant, each pellet was resuspended by gentle swirling in 30 ml of sterile ice-cold solution containing 80 mM $MgCl_2$ and 20 mM $CaCl_2$. The treated cells were collected by centrifugation at 2700 *x g* at 4°C for 10 minutes and gently resuspended in 2 ml of sterile, ice-cold 100 mM $CaCl_2$ solution. For long term storage, 80% (v/v) glycerol was added to the cells with a final concentration of 10% (v/v) glycerol. This cell mixture was then dispensed into 100 µl aliquots in sterile 1.5 ml microcentrifuge tubes. After flash freezing with liquid nitrogen, the competent cells were stored at -80°C.

Prior to transformation, aliquots of competent cells were thawed on ice. After thawing, 10 µl of ligation mixture was added to each aliquot with gentle mixing and left to incubate on ice for 20 minutes. Each mixture was heated at 42°C for 90 seconds, and then cooled on ice for two minutes. Approximately 900 µl of LB medium was added to each tube and left to incubate at 37°C for an hour. After the recovery period, the transformed cells were centrifuged at 2700 *x g* for five minutes at room temperature. After removing 850 µl of supernatant, the cell pellet was gently resuspended in the remaining 150 µl solution and plated on selective media in duplicate. The plates were then left to incubate overnight at 37°C.

2.3.2 Electroporation

DNA transformation into *Pseudomonas* strains was done by electroporation. Electrocompetent cells were prepared using an adapted 10 minute preparation method

(Choi et al. 2006). For each tube of electrocompetent cells, 6 ml of sterile TSB medium was inoculated with one bacterial colony. This culture was grown at 24°C (for *P. putida* GR12-2) or 30°C (for *P.* sp. UW4) until the OD at 600 nm reached 0.8. Cells were then harvested by equally distributing the bacterial culture into four sterile 1.5 ml microcentrifuge tubes. Each tube was then centrifuged at 10,300 *x g* for two minutes. All remaining steps occurred on ice or with pre-chilled equipment. After discarding the supernatant, each collected pellet was washed with one milliliter of ice-cold, sterile 300 mM sucrose solution. Following three to four repeated washings, each chilled pellet was resuspended in 25 μl of ice-cold, sterile 300 mM sucrose and collected into one final tube. At this point, electrocompetent cells are ready for electroporation.

Prior to transformation, it was necessary to ensure that the DNA solutions used contain only minimal amounts of salt to prevent arcing during electroporation pulsing. If the DNA solution contained salt, DNA was first purified using ethanol precipitation and then resuspeneded in sterile water prior to transformation. For each batch of electrocompetent cells, either 50 ng of purified DNA or sterile water was added and gently mixed. This mixture was transferred to a sterile, ice-chilled Fisherbrand® electroporation cuvette with a 2 mm gap width. After wiping any condensation on the cuvette, the cuvette was placed in the cuvette holder, and the cells were pulsed at 2.5 kV, 200 ohms of resistance and 25 μF capacitance using a Bio-Rad Gene Pulser kindly provided by Dr. B. A. Moffatt (Choi et al. 2006). Immediately after pulsing, 900 μl of sterile TSB was added and mixed by inverting the cuvette several times. The electroporated cells were then transferred to a sterile 1.5 ml microcentrifuge tube and incubated in a shaking water bath at 24°C (for *P. putida* GR12-2) or 30°C (for *P.* sp.

UW4) for one hour of recovery. Recovered cells were plated on antibiotic containing selective media for overnight incubation of successful transformants.

2.4 Native AfpA

2.4.1 Isolation and Partial Purification of Native AfpA

Antifreeze protein AfpA from *P. putida* GR12-2 was isolated as described by Xu et al. (1998) with minor variations. One ml of an overnight culture of wild-type *P. putida* GR12-2 was diluted 1000-fold in one liter of sterile TSB and incubated at 5°C. After six days of growth, 600 µl of an acetone solution of one molar phenylmethylsulfonyl fluoride (PMSF) was added until a final concentration of 600 µM to inhibit serine proteases. This culture batch was centrifuged at 15,000 *x g* for 20 minutes at 4°C in a Sorvall RC6 plus centrifuge with a F10S rotor. The supernatant was kept and placed in a 4°C cold room. Powdered ammonium sulfate was slowly added into the magnetically stirring supernatant until 40% solution saturation was attained. At this point, the solution was mixed for an additional hour and then centrifuged at 15,000 *x g* for 30 minutes at 4°C. Due to the presence of a flaky protein precipitate, the supernatant was very carefully decanted. The remaining protein pellet was dissolved in 10 ml of 50 mM Tris-HCl buffer (pH 7.5). This protein solution was sterilized by syringe filtration with a 0.8/0.2 µm polyethersulfone filter (Pall Corporation, Port Washington, NY). Ammonium sulfate in the solution was removed by overnight dialysis at 4°C using a Spectra-Por® membrane (Spectrum Medical Industries Inc., Laguna Hills, CA) with a nominal molecular weight cut off

(MWCO) of 8000 kDa. The solution was dialyzed against three exchanges of one liter each of 50 mM Tris-HCl (pH 7.5) buffer.

2.4.2 Protein Quantification

The concentration of protein samples was determined using the method of Bradford. Dye reagent concentrate, supplied by Bio-Rad Laboratories Inc. (Hercules, CA), was diluted 1/5 with distilled water. In a microtiter well plate, 10 µl of protein sample was mixed with 200 µl of diluted dye reagent and left to incubate for 10 minutes at room temperature. After incubation, the absorbance of each well was measured at 595 nm. Each protein sample was measured in duplicate and compared to a bovine serum albumin (BSA) protein standard. The BSA standards were prepared in concentrations of 0.025, 0.05, 0.075, 0.1, 0.25, 0.5 and 0.75 mg/ml. When necessary, experimental protein samples were diluted until the measured OD fell within the linear range of the BSA standards.

2.4.3 Protein Concentration and Buffer Exchange

Dilute protein samples were concentrated using Amicon® centrifugal filter devices (Millipore Corporation, Billerica, MA). These devices concentrate protein samples by filtration, removing most of the buffer and any solutes below the designated MWCO. Larger solutes are retained in the residual buffer solution. For native AfpA, a filter device with a MWCO of 50 kDa was used. For the recombinant AfpA fragment, a filter device with a MWCO of 30 kDa was used. The protein samples were centrifuged at 4000 *x g* at 4°C in a swinging bucket rotor for 10 to 15 minutes depending on the desired

retentate volume. These devices also used for protein buffer exchange. After concentrating the protein sample, new protein buffer was added to the retentate and re-centrifuged. Following six centrifugation runs, the remaining retentate was recovered and diluted to the desired volume in new protein buffer.

2.4.4 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Both native and recombinant antifreeze protein samples were visualized and resolved using SDS-PAGE with a Bio-Rad Mini-Protean II apparatus (Bio-Rad Laboratories Inc., Hercules, CA). Prior to gel electrophoresis, loading dye was added to protein samples with a final concentration of 1X SDS-PAGE loading dye (0.3125 M Tris (pH 6.8), 50% (v/v) glycerol, 10% (w/v) SDS, 500 mM DTT, 0.1% Bromophenol blue) in a total volume of 25 µl. These samples were then denatured by incubation at 95°C for five minutes. The samples were loaded on a 10% or 12% polyacrylamide gel (with a 5% stacking gel) in 1X SDS running buffer (250 mM Tris (pH 6.8), 1.9 M glycine, 1% (w/v) SDS). They were run at 80V for the length of the stacking gel and at 200V through the resolving gel. When finished, the resolving gel was stained with Coomassie Blue (10% (w/v) Coomassie Blue R250, 10% (v/v) glacial acetic acid, 40% (v/v) methanol) for 20 minutes at room temperature and destained with destaining solution (10% (v/v) glacial acetic acid, 20% (v/v) methanol) overnight.

2.5 Recombinant AfpA

2.5.1 Recombinant Protein Expression and Cell Lysis

For recombinant antifreeze protein expression, plasmid constructs pETAfpAN, pETAfpAC, and pETAFrag were transformed into *E. coli* BL21 (DE3) as the expression host. After pilot experiments, optimal expression was found when *E. coli* BL21 carrying these constructs was grown at 30°C in LB broth (ThermoFisher Scientific, Waltham, MA) containing 50 μg/ml of kanamycin. This culture was grown until the mid-log phase (OD at 600 nm was between 0.4-0.6). Recombinant protein expression was induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and the culture was then incubated at room temperature for 4 hours. The cells were collected by centrifugation at 3000 *x g* for 10 minutes. The pellet was then resuspended in 50 mM Tris-HCl (pH 8) and 50 mM NaCl.

Cell lysis was achieved by sonication using a sonic dismembrator Model 100 (ThermoFisher Scientific, Waltham, MA). To facilitate cell lysis, a final concentration of 1 mg/ml of lysozyme was added to the cell suspension to help disrupt the bacterial cell walls. This mixture was left to incubate on ice for 30 minutes. The lysate was then pulsed 10 times in 15 second intervals at 200W. To avoid overheating and protein denaturation, the sample was constantly kept on ice. The soluble and insoluble fractions of the lysate were separated by centrifugation at 10,000 *x g* for 30 minutes. These fractions were checked for expression of the recombinant protein by running them on SDS-PAGE. If the soluble fraction contained the recombinant protein, it was saved and stored at 4°C for His-tag purification.

2.5.3 His-tag Purification by Nickel Resin

The recombinant full AfpA and AfpA fragment were purified by its respective 6X His-tag using HisPur® Ni-NTA resin, a nickel-charged resin (ThermoFisher Scientific, Waltham, MA). This resin has a 60 mg protein binding maximum capacity per ml of resin. Ni-NTA resin was prepared as per the manufacturer's instructions. Before purification, the soluble fraction of the cell lysate was filtered with a 0.2 μm syringe filter to remove any insoluble components that could interfere with the Ni-NTA resin. This filtered protein extract was then added to the prepared resin and incubated overnight at 4°C with continuous gentle mixing. Afterwards, at 4°C, His-tag purification of this extract was performed as per the manufacturer's directions in a 30 cm tall gravity-flow column. The equilibration buffer used was 20 mM Tris-HCl (pH 7.5) and 5 mM imidazole; the wash buffer used was 20 mM Tris-HCl and 10 mM imidazole; the elution buffer used was 20 mM Tris-HCl and 20/50/100 mM imidazole. Protein quantity and quality of flowthrough, washes and elutions were resolved using SDS-PAGE.

2.6 Size Exclusion Chromatography

Size exclusion chromatography was used to estimate the size of both native AfpA and recombinant AfpA fragment. A HiPrep 26/60 Sephacryl S-300 High Resolution prepacked gel filtration system (GE Healthcare, Cleveland, OH) was connected to a Bio-Rad Biologic Duo Flow FPLC system (Bio-Rad Laboratories Inc., Hercules, CA). Preparation, running, cleaning and storage of the SEC column were conducted as per the Manufacturer's protocol. In this system, 1 ml of concentrated, filtered protein sample was injected with 50 mM Tris-HCl (pH 8) and 150 mM NaCl buffer as the mobile phase. This

was run at a flow rate of 0.8 ml/min. A standard curve was generated with 640 kDa bovine thyroglobulin, 440 kDa horse ferritin, 158 kDa chicken ovalbumin and 17 kDa horse myoglobulin (Bio-Rad Laboratories, Hercules, CA; Pharmacia, Uppsala, Sweden).

2.6 Protein Assays

2.6.1 Thermal Hysteresis (Antifreeze) Assay

Thermal hysteresis activity was measured using an Otago nanoliter osmometer (Otago Osmometer Ltd., Dunedin, New Zealand) and an Olympus phase contrast light microscope (Olympus Corporation Canada, Richmond Hill, ON). The nanoliter osmometer consisted of a temperature controller and a thermoelectric cooling stage (Ramløv 2010). The temperature controller had two options: freeze or return to a predetermined temperature. The temperature of the second option could be changed and fine tuned. The cooling stage is an aluminum stage with a disc shaped hole in the middle. During the assay, high viscosity oil is placed in the middle of the hole. A clean sample holding disc, with thermal grease on the underside, is fitted in the hole so that oil fills the wells of the disc. Using a heat modified 1.0 mm glass capillary tube and tubing, around 10 nl of antifreeze protein sample was carefully injected into each oil well to form a droplet encased in oil.

Thermal hysteresis was determined by monitoring changes in ice crystal morphology at different temperatures. Switching to the freeze option, these droplets were frozen beyond -20°C. After complete freezing, the temperature was slowly increased until one individual ice crystal remained. While carefully adjusting the temperature, ice crystal morphology and growth was observed. Thermal hysteresis was measured as the

temperature difference between when the ice crystal starts melting and when it starts growing.

2.6.2 Ice Recrystallization Inhibition Assay

Ice recrystallization inhibition activity was measured with a thermoelectric cold plate and an Olympus phase contrast microscope (Olympus Corporation Canada, Richmond Hill, ON). The assay was carried out as described by Livernois and colleagues (2009). Protein samples were supplemented with 30% (w/v) sucrose. Qualitative determination of ice recrystallization activity was done by comparing pictures of ice crystals sizes at -6°C taken at 0 and 60 minutes (Livernois et al. 2009).

2.6.3 Luciferase Assay

The promoter strength of the *afpA* promoter was quantified indirectly with the use of promoterless-probe vector pQF70 (Farinha and Kropinski 1990). This vector contains *luxAB* genes from the bacterium *Vibrio harveyi*, encoding luciferase protein. Lacking a promoter directly upstream of the *luxAB* genes, luciferase protein encoded in this vector is not expressed without a cloned upstream promoter. Using these genes as reporter genes, promoter-insertion positive clones can be indirectly tested for promoter activity by quantifying luciferase luminescence. The blue green (490 nm) luminescence was quantified using luminometer settings on a Victor^3 V multilabel reader (Perkin Elmer).

Triplicates of overnight cultures were diluted 100-fold in new sterile broth. All experimental growth conditions tested are listed in Table 3. These diluted cultures were cultivated until the exponential growth phase when the OD at 600nm reached around 0.4

Table 3. Experimental conditions for luciferase assay

Temperature	Medium	Medium supplementation
4°C	TSB	None
24°C	TSB	None
	M9 minimal medium lacking a carbon source[1]	0.4% Glucose
		0.4% Glycerol
		0.4% Galactose
		0.4% Sucrose
		0.4% Fructose
		0.4% Sorbitol
		0.4% Mannitol
	DF minimal medium lacking nitrogen source[2]	0.3% Calcium nitrate
		0.3% Potassium nitrate
		0.3% Sodium nitrate
		0.3% Ammonium nitrate
		0.3% Ammonium sulfate
		0.3% Ammonium chloride
		0.3% Peptone
		None
	M9 minimal medium[3]	None
		2 mM Cytidine
		10 mM N-3-Oxo-dodecanoyl-L-homoserine lactone
		10 mM N-Butyryl- DL-homoserine lactone
30°C	TSB	None

1 – Medium Composition: 42 mM Na_2HPO_4, 30 mM KH_2PO_4, 0.1% NH_4Cl, 1.0 mM $MgSO_4$, 1.0 mM $CaCl_2$ and 0.001% Thiamine-HCl

2 – Medium Composition: 42 mM Na_2HPO_4, 30 mM KH_2PO_4, 0.8 mM $MgSO_4$, 0.2% Glucose, 10 mM gluconic acid, 10 mM citric acid, 0.1 mM $FeSO_4$, and Trace elements (H_3BO_3, $MnSO_4$, $ZnSO_4$, $CuSO_4$ ad MnO_3)

3 – Medium Composition: 42 mM Na_2HPO_4, 30 mM KH_2PO_4, 0.2% Glucose, 0.1% NH_4Cl, 1.0 mM $MgSO_4$, 1.0 mM $CaCl_2$ and 0.001% Thiamine-HCl

since luciferase activity requires cellular ATP during the exponential phase of cell growth. Depending on the experimental temperature and broth conditions, the growth time varied between 5 to 48 hours. For each triplicate sample, one milliliter of culture was transferred to a 1.5 ml microcentrifuge tube and centrifuged at 10,300 *x g* for three min. During this time, the Victor^3 V reader was turned on and set to 30°C. After removing the supernatant, each culture pellet was resuspended in 100 µl of 0.85% NaCl and transferred to opaque white wells in a 96-well microtiter plate (Corning Co-star). In each well, 10 µl of n-decanal substrate was added. Following circular mixing for 99 seconds at normal speed, light output from each well was measured with a 460 nm excitation/535 nm emission filter.

Using the negative control (empty pQF70 vector) as the baseline for luminescence, each experimental reading was normalized by dividing each experimental condition reading with the negative control reading. Each luminescence reading was then reported as the relative luminescence level dependent on negative control levels. The final relative luminescence level was presented as the mean for each experimental condition with calculated standard error.

2.6.3.1 Statistical Tests

For each condition, a statistical test was used to determine significance depending on the data collected. For the assays with different temperature growth using either a full or partial promoter constructs, both temperature effect and promoter length effect were examined using one-way ANOVA. For the cold acclimation condition (testing for CspA regulation), the change in activity over time (slope) was considered the main effect.

Although the data was discrete, to test this cold acclimation effect the data was treated as continuous with a calculated slope. Statistical significance of these slopes were compared using standard t-tests. For carbon and nitrogen effects, a simple one-way ANOVA with a Tukey post-hoc test was used to determine if any of the growth sources were statistically different. For the induction assays with different metabolites, a mixed factorial ANOVA was used to examine the effect of metabolites between subjects (compared to the negative control) and within subjects (i.e., significant change after adding the metabolite). In the osmoregulation assay with different saline concentrations, statistical significance was determined using a simple t-test.

2.7 Bioinformatics Prediction Programs

A number of online bioinformatics programs have been used for different predictions. Deviations from any default parameters are detailed in the Results section under the respective program used. At the DNA sequence level, Softberry's BPROM was used to predict sigma 70 based promoter similar to *E. coli*'s canonical -35 and -10 promoter elements (Solovyev and Salamov 2011). Virtual footprint (Version 3.0) was used to predict *cis* promoter elements that may bind to known transcription factors in *E. coli* and *Pseudomonas aeruginosa* (Münch et al. 2005). RibEx: Riboswitch explorer was used to predict potential transcriptional or translational attenuators (Abreu-Goodger et al. 2004). Similarly, two programs (RNAstructure and MFOLD version 2.3) were used to analyze probable translational attenuator by predicting the secondary structure of 5' UTR of the *afpA* mRNA transcript (Reuter and Mathews 2010; Zuker 2003). Phyre2.0 was used for protein structure prediction (Kelley and Sternberg 2009). AFPredictor was

subsequently used to calculate a potential ice binding surface using spatial regularity of ordered surface carbons in the predicted protein structure surface (Doxey et al. 2006).

3.0 Results

3.1 Bioinformatic Characterization of the afpA Gene Sequence

Building upon the previous analysis by Muryoi and colleagues (2004), additional characteristics were predicted in the *afpA* gene sequence (Figure 6). Within the nucleotides upstream of the *afpA* coding sequence, the program BPROM predicted two RpoD based promoter elements, the -35 and -10 box, upstream of the transcriptional start site (Solovyev and Salamov 2011). In the same region, several putative transcription factors were found on the plus strand by online program VirtualFoot (Table 4). This program used position based matrices to score the query sequence (*afpA* promoter region) against conserved core nucleotides in canonical transcription factor sequences from *Pseudomonas aeruginosa* and *Escherichia coli*. Results from the gram positive bacterium *Bacillus subtilis* were excluded. After lowering the sensitivity parameters to increase selectivity, several transcription factors were predicted. Factors with high scores that are testable with exogenous molecules were reported. Despite a lower score of 5, *Pseudomonas aeruginosa* NarL was also tested due to its role in nitrate regulation; nitrogen growth sources have been documented to affect expression of ice nucleation activity (Chen et al. 2003). For each predicted transcription factor, their effect on *afpA* promoter strength was tested using the exogenous substrates listed in Table 4. NarL was tested using different nitrogen sources in the growth medium.

Downstream of the *afpA* coding sequence, the programs Riboswitch and ARNold predicted a transcriptional terminator, antiterminator and anti-antiterminator approximately 30 to 50 nucleotides after the stop codon (Figure 6 green box). Combined

Figure 6. *afpA* sequence and predicted gene elements. The *afpA* coding sequence is highlighted in pink. Predicted transcriptional terminator and antiterminator region is highlighted in green. The coding sequence of a putative glycosyl transferase gene is represented in orange. Previously deduced transcriptional start site and Shine Dalgarno sequence are circled and underlined respectively (Muryoi et al. 2004). Sequences encased in black boxes represent predicted -35 and -10 promoter elements of the putative downstream gene.

Table 4. Predicted transcription factors within the promoter region of the *afpA* gene. For each transcription factor, prediction details are listed. Experimental conditions used in this study to test these predictions are indicated accordingly.

Transcriptional factor name	**Regulation**	**Score**	**Sequence**	**Experimental condition**
CspA	Cold shock	10.00	CCAAT	Cold acclimation
CytR	Cytidine catabolism	7.65	GGCAAATCTCTA	Cytidine
CytR	Cytidine catabolism	8.09	AGAAAATTAGTA	Cytidine
OmpR	Osmoregulation	8.28	GAAAACT	Salt
OmpR	Osmoregulation	8.17	GAAAATT	Salt
RhlR	Quorum sensing	10.22	CTCTGGTATCTGCCGG	Homoserine lactones
RhlR	Quorum sensing	9.27	CTTTGCCATCGTCTGG	Homoserine lactones

with the predicted transcriptional start site, the *afpA* RNA transcript would be 1.5 kb in size.

Following closely behind the full *afpA* gene is a putative glycosyl transferase gene (Figure 6 orange box). The downstream nucleotide sequence was translated into an amino acid sequence using six different reading frames. One potential amino acid sequence was then subsequently used as query sequence in a default blastp search against the National Center for Biotechnological Information (NCBI) non-redundant database (Altschul et al. 1990). Top results showed greater than 90% coverage and 50% sequence identity were glycosyl transferases in *Pseudomonas* species. The majority of these glycosyl transferases belonged to family 2 or were unidentified. RpoD promoter elements were also predicted upstream of this putative glycosyl transferase gene. However, the promoter is within the transcriptional terminator region of the *afpA* gene. There are three potential hairpin loop structures in this region that can be formed by different nucleotide binding (Figure 7). The colours represent the shared nucleotides between each structure (i.e., the green nucleotides of structure one and two are the same in both conformations). The first structure is the transcriptional terminator of the *afpA* gene. The second structure is an antiterminator, which is an alternate conformation of the transcriptional terminator. When this antiterminator is formed, the transcriptional terminator is abolished. This structure change likely allows transcription to continue. The third structure is an anti-antiterminator. This structure likely prevents the formation of the antiterminator by binding the red nucleotides in this alternate conformation. This permits the green nucleotides of the antiterminator to fold into the transcriptional terminator conformation. Therefore, in one condition, structures one and three are formed while in another

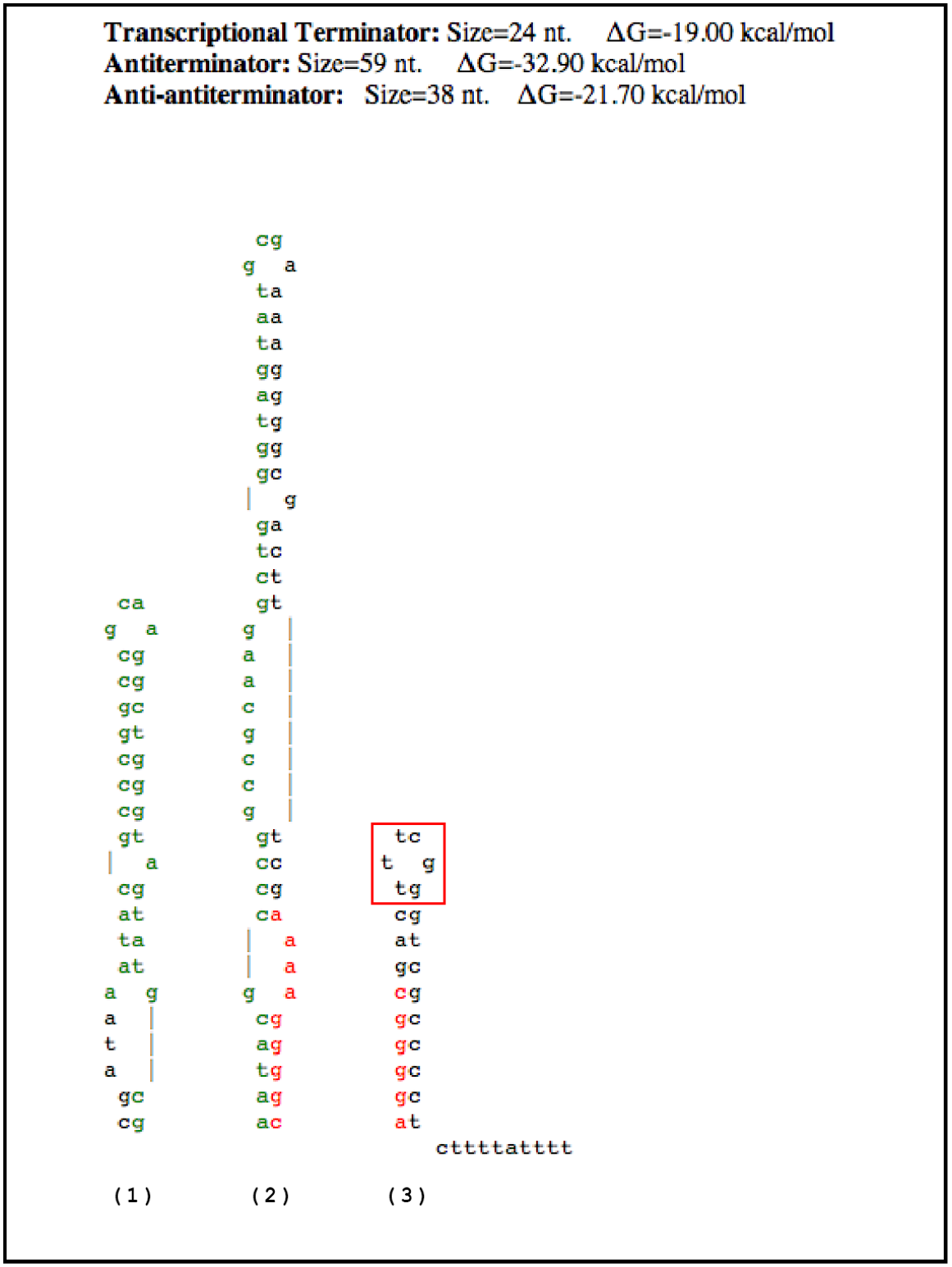


Figure 7. Three putative secondary structures of the transcriptional terminator region of the *afpA* gene. Green and red sequences represent shared nucleotides between alternate secondary structures. (1) – transcriptional terminator. (2) – antiterminator. (3) – Anti-antiterminator. Red box – the -35 box of the glycosyltransferase promoter.

condition, only structure two is formed. Based on these predicted secondary structures of this terminator region, the -35 box is on the loop portion of the anti-antiterminator hairpin loop during transcriptional termination of *afpA* (Figure 7). However, when the antiterminator is formed, this -35 element is no longer involved in these conformations. This change in secondary structure indicates that the position and availability of this promoter can potentially be altered under different conditions. The Riboswitch program did not identify the terminator and antiterminator complex as a known riboswitch. A riboswitch is a RNA structure that can "sense" a metabolite for small-molecule-mediated transcriptional attenuation. A lack of an identified riboswitch indicates that the terminator and antiterminator complex structure is not likely to be influenced by metabolites.

3.2 Construction of pQF70K and Experimental Derivatives

The vector pQF70K was constructed by cloning a kanamycin resistance gene from expression vector pET30b(+) into the NdeI site of the pQF70 vector (Figure 8). This restriction enzyme site was used to avoid interference with sequences cloned into the multiple cloning site. The kanamycin resistance gene, with promoter and transcriptional terminator, was amplified using KOD hot start polymerase from pET30b(+) with primers containing an NdeI restriction site (Figure 9a). These PCR products and the pQF70 vector were digested with NdeI, ligated together with T4 DNA ligase and then transformed into *E. coli* DHA5α. Transformants that survived growth with 50 μg/ml kanamycin were confirmed to have the new construct by digesting the isolated plasmid with NdeI (Figure 10 green boxes). This digestion confirmed that the kanamycin resistance gene was present.

To create the derivative vectors, pQKAfp431 and pQKAfp131, the full or partial promoter region of the *afpA* gene was cloned into the multiple cloning site in pQF70K at the BamHI and XbaI site. These inserts were amplified from *P. putida* GR12-2 genomic DNA, digested with the two restriction enzymes, ligated together with T4 DNA ligase and then transformed into *E. coli* DH5α (Figure 9b). The new constructs were isolated from transformants and confirmed by a double digest with BamHI and XbaI (Figure 10 and Figure 11). At the same time, the constructs were digested by NdeI to confirm that the promoter inserts were cloned into pQF70K and not pQF70. A construct containing the promoter of the *iaaM* gene from *P.* sp. UW4 was made as a positive control because of its comparatively strong promoter activity (unpublished data). This construct was also confirmed by restriction enzyme digestion. After confirmation, these new constructs were then transformed into the experimental strains to be tested.

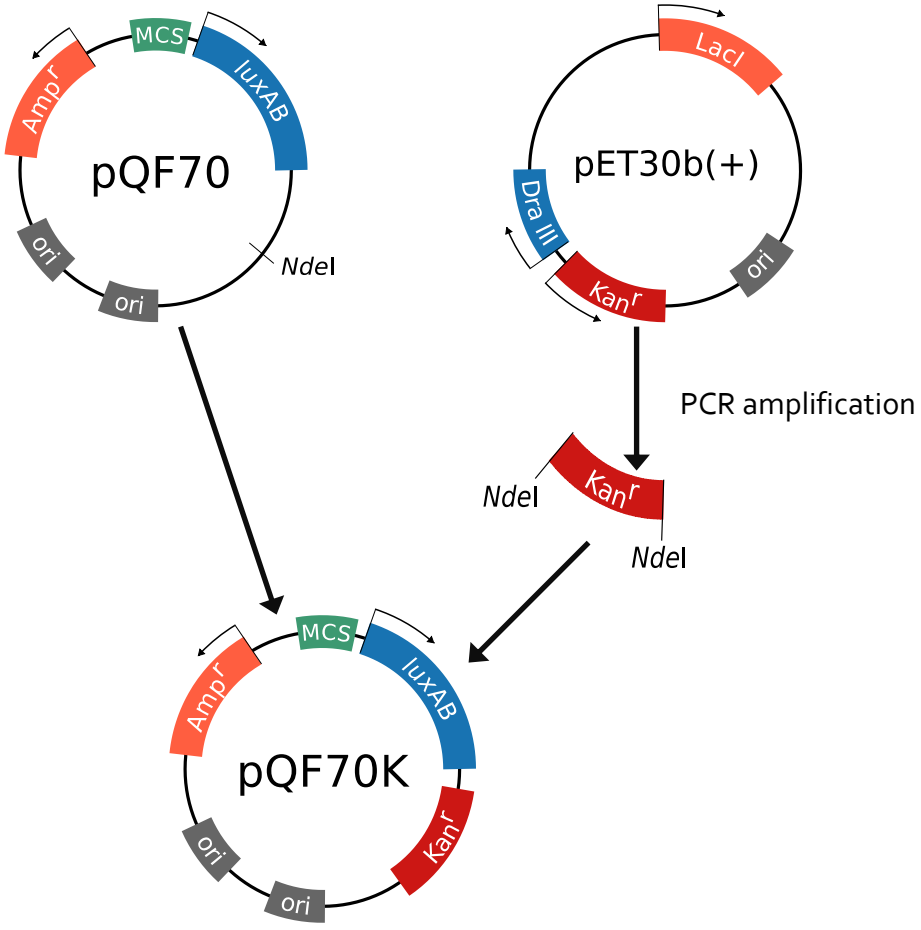


Figure 8. Construction of the pQF70K vector. The vector pQF70 was modified to include a kanamycin resistance gene, originating from vector pET30b(+), at the NdeI site. Amp^r - ampicillin resistance gene, Kan^r – kanamycin resistance gene, *luxAB* – bacterial luciferase gene, MCS – multiple cloning site (SstI, KpnI, BamHI, XbaI, SalI, AatII, MluI NcoI BglII XhoI StuI, SphI, HindIII), ori – origin of replication.

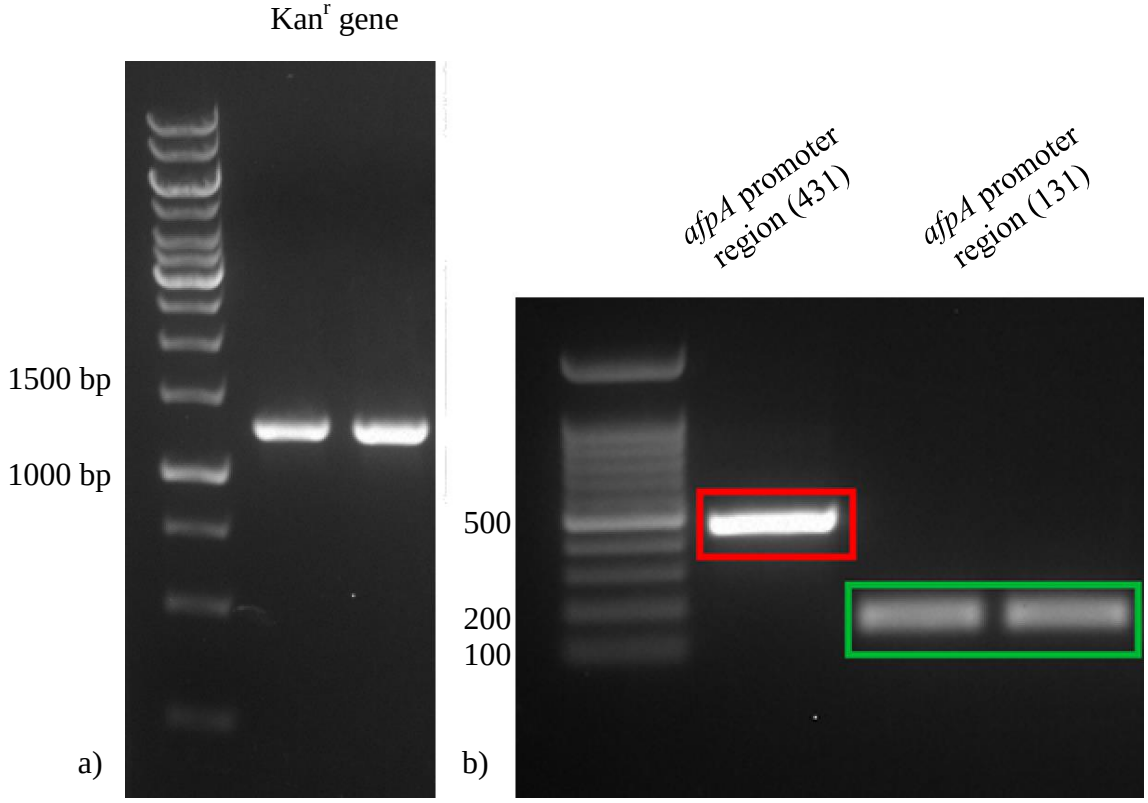


Figure 9. PCR amplification of the Kan^r gene from pET30b(+) and *afpA* promoter regions (431 bp and 131 bp) from *P. putida* GR12-2 genomic DNA. (a) The Kan^r gene product is 1241 bp in size. The DNA ladder used is Fermentas generuler 1 kb ladder. (b) the full promoter region is 431 bp (red box) and the partial promoter region is 131 bp (green box). The DNA ladder used is Fermentas generuler 100 bp ladder. Samples were prepared in duplicates.

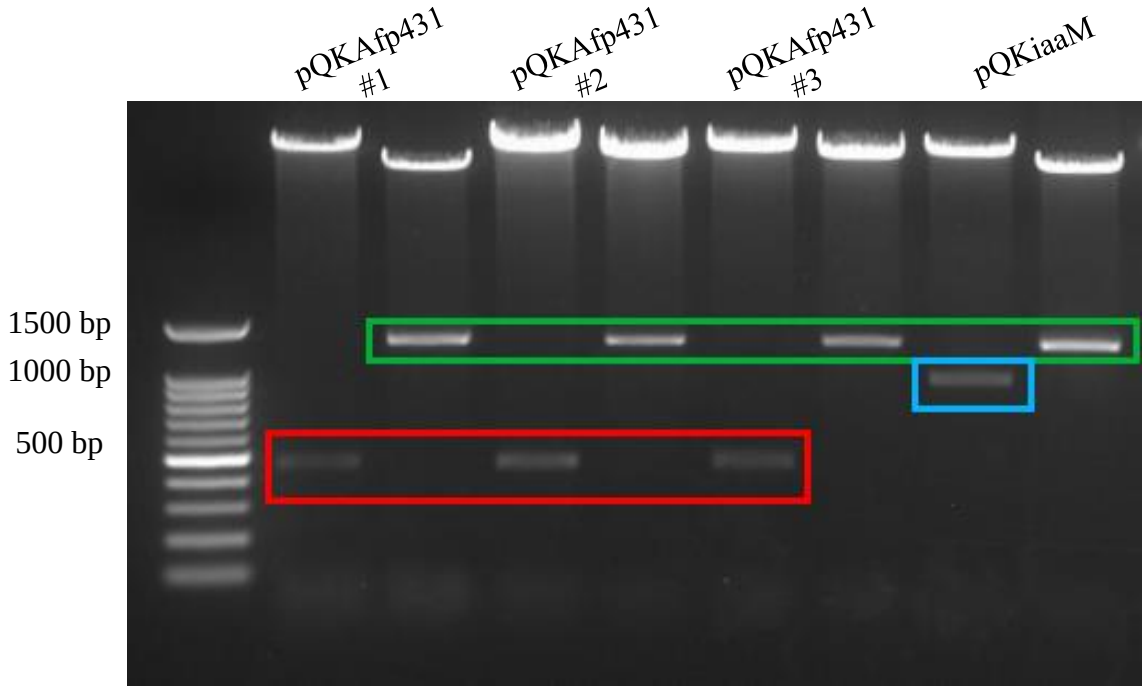


Figure 10. Confirmation of inserts in respective plasmid DNAs by restriction enzyme digestion. Each plasmid construct was digested to confirm both the vector's Kan^r gene (green box) and a promoter region insert. DNA bands encased in a red box indicate digested *afpA* full promoter insert with BamHI and XbaI. The DNA band in blue indicates digested *iaaM* promoter insert (850 bp) with NcoI and SalI. The DNA ladder used is Fermentas generuler 1 kb ladder.

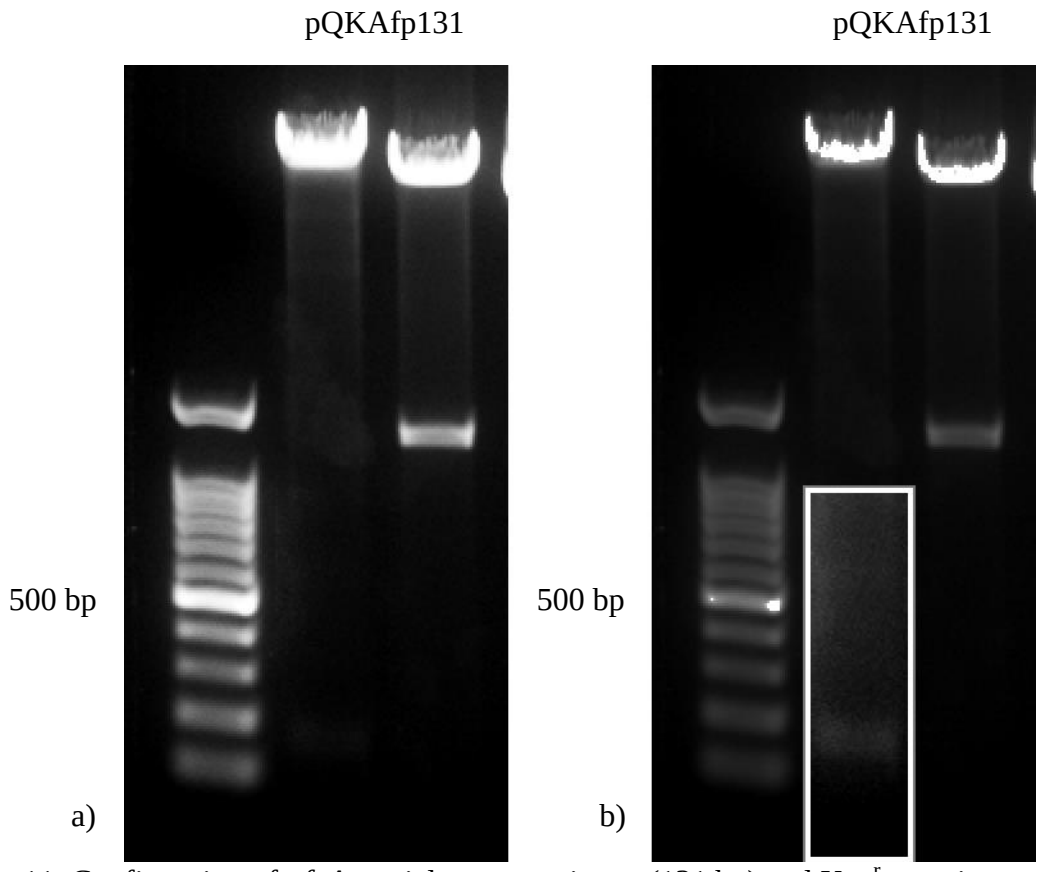
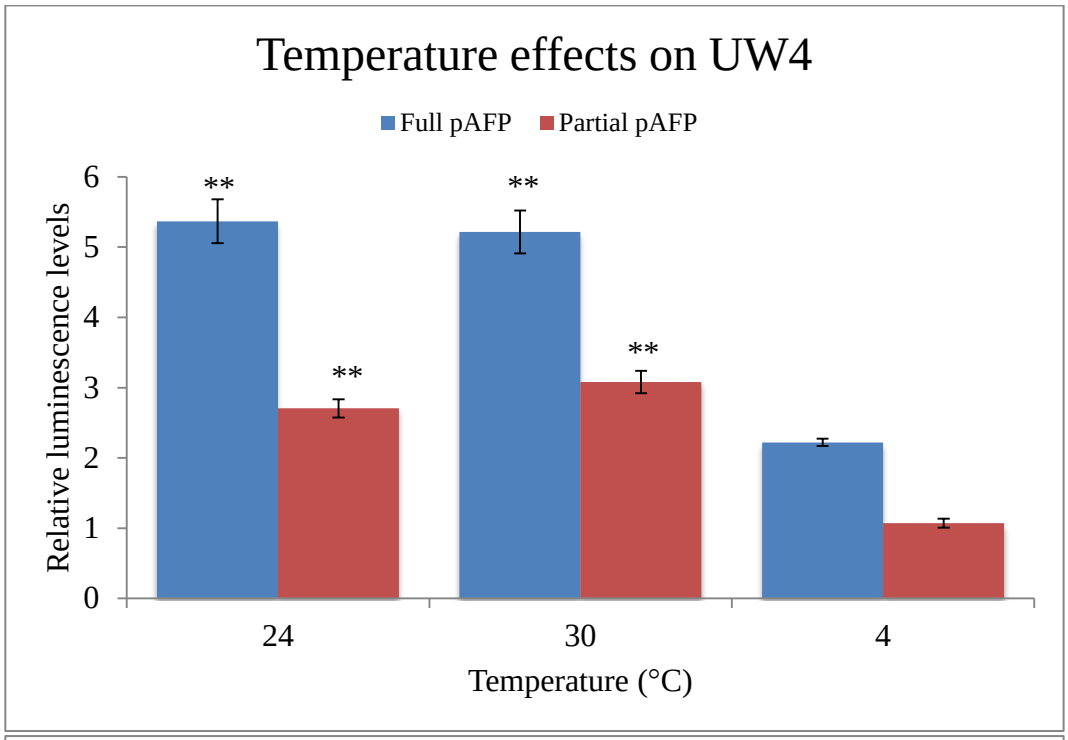


Figure 11. Confirmation of *afpA* partial promoter insert (131 bp) and Kan^r gene in pQF70K vector using restriction enzyme digestion. Kanamycin resistance gene was digested with NdeI. Promoter insert was isolated by double digestion with BamHI and XbaI. (a) Original agarose gel picture (b) Selective brightness and contrast (in box) was adjusted to enhance the faint DNA band of the partial promoter insert.

3.3 Promoter Activity Strength Quantification by Luciferase Assay

The promoter strength of each *afpA* promoter region construct (pQKAfp431 and pQKAfp131) was indirectly tested with a luciferase assay. With this assay, the promoter strength measured is reported as the relative luminescence level because promoter activity is proportional to the activity of expressed luciferase proteins. The relative luminescence level is the magnitude that the experimental luminescence level is greater than the negative control luminescence level. A relative luminescence level of one means that the level of activity found in the test bacterium (at the respective test condition) is equivalent to the negative control; A level two means a two-fold greater activity. Both constructs were transformed in *P.* sp. UW4 and *P. putida* GR12-2 which were subsequently tested under various growth conditions. On all graphs, pQKAfp431 and pQKAfp131 are designated as either full or partial promoter region (pAFP) respectively. The first condition tested was temperature (Figure 12). Although luciferase activity is affected by temperature, these results should remain accurate since temperature affects both the experimental and negative control activity levels. The reported relative levels should be similar when experimental signals are considered in the context of the respective negative control. Both strains were tested at 24°C and 4°C. UW4 was also tested at 30°C because it is the bacterium's optimal growth temperature. For both UW4 and GR12-2, across all temperature conditions, a full pAFP has nearly double the activity of a partial pAFP construct. At temperatures 24°C and 30°C, the relative luminescence levels for full pAFP constructs in UW4 are comparable. GR12-2 has slightly lower activity level at 24°C than UW4, but is within the sample error range. For partial pAFP constructs, at the same temperatures, it appears that UW4 construct are approximately

i)



ii)

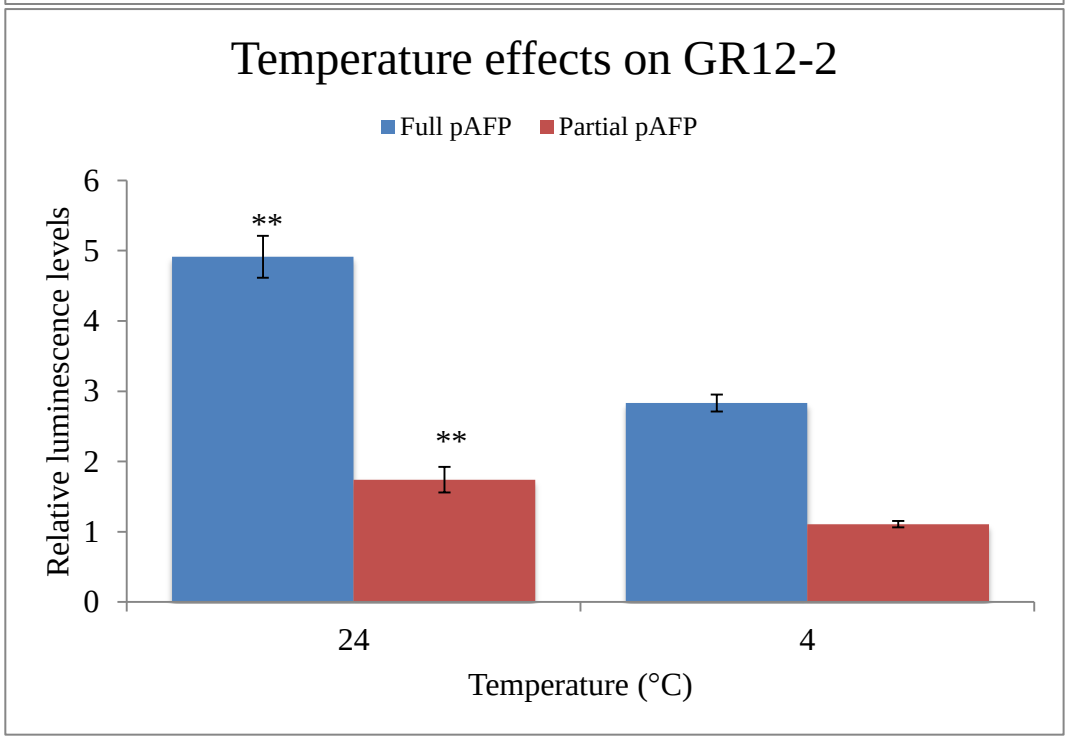
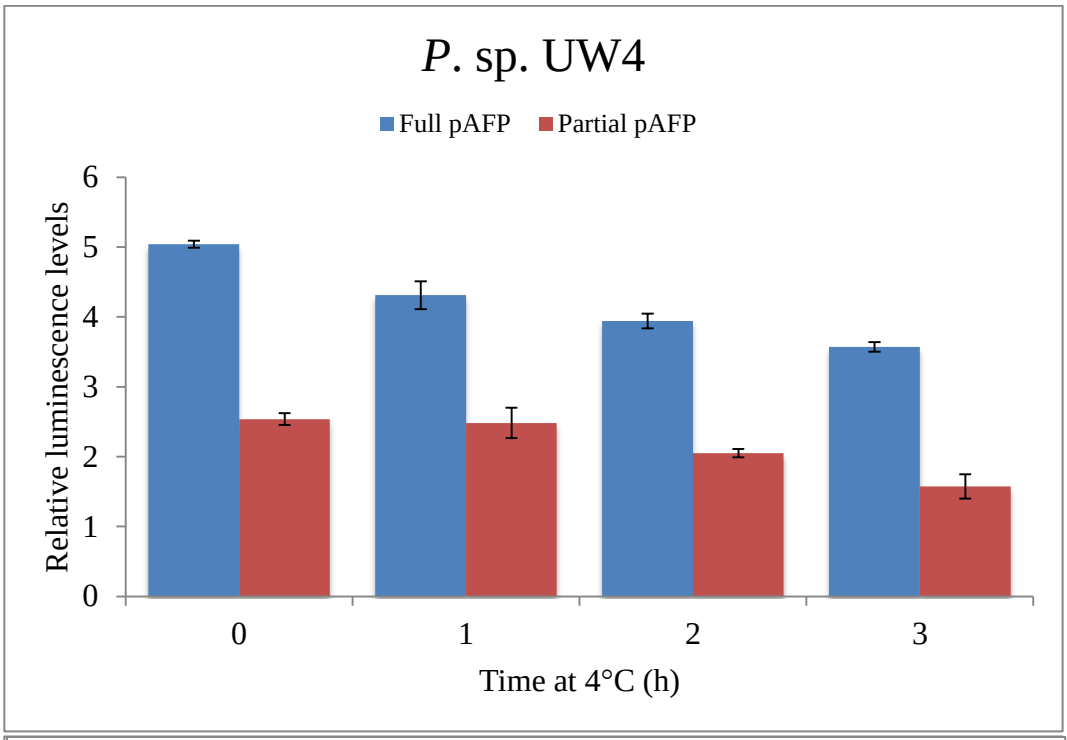


Figure 12. Relative luminescence levels following growth at different temperatures with either the full promoter region (pQKAfp431) or the partial promoter region (pQKAfp131). (i) *P.* sp. UW4 (ii) *P. putida* GR12-2. ** p-value < 0.01 (for temperature effects only)

one level higher than GR12-2 construct. At 4°C, both strains have consistently lower activity levels for full pAFP constructs compared to other temperatures. Partial pAFP constructs in both strains showed relative luminescence levels of approximately one. This low level indicates that at 4°C, luciferase activity from partial pAFP constructs are equal to the activity from the negative controls. This suggests that expression has been terminated or there is not sufficient activity to differentiate it from background noise. For both constructs, the decrease in activity seen at 4°C (main temperature effect) is statistically significant.

The second condition tested was cold acclimation. Cold acclimation was tested by growing the bacterial strains at 24°C until the exponential phase, and then shifting the growth temperature to 4°C. This change triggers a cold acclimation phase where the bacteria adjust to a temperature decrease. This phase is characterized by its expression of cold shock proteins and other cold related proteins as well as a transition into the stationary growth phase. The effects of cold acclimation on luciferase activity are shown on Figure 13. For UW4, both promoter region constructs showed a gradual decrease in activity the longer UW4 stayed at 4°C. With both partial and full pAFP, GR12-2 showed a consistent very slight increase in activity level over time. To test the behaviour of these trends between the partial and full promoter constructs, the results were treated as continuous data with a calculated slope. For each bacteria strain, the slopes of both partial and full pAFP were compared using a standard t-test. In each respective host bacteria, there was no significant difference between the slopes found with partial and full pAFP constructs. Since both bacteria had different slopes in activity, they were compared to determine if these changes are statistically different between bacteria. When the

i)



ii)

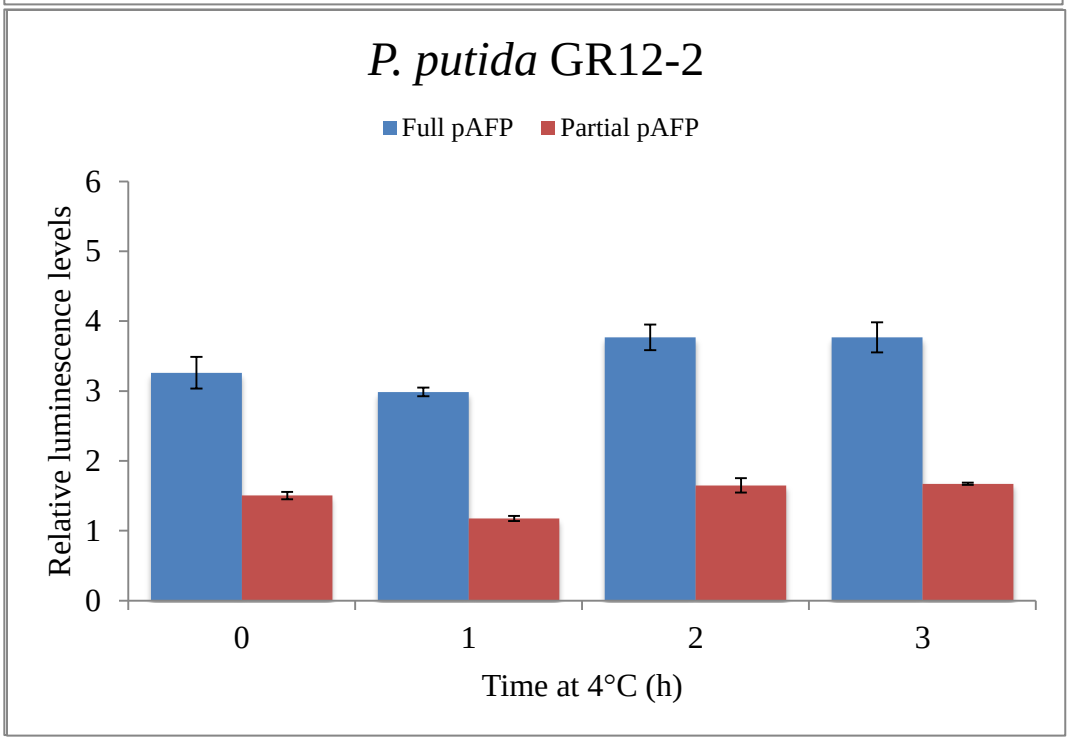
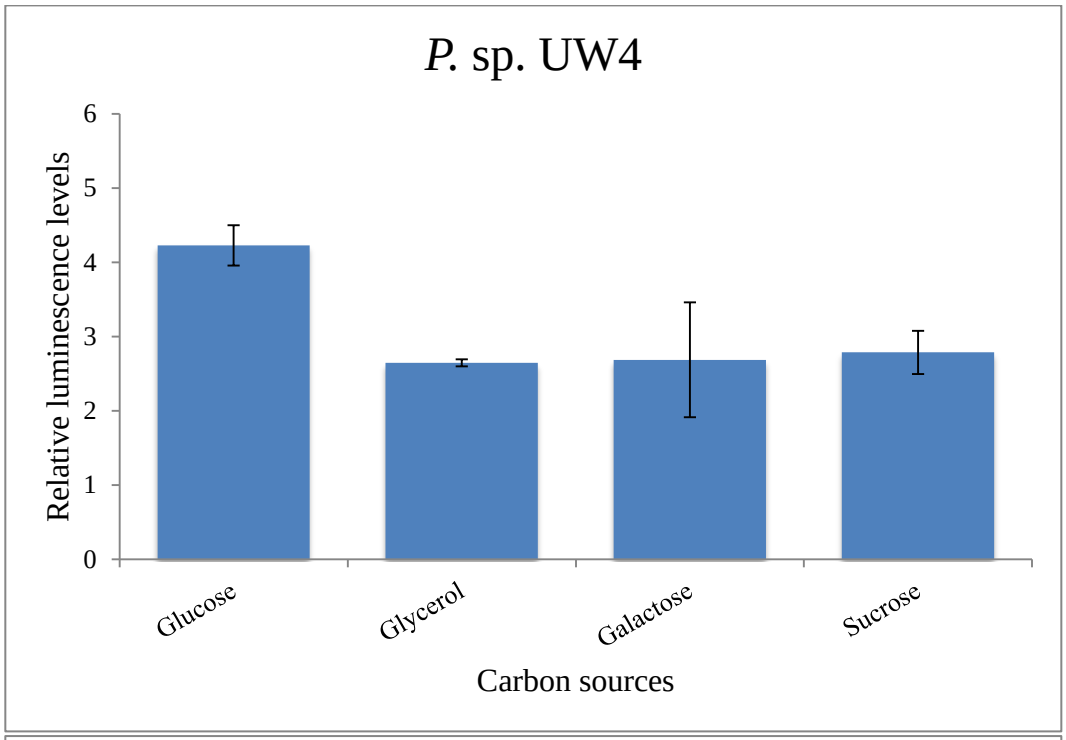


Figure 13. Relative luminescence levels during cold acclimation phase at 4°C with either the full promoter region (pQKAfp431) or the partial promoter region (pQKAfp131). (i) *P.* sp. UW4 (ii) *P. putida* GR12-2. The change in activity over time using full pAFP were statistically different between UW4 and GR12-2 with a p-value < 0.01

decreasing activity in UW4 (slope = -0.48) was compared to the more stable activity in GR12-2 (slope = 0.23) using a t-test, these changes were found to be statistically different (p-value < 0.01). This stability suggests that the promoter region in GR12-2 is able to maintain expression during a decrease in temperature, unlike in UW4.

Potential effects of carbon and nitrogen sources on promoter strength were also tested in both strains (Figure 14 and Figure 15). The best carbon source for UW4 was glucose which showed the highest luciferase activity, however this activity was not statistically significant when compared to the other carbon source conditions. For GR12-2, this bacterium could grow with a larger variety of carbon sources. Of these sources, both glucose and glycerol showed the highest activity level and were statistically different with a p-value less than 0.01 when tested using one-way ANOVA with a Tukey post-hoc test. Nitrogen sources have a lesser effect on promoter activity for both bacteria (Figure 15). In strain UW4, growth with ammonium chloride showed the greatest activity while growth in a nitrogen free medium had the lowest activity. Unlike the situation with carbon sources, the remaining nitrogen sources showed relative luminescence level between the highest and lowest activities. This variation suggests preferential use of nitrogen sources can potentially affect the *afpA* promoter activity in these strains. GR12-2 showed a trend similar to UW4. This bacterium had the highest promoter activity when grown with potassium nitrate, and the lowest activity when grown with ammonium nitrate. Similar to strain UW4, the remaining nitrogen sources were observed with activities between the highest and lowest levels. Despite the variation in GR12-2, these activity levels were found to not be statistically different. In UW4, a one-way ANOVA with a Tukey post hoc test revealed three groups that are statistically different (Figure

i)



ii)

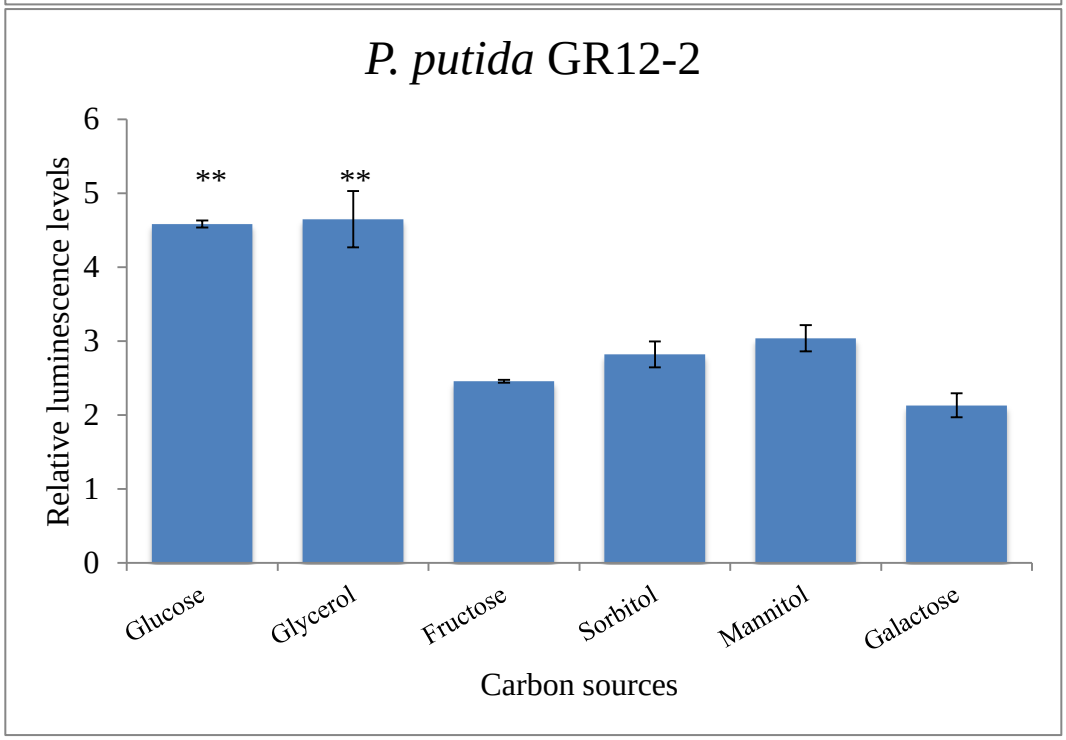
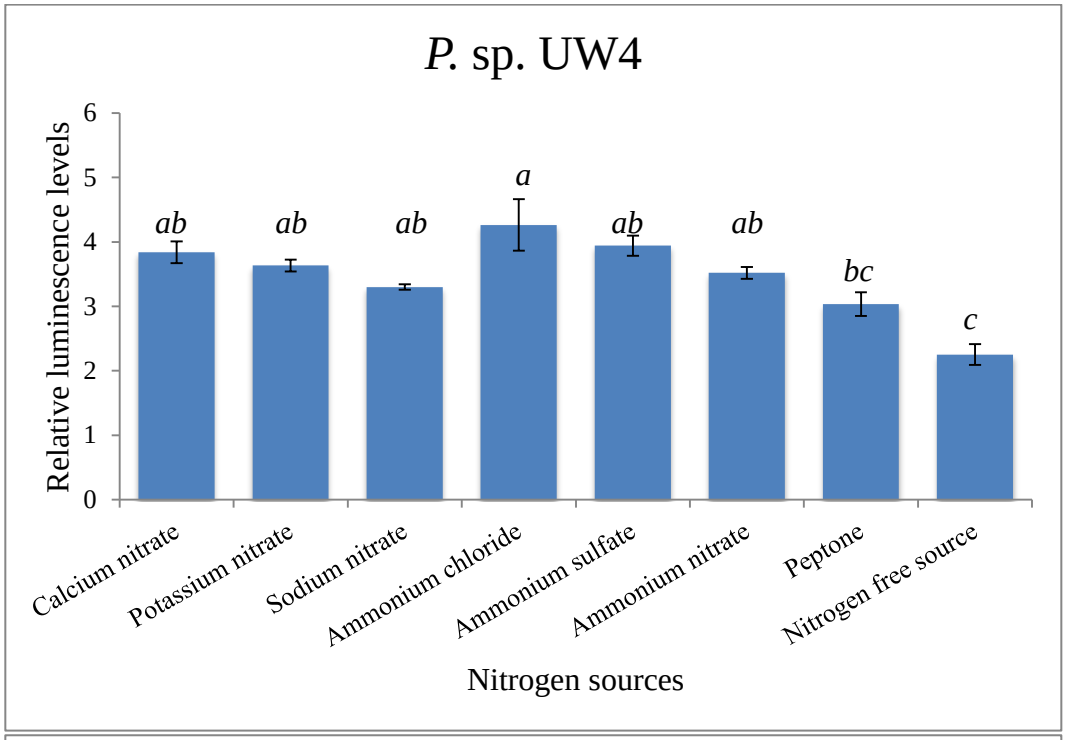


Figure 14. Relative luminescence levels following growth at 24°C with 0.4% of different carbon sources and ammonium chloride as the nitrogen source. These levels were measured using a full promoter region (pQKAfp431). (i) *P.* sp. UW4 (ii) *P. putida* GR12-2. ** p-value < 0.01

i)



ii)

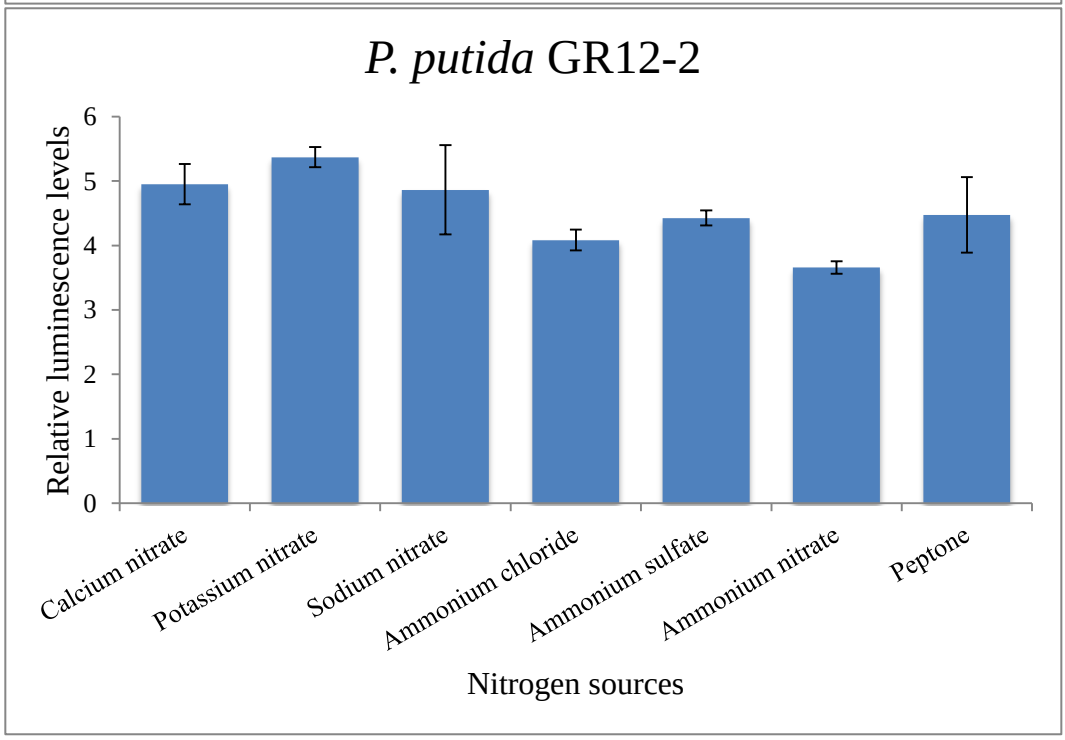
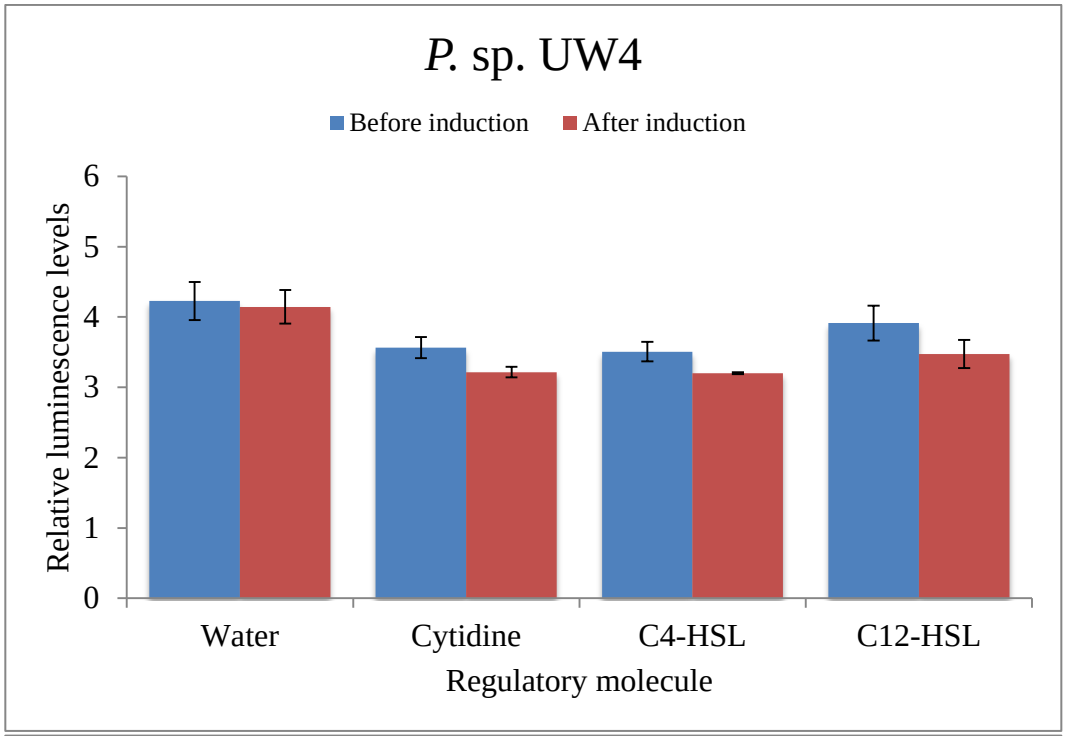


Figure 15. Relative luminescence levels following growth at 24°C with 0.3% of different nitrogen sources and glucose as the carbon source. These levels were measured using a full promoter region (pQKAfp431). (i) *P.* sp. UW4 (ii) *P. putida* GR12-2. *a,b,c* – statistically different groups with a p-value < 0.05

15a). Overall, both carbon and nitrogen source have little to no impact on reported luminescence levels with a few exceptions.

Since a number of testable transcription factors were predicted in the promoter region, the effect of these factors were tested by adding potential exogenous molecules known to be involved with these factors. Possible CspA involvement was already tested during cold acclimation. Cytidine was used to test the possible involvement of CytR, which is involved in cytidine catabolism. Two homoserine lactones, N-3-oxo-dodecanoyl-L-homoserine lactone and N-butyryl-DL-homoserine lactone, were used to test the predicted RhlR factor involvement in quorum sensing. Since this RhlR factor was predicted based on the *P. aeruginosa* sequence, N-butyryl-DL-homoserine lactone was specifically chosen for its involvement with RhlR in this pseudomonad. To directly test these exogenous factors on promoter strength, the strains containing the full promoter region constructs were grown to exponential phase prior to adding the regulatory molecules. Luciferase activity was measured before and 1.5 hours after adding the molecules (Figure 16). For GR12-2, adding these molecules had little to no effect on the relative luminescence levels. The decrease in activity after induction is significant (within subjects) but statistically similar to the negative control (between subjects) which suggests that this decrease is a result of the assay itself. In UW4, these decreases were deemed statistically different than the negative control. However, since the same magnitude of difference was also seen in GR12-2, the same assay effect is likely influencing the UW4 results. These results indicate that these transcription factor predictions are either incorrect or have little to no effect on *afpA* promoter strength.

i)



ii)

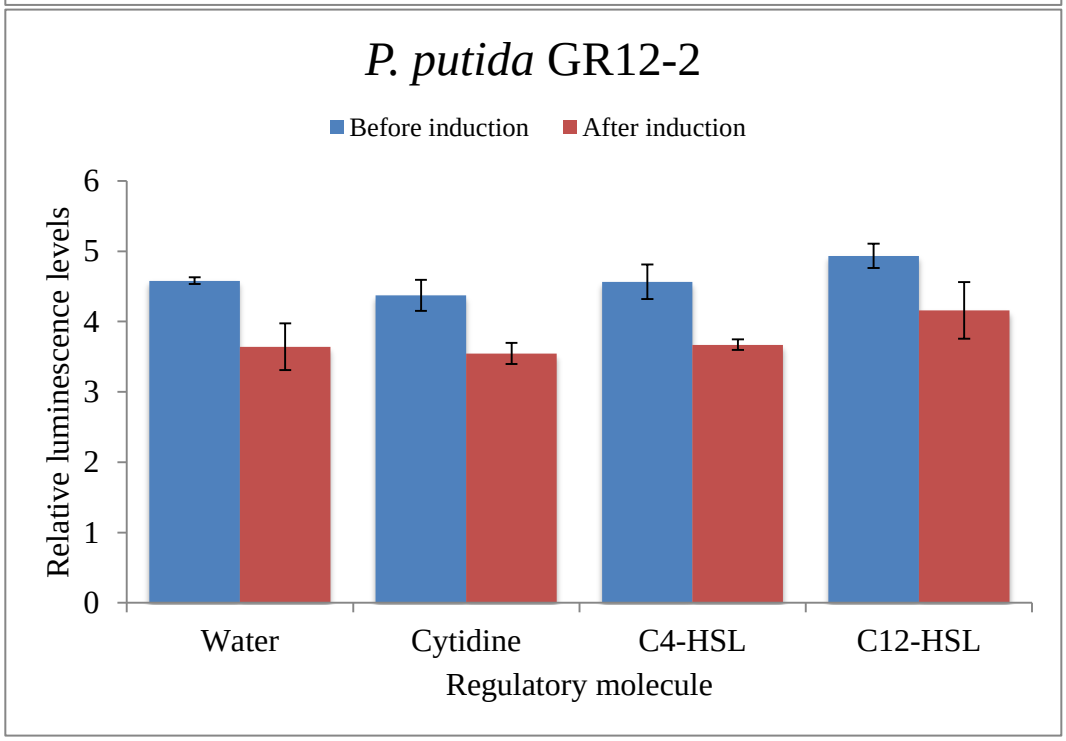


Figure 16. Relative luminescence levels before and after induction with the respective regulatory molecule. These levels were measured after growing the bacterial strain at 24°C in a M9 minimal medium. These bacterial strains contained a full promoter region construct (pQKAfp431). Water – sterile water; Cytidine – 2 mM cytidine; C4-HSL – 10 mM N-butyryl- DL-homoserine lactone; C12- HSL – 10 mM N-3-oxo-dodecanoyl-L-homoserine lactone (i) *P.* sp. UW4 (ii) *P. putida* GR12-2

To test the effects of osmoregulation (OmpR) on promoter strength, a growth medium with high salt concentration was used (Figure 17). At the time, only the UW4 strain containing the full promoter construct was tested because of contamination problems with GR12-2. Comparing growth with and without salt, salt had a small effect on activity levels. When grown with higher salt concentration, the activity levels in the UW4 strain were lower by approximately one unit. This difference is similar to the difference found using different carbon sources, but in this case it is statistically significant with a p-value less than 0.01. None of the conditions tested has shown a dramatic change in luminescence activity. Despite small differences in activity, the relative similarity in activity level across all conditions suggests that regulation of *afpA* is more complex than suggested by the simple models posited.

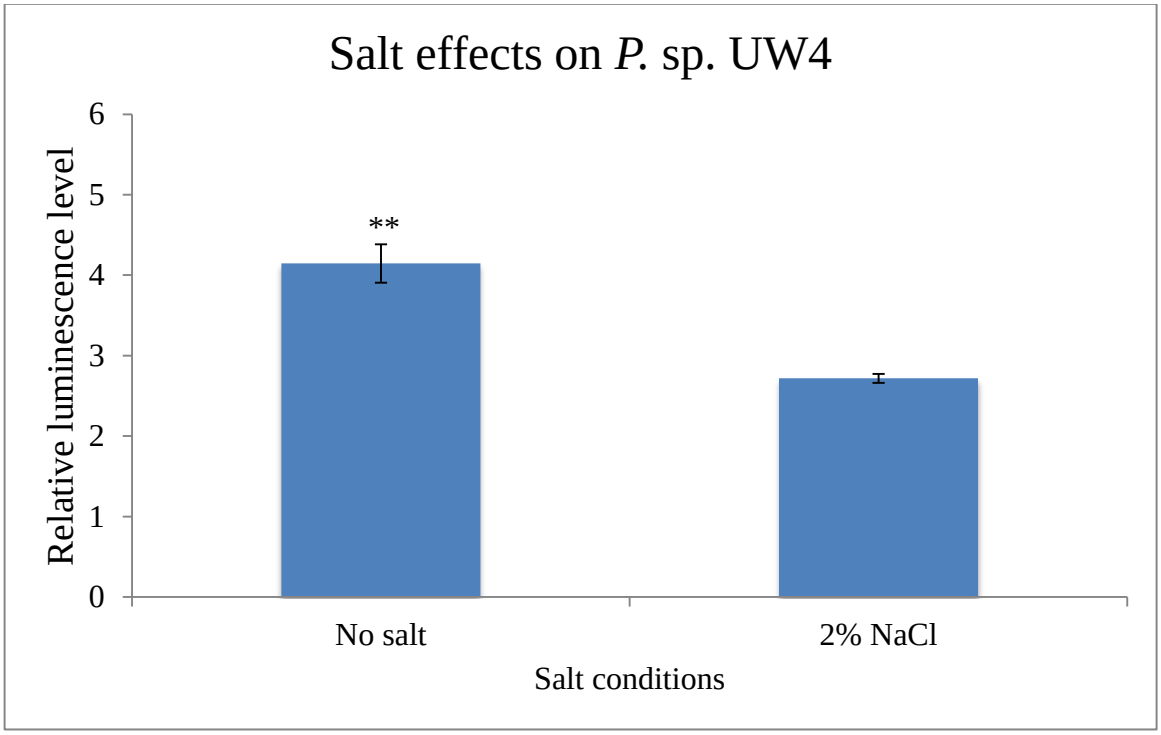


Figure 17. Relative luminescence levels at different salt conditions following growth at 24°C. M9 minimal medium with the two salt concentrations was used. These luminescence levels were measured using full promoter region construct (pQKAfp431). **p-value < 0.01

3.4 Secondary Structure Prediction of afpA RNA Transcript

Regulation of *afpA* can occur at different levels. Besides regulation at the transcriptional level, cold related proteins can be regulated at the post-transcriptional (generally translational) level (Kortmann and Narberhaus 2012; Merino and Yanofsky 2005; Yanofsky 1981). At this level, protein translation may be regulated by the secondary structure conformation of the RNA transcript. To investigate the regulation of *afpA* at the post-transcriptional level, prediction software MFOLD and RNAstructure were used to predict the secondary structure of the *afpA* RNA transcript. In both programs the temperature parameter was changed to reflect potential secondary structure at different temperatures. The temperatures tested were 37°C, 24°C and 4°C. Other optional parameters were kept at default values to minimize restrictions in structure folding.

A translational attenuator was predicted in the 5' untranslated region (UTR) with the Riboswitch program. To deduce the structure of this attenuator, the 5' UTR of the *afpA* RNA transcript was predicted using MFOLD (Figure 18). The RNA sequence used for the 5' UTR ranged from the transcriptional start site to several nucleotides past the start codon. In all temperature conditions, there are two consistent hairpin loop structures. One of these hairpin loops is the predicted translational attenuator around the 30^{th} nucleotide. Two nucleotides downstream is another hairpin loop consisting of both the Shine Dalgarno sequence and the start codon. The location of these elements permits recognition of the ribosome binding site suggesting that translation can occur in all conditions. However, the proximity of an attenuator structure upstream of this loop can influence translation initiation at all three temperatures. The program RNAstructure was

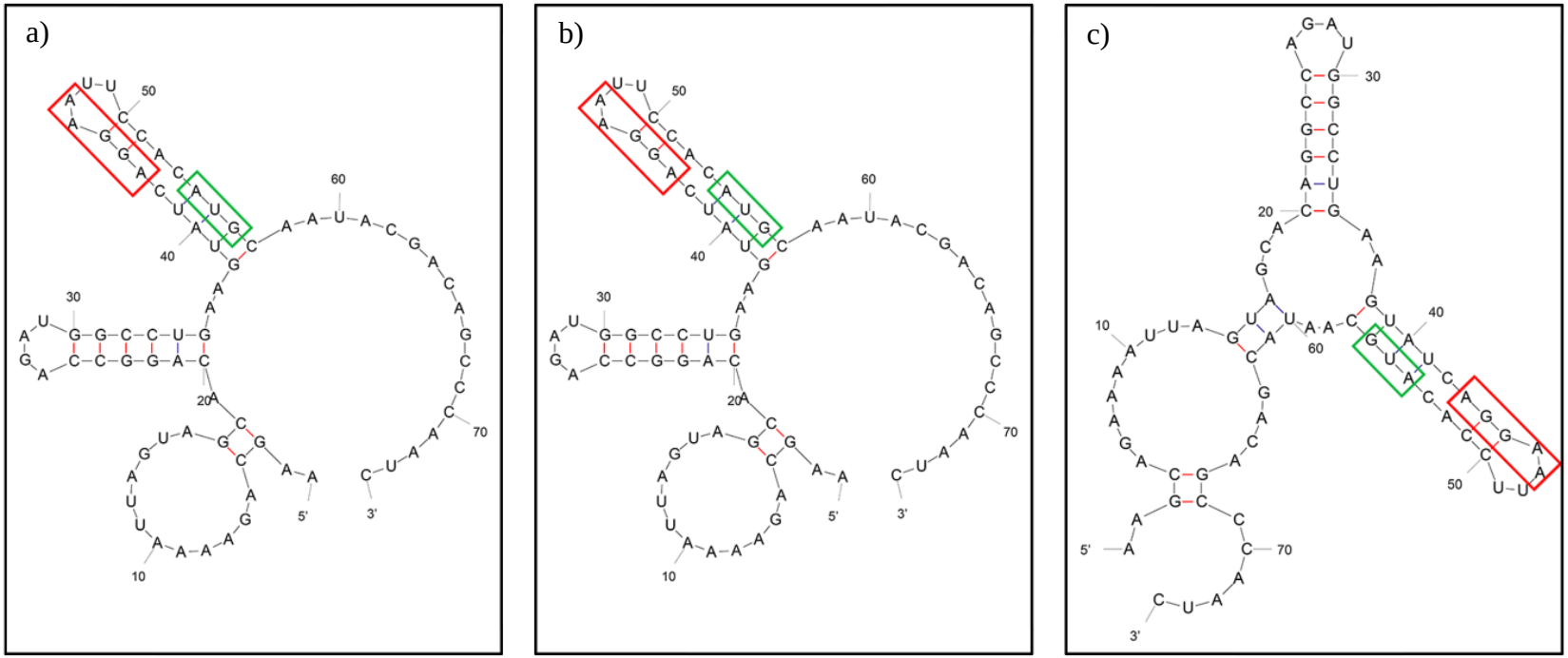


Figure 18. MFOLD predictions of RNA secondary structure of the 5' UTR region of the *afpA* RNA transcript at different temperatures. The location of Shine Dalgarno sequence is indicated in red while the start codon is indicated in green. (a) RNA at 37°C. (b) RNA at 24°C (c) RNA at 4°C.

able to confirm these predictions and predicted that the probability of these loop structures was greater than 70%. An RNA pseudoknot is formed when an RNA loop folds and base pairs with either an upstream or downstream nucleotide region of the same RNA (Cao and Chen 2006). Pseudoknots that may interfere with the predicted structures were not found.

Studies of other cold essential proteins has shown that this 5'UTR can interact with the coding sequence to alter the RNA secondary structure during post-transcriptional regulation (Giuliodori et al. 2010; Kortmann and Narberhaus 2012). Therefore, RNA structure prediction of only the 5'UTR is not sufficient. The RNA conformation of the entire *afpA* transcript was predicted using the sequence ranging from the transcriptional start site to the end of the transcriptional terminator (Figure 6). Figure 19 focuses on the 5'UTR region of these full transcript predictions. Unlike the previous MFOLD predictions, all three temperatures showed different secondary structures. Although both stem loops were present in the three predictions, the conformation of these structures were different. At 37°C, both loop structures are a part of a larger RNA fold. The Shine Dalgarno sequence is buried within the stem while the start codon is in the loop structure. This placement can affect the translation efficiency of the transcript. In contrast, at 4°C, the conformation of the 5'UTR in the full transcript was the same as in the 5'UTR prediction. In this conformation, the Shine Dalgarno sequence is on the loop portion of the stem loop and is therefore likely accessible to ribosomes. At 24°C, full transcript predictions showed that there is some interaction with the coding sequence, but not as extensive as at 37°C. Similarly, the Shine Dalgarno is trapped in the stem of a stem loop structure. However, half of the alternate predictions were the same as 4°C. This suggests

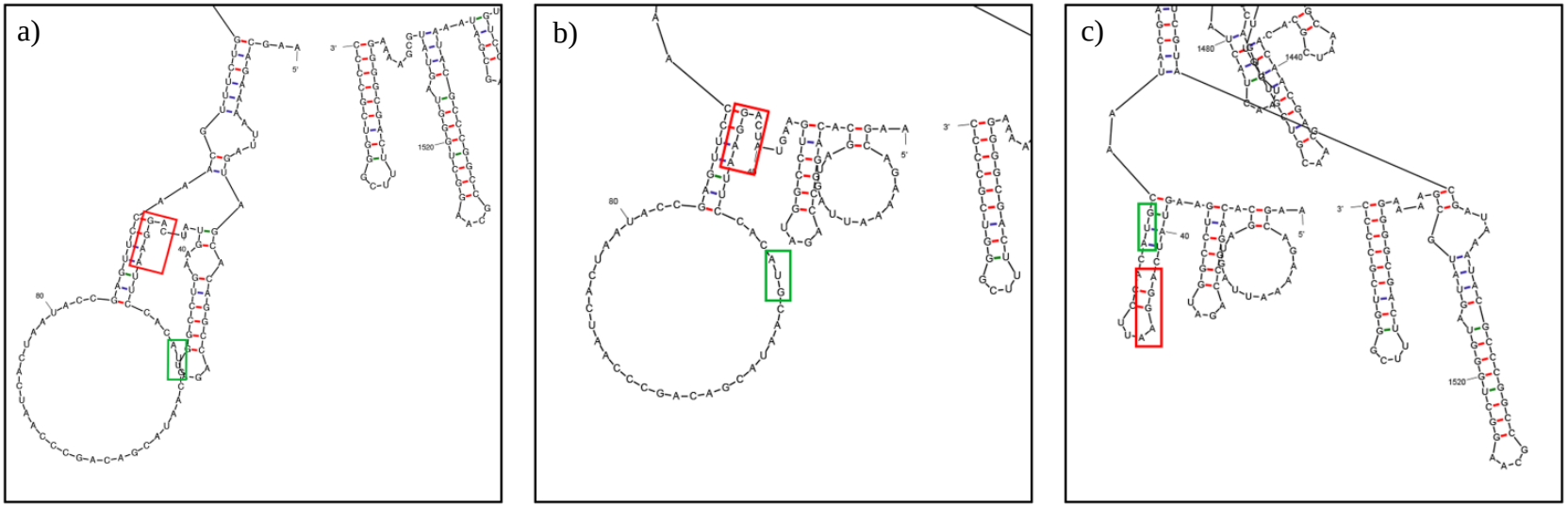


Figure 19. MFOLD predicted RNA secondary structure of the full *afpA* RNA transcript at different temperatures. Each picture selectively depicts the 5'UTR region of the RNA transcript. The location of the Shine Dalgarno sequence is indicated in red while the start codon is indicated in green. (a) RNA at 37°C. (b) RNA at 24°C (c) RNA at 4°C.

that at 24°C the "true" RNA structure is flexible and can be influenced by other factors. Once again, in all three structures the attenuator was predicted as a hairpin loop structure. The program RNAstructure predicted the same RNA conformation as the 5'UTR predictions. Pseudoknots were not found in any of the structure predictions.

3.5 Potential Polycistronic Expression of afpA Gene

When the NCBI non-redundant databases were searched using AfpA as the query sequence, numerous hypothetical protein hits from pseudomonads were returned. Some of these hits were annotated as hemolysin-type calcium binding proteins. Protein results with significant E-values were found to have greater than 45% sequence identity while covering at least 71% of the query protein (AfpA). Notably, these protein hits shared both Gly-X-Gly-X-Asp hemolysin-type calcium-binding motifs and D-U-U-U (where U represent hydrophobic residues) motifs with an unknown function found near the C-terminus. When the query and results were aligned, the calcium-binding motifs identified by Muryoi and colleagues (2004) and surrounding sequences are well conserved (Figure 20). The spacing between these motifs was relatively consistent. Between the four motifs, the outer two calcium-binding motifs are spaced 14 to 15 amino acids apart from the middle two motifs, while the middle two motifs are spaced 22 residues apart. Upon closer analysis, two conserved Gly-X-Gly-Asn-Asn sequences and one non-conserved Gly-X-Gly-X-X-Asp sequence were found within this space (Figure 20 red boxes). Whether these sequences are calcium-binding motifs is uncertain. The placement of these putative motifs increases the regularity of spacing between motifs. Six of the motifs would be spaced four to five residues apart. These spacing residues are usually either uncharged

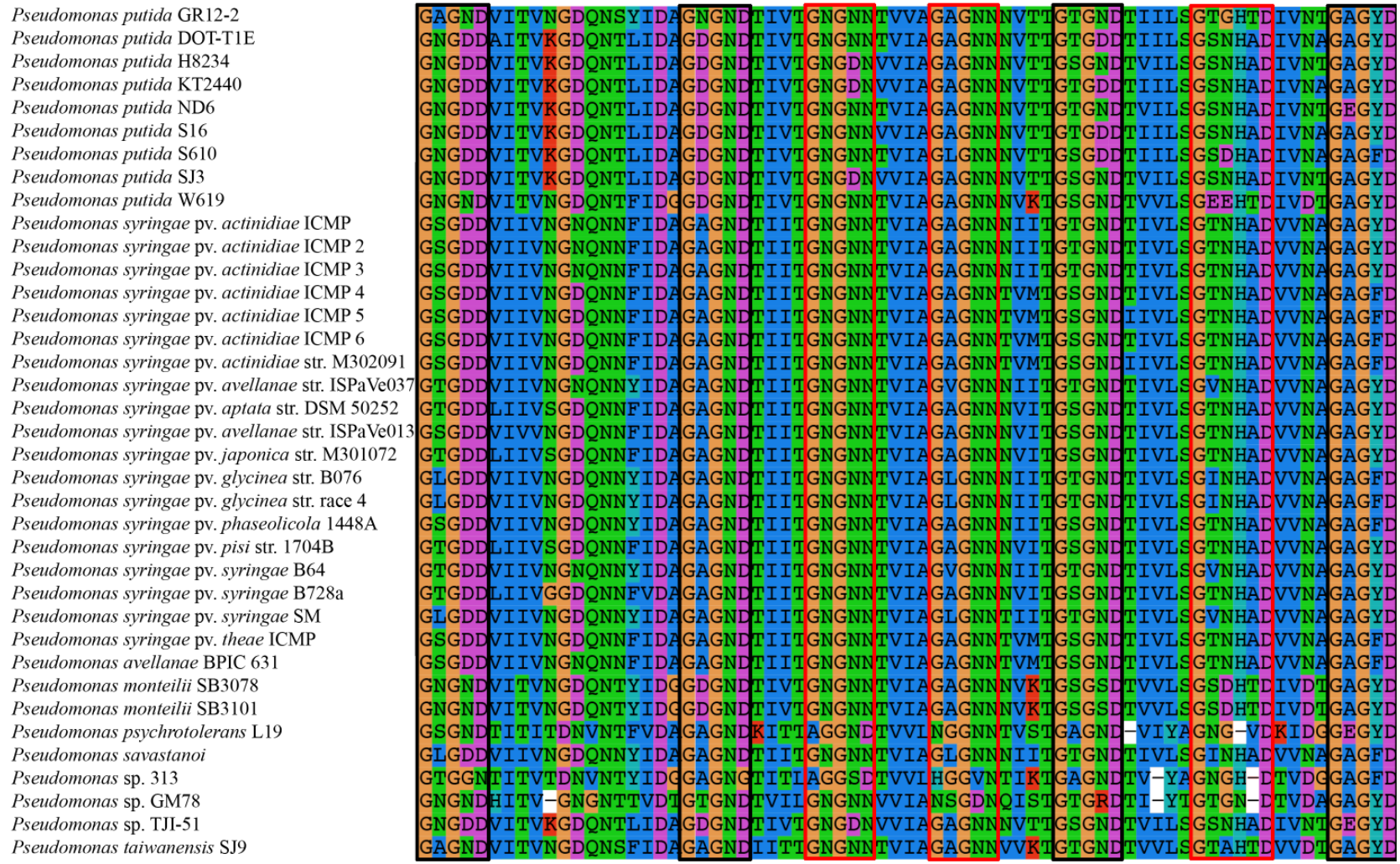
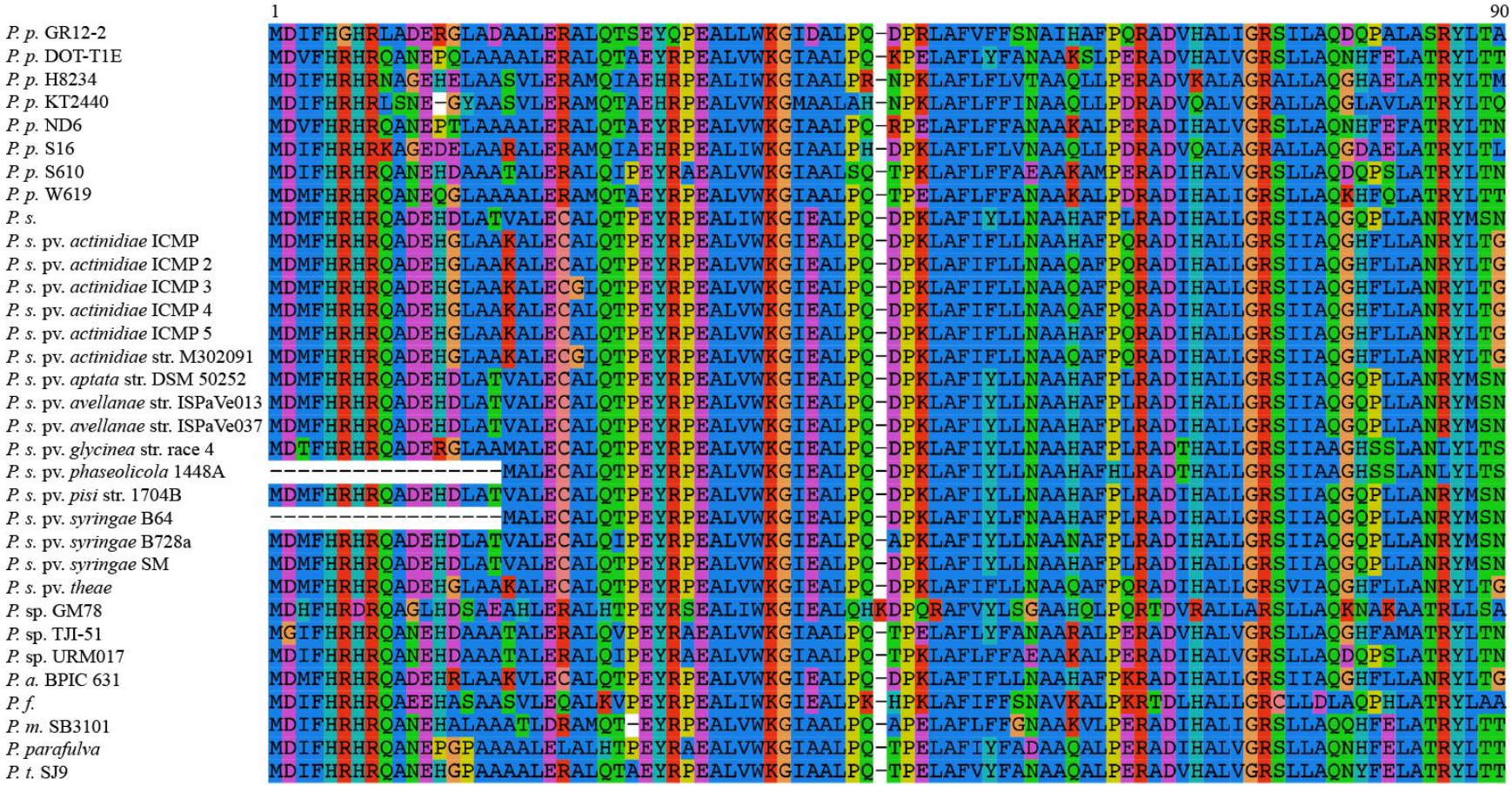


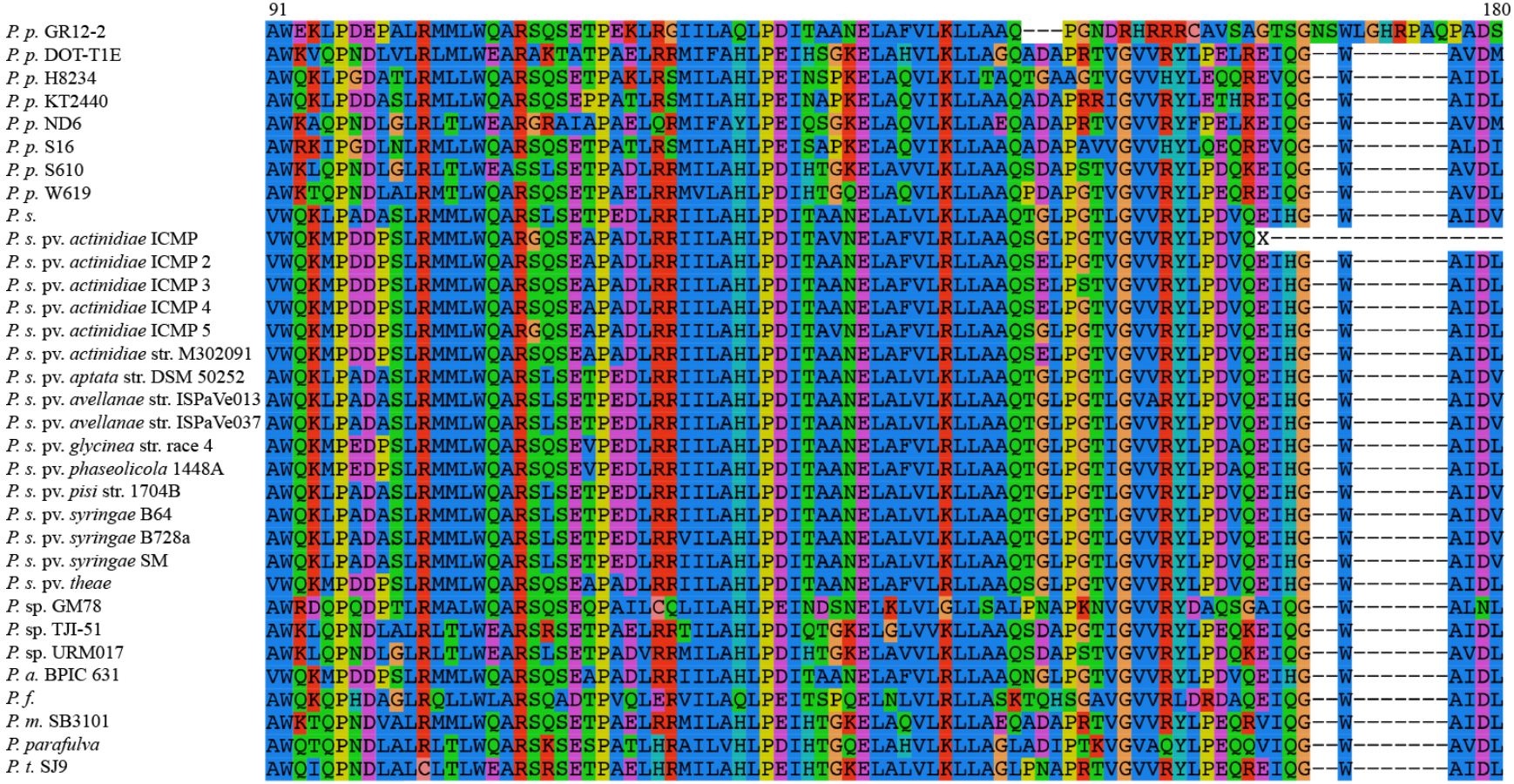
Figure 20. Multiple sequence alignment of Gly-X-Gly-X-(Asp/Asn) motifs of AfpA compared to other similar hemolysin binding protein. Black boxes represent the previously identified Gly-X-Gly-X-Asp previously identified by Muryoi et al. (2004). Red boxes represent other potential calcium binding motifs.

polar or nonpolar hydrophobic residues. The conserved sequence of this hemolysin-type calcium-binding domain suggests that these hypothetical proteins may be related to AfpA.

In fact, a blastn search of the *afpA* promoter region revealed a conserved hypothetical protein gene directly upstream of the hemolysin-type calcium binding protein genes in ten pseudomonads. One to two glycosyl transferase genes were also found downstream of these hemolysin-type calcium binding protein genes. All four genes are transcribed in the same direction. Gene proximity, arrangement and expression direction suggests that this gene cluster may be expressed as an operon in at least ten different pseudomonads (Figure 21). *P. putida* GR12-2's *afpA* gene appears to be arranged in the same manner. A blastx search highly matched the *afpA* promoter region with conserved hypothetical proteins. A putative glycosyl transferase gene with a predicted promoter was also found downstream of *afpA*. However, this putative glycosyl transferase may be a partial protein. An alignment of this putative partial protein against other *Pseudomonas* glycosyl transferases showed high similarity but low coverage. This partial protein covers the N-terminal end until around residue 256 (Figure 22). In *P. putida* GR12-2, the translated amino acids beyond the stop codon were dissimilar and did not align with the other glycosyl transferases. This indicates that the premature stop was not a sequencing error. When these putative genes in *P. putida* GR12-2 are arranged together, the pattern matches the arrangement found in the ten pseudomonads: a conserved hypothetical protein gene closely followed by a hemolysin-type calcium binding protein gene (*afpA* in GR12-2) with a downstream glycosyl transferase gene (Figure 21).

Figure 21. The arrangement of a potential operon based on BLAST results of the *afpA* promoter region. Hypo protein = hypothetical protein. Each colour represents a different gene. In *P. putida* GR12-2, *afpA* would be represented by hypothetical protein 2.


1
90
P. p. GR12-2
P. p. DOT-T1E
P. p. H8234
P. p. KT2440
P. p. ND6
P. p. S16
P. p. S610
P. p. W619
P. s.
P. s. pv. actinidiae ICMP
P. s. pv. actinidiae ICMP 2
P. s. pv. actinidiae ICMP 3
P. s. pv. actinidiae ICMP 4
P. s. pv. actinidiae ICMP 5
P. s. pv. actinidiae str. M302091
P. s. pv. aptata str. DSM 50252
P. s. pv. avellanae str. ISPaVe013
P. s. pv. avellanae str. ISPaVe037
P. s. pv. glycinea str. race 4
P. s. pv. phaseolicola 1448A
P. s. pv. pisi str. 1704B
P. s. pv. syringae B64
P. s. pv. syringae B728a
P. s. pv. syringae SM
P. s. pv. theae
P. sp. GM78
P. sp. TJI-51
P. sp. URM017
P. a. BPIC 631
P. f.
P. m. SB3101
P. parafulva
P. t. SJ9


91
180
P. p. GR12-2 AWEKLPDEPALRMMLWQARSQSETPEKLRGIILAQLPDITAANELAFVLKLLAAQ---PGNDRHRRRCAVSAGTSGNSWLGHRPAQPADS
P. p. DOT-T1E AWKVQPNDLVLRLMLWEARAKTATPAELRRMIFAHLPEIHSGKELAHVLKLLAGQADAPRTVGVVRYLPELREIQG--W--------AVDM
P. p. H8234 AWQKLPGDATLRMLLWQARSQSETPAKLRSMILAHLPEINSPKELAQVLKLLTAQTGAAGTVGVVHYLEQQREVQG--W--------AIDL
P. p. KT2440 AWQKLPDDASLRMLLWQARSQSEPPATLRSMILAHLPEINAPKELAQVIKLLAAQADAPRRIGVVRYLETHREIQG--W--------AIDL
P. p. ND6 AWKAQPNDLGLRLTLWEARGRAIAPAELQRMIFAYLPEIQSGKELAQVLKLLAEQADAPRTVGVVRYFPELKEIQG--W--------AVDM
P. p. S16 AWRKIPGDLNLRMLLWQARSQSETPATLRSMILAHLPEISAPKELAQVIKLLAAQADAPAVVGVVHYLQEQREVQG--W--------ALDI
P. p. S610 AWKLQPNDLGLRLTLWEASSLSETPADLRRMILAHLPDIHTGKELAVVLKLLAAQSDAPSTVGVVRYLPDQKEIQG--W--------AVDL
P. p. W619 AWKTQPNDLALRMTLWQARSQSETPAELRRMVLAHLPDIHTGQELAQVLKLLAAQPDAPGTVGVVRYLPEQREIQG--W--------AVDL
P. s. VWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. actinidiae ICMP VWQKMPDDPSLRMMLWQARGQSEAPADLRRIILAHLPDITAVNELAFVLRLLAAQSGLPGTVGVVRYLPDVQX----------
P. s. pv. actinidiae ICMP 2 VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQSELPGTVGVVRYLPDVQEIHG--W--------AIDL
P. s. pv. actinidiae ICMP 3 VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQSELPSTVGVVRYLPDVQEIHG--W--------AIDL
P. s. pv. actinidiae ICMP 4 VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQSELPGTVGVVRYLPDVQEIHG--W--------AIDL
P. s. pv. actinidiae ICMP 5 VWQKMPDDPSLRMMLWQARGQSEAPADLRRIILAHLPDITAVNELAFVLRLLAAQSGLPGTVGVVRYLPDVQEIHG--W--------AIDL
P. s. pv. actinidiae str. M302091 VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQSELPGTVGVVRYLPDVQEIHG--W--------AIDL
P. s. pv. aptata str. DSM 50252 AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. avellanae str. ISPaVe013 AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVARYLPDVQEIHG--W--------AIDV
P. s. pv. avellanae str. ISPaVe037 AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. glycinea str. race 4 AWQKMPEDPSLRMMLWQARSQSEVPEDLRRIILAHLPDITAANELAFVLRLLAAQTGLPGTIGVVRYLPDAQEIHG--W--------AIDL
P. s. pv. phaseolicola 1448A AWQKMPEDPSLRMMLWQARSQSEVPEDLRRIILAHLPDITAANELAFVLRLLAAQTGLPGTIGVVRYLPDAQEIHG--W--------AIDL
P. s. pv. pisi str. 1704B AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. syringae B64 AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. syringae B728a AWQKLPADASLRMMLWQARSLSETPEDLRRVILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. syringae SM AWQKLPADASLRMMLWQARSLSETPEDLRRIILAHLPDITAANELALVLKLLAAQTGLPGTLGVVRYLPDVQEIHG--W--------AIDV
P. s. pv. theae VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQSGLPGTVGVVRYLPDVQEIHG--W--------AIDL
P. sp. GM78 AWRDQPQDPTLRMALWQARSQSEQPAILCQLILAHLPEINDSNELKLVLGLLSALPNAPKNVGVVRYDAQSGAIQG--W--------ALNL
P. sp. TJI-51 AWKLQPNDLALRLTLWEARSRSETPAELRRTILAHLPDIQTGKELGLVVKLLAAQSDAPGTIGVVRYLPEQKEIQG--W--------AIDL
P. sp. URM017 AWKLQPNDLGLRLTLWEARSLSETPADVRRMILAHLPDIHTGKELAVVLKLLAAQSDAPSTVGVVRYLPDQKEIQG--W--------AVDL
P. a. BPIC 631 VWQKMPDDPSLRMMLWQARSQSEAPADLRRIILAHLPDITAANELAFVLRLLAAQNGLPGTVGVVRYLPDVQEIHG--W--------AIDL
P. f. AWQKQPHDAGLRQLLWLARSQADTPVQLERVILAQLPEITSPQELNLVLRLLASKTQHSGAVGVVRLDRDAQEIQG--W--------AIDL
P. m. SB3101 AWKTQPNDVALRMMLWQARSQSETPAELRRMILAHLPEIHTGKELAQVLKLLAEQADAPRTVGVVRYLPEQRVIQG--W--------AIDL
P. parafulva AWQTQPNDLALRLTLWQARSKSESPATLHRAILVHLPDIHTGQELAHVLKLLAGLADIPTKVGVAQYLPEQQVIQG--W--------AVDL
P. t. SJ9 AWQIQPNDLALCLTLWEARSRSETPAELHRMILAHLPEIHTGKELALVLKLLAGLPNAPRTVGVVRYLPEQREIQG--W--------AIDL

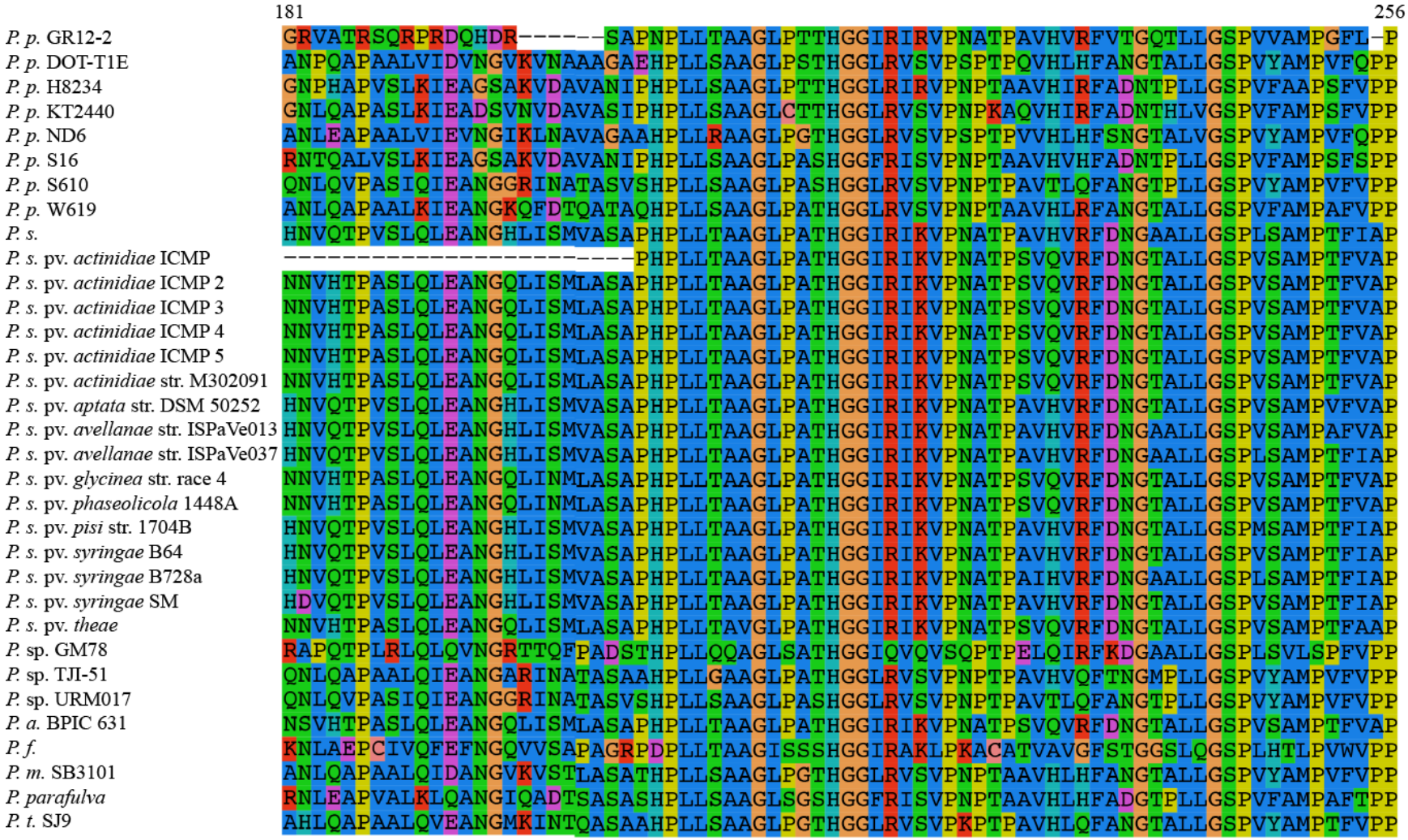


Figure 22. Multiple sequence alignment of the downstream glycosyltransferase proteins. Alignment includes the first 256 amino acid sequence except the partial glycosyl transferase protein from *P. p.* GR12-2. *P. p. = Pseudomonas putida*; *P. s. = Pseudomonas syringae*; *P.a. = Pseudomonas avellanae*; *P.f. = Pseudomonas fulva*; *P.m. = Pseudomonas monteilii*; *P.t. = Pseudomonas taiwanensis*.

A preliminary phylogenetic analysis was done on AfpA and hemolysin-type calcium binding proteins to elucidate a potential relationship between the proteins (Figure 23). In this tree, there are two major clusters: a *P. syringae* cluster and a cluster containing the other pseudomonads. *P. putida* GR12-2 was clustered closer to plant based *P. syringae* strains rather than soil and wastewater based *P. putida* strains. However, *P. putida* GR12-2 appeared as an outlier of the *P. syringae* cluster implying a weak link to either cluster. When the *Pseudomonas* strains containing the putative operon were identified within this tree, the *P. syringae* strains were unsurprisingly grouped within the *P. syringae* cluster. The other *Pseudomonas* strains were in the other major cluster. Across all strains in this phylogenetic tree a glycosyl transferase has been identified downstream of the hemolysin-type calcium binding protein differing only by an upstream conserved hypothetical protein. This grouping suggests that this putative operon may have been altered or lost in the *P. syringae* cluster. The majority of the other strains in the *Pseudomonas* cluster retained this putative operon.

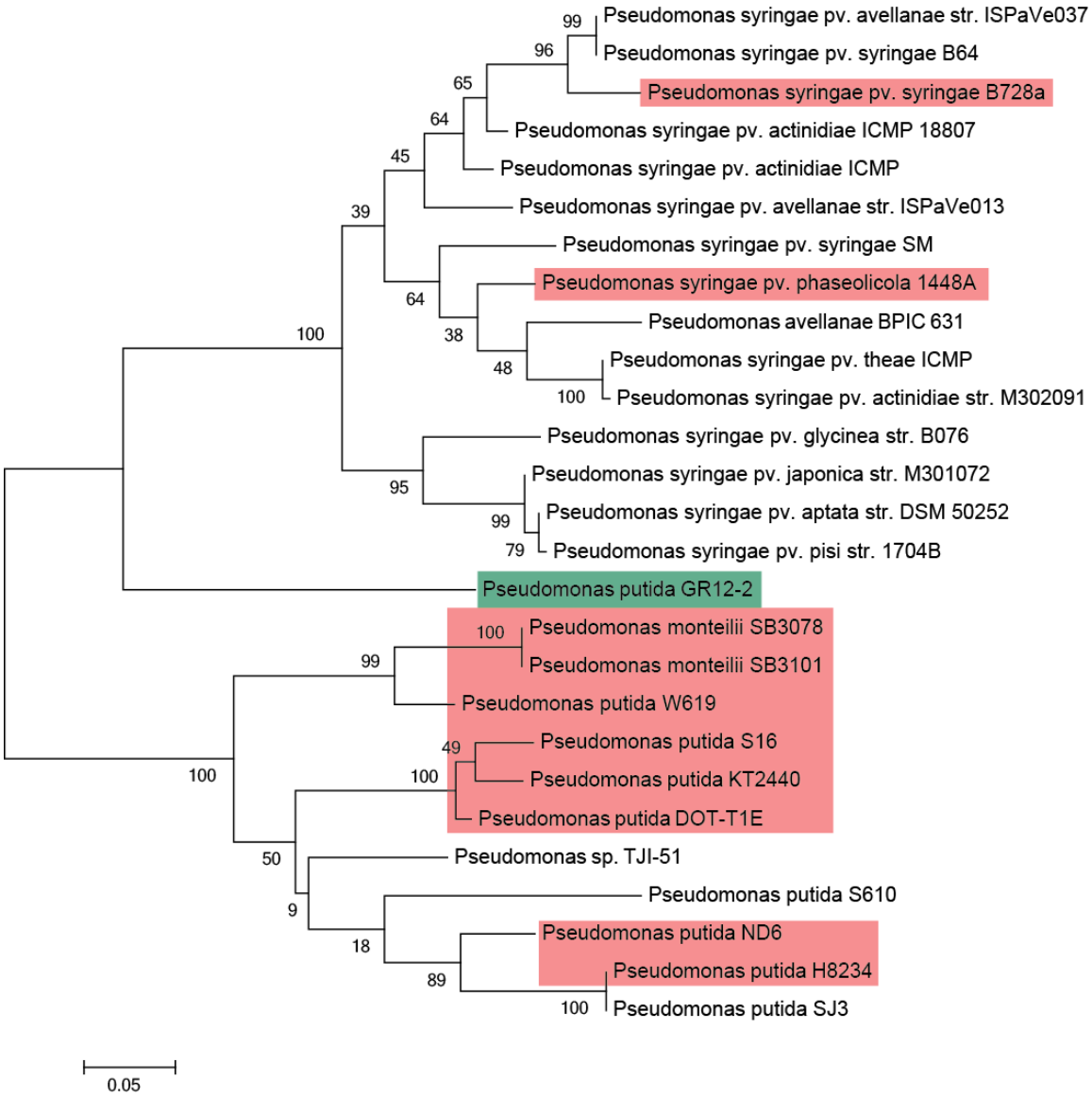


Figure 23. Phylogeny tree of AfpA and hemolysin-type calcium binding proteins in pseudomonads. Target organism *P. putida* GR12-2 is highlighted in green. Ten *Pseudomonas* strains found with a putative operon with an upstream conserved hypothetical protein are highlighted in pink.

3.6 Isolation of Partially Purified Native AfpA

Native AfpA was isolated following five days of growth at 4°C in TSB medium. After ammonium precipitation, the concentrated extracellular proteins were visualized by SDS-PAGE. A slightly different protein profile was seen in this extracellular protein sample than that observed in a previous study by Xu et al. (1998). AfpA was still present as a ~164 kDa diffuse band, however, the other proteins were slightly different (Figure 24). Instead of a prominent 45 kDa band, the previously faint 33 kDa band was more prominent. The 45 kDa protein was a faint band. Four other unidentified bands approximately 75 kDa, 65 kDa, 50 kDa and less than 29 kDa in size were also detected. These new bands may be a product of slight alterations in protocol and are not suspected to significantly affect antifreeze protein activity.

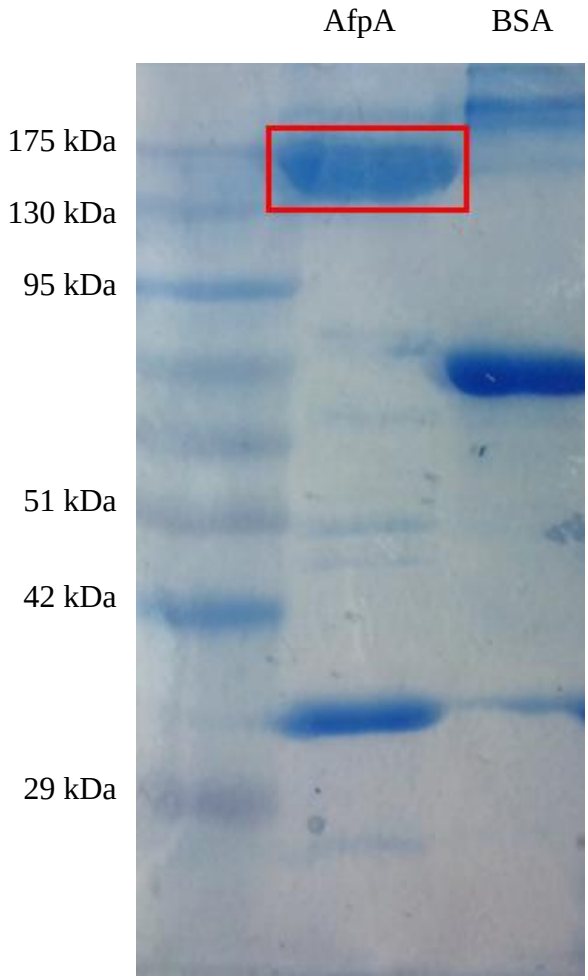


Figure 24. Polyacrylamide gel of the partially purified native AfpA. BSA was used as the loading control.

3.7 Expression and Partial Purification of Recombinant AfpA Protein Fragment (pETAFrag)

Using the protein structure prediction software Phyre 2.0, a fragment of the full native AfpA protein was modeled into a β-solenoid structure using antifreeze protein *Mp*AFP as the template. This protein fragment contained at least four Gly-X-Gly-X-Asp motifs that are similar to the calcium binding motif X-Gly-Thr-Gly-Asn-Asp of *Mp*AFP. Because of these similarities, an AfpA protein fragment was expressed using the pET vector system with pET30b(+). To create the pETAFrag construct, the latter 1062 bp of the *afpA* coding sequence were PCR amplified with primers containing the NdeI and XhoI restriction enzyme sites (Figure 25a). This amplified sequence did not contain a stop codon and was positioned in frame with the C-terminal His-tag and the vector's internal stop codon. A C-terminal His-tag was chosen to avoid potential tag interference with the N-terminal Gly-X-Gly-X-Asp motifs in the recombinant protein fragment. These PCR products and pET30b(+) were then digested with NdeI and XhoI, ligated together with T4 DNA ligase and then transformed into *E. coli* DH5α. After isolating the plasmids in transformants, these constructs were confirmed by double digestion with NdeI and XhoI (Figure 25b). The remaining plasmids were then transformed into *E. coli* BL21, an expression host strain lacking Ion and OmpT proteases.

Unlike the study by Muryoi and colleagues (2004), the use of a different vector and expression host required slight changes in the optimization of the expression protocol. Recombinant protein expression in whole cell extracts was monitored for the highest level of soluble expression under different growth and induction conditions. Uninduced cells were unable to express the recombinant protein compared to induced

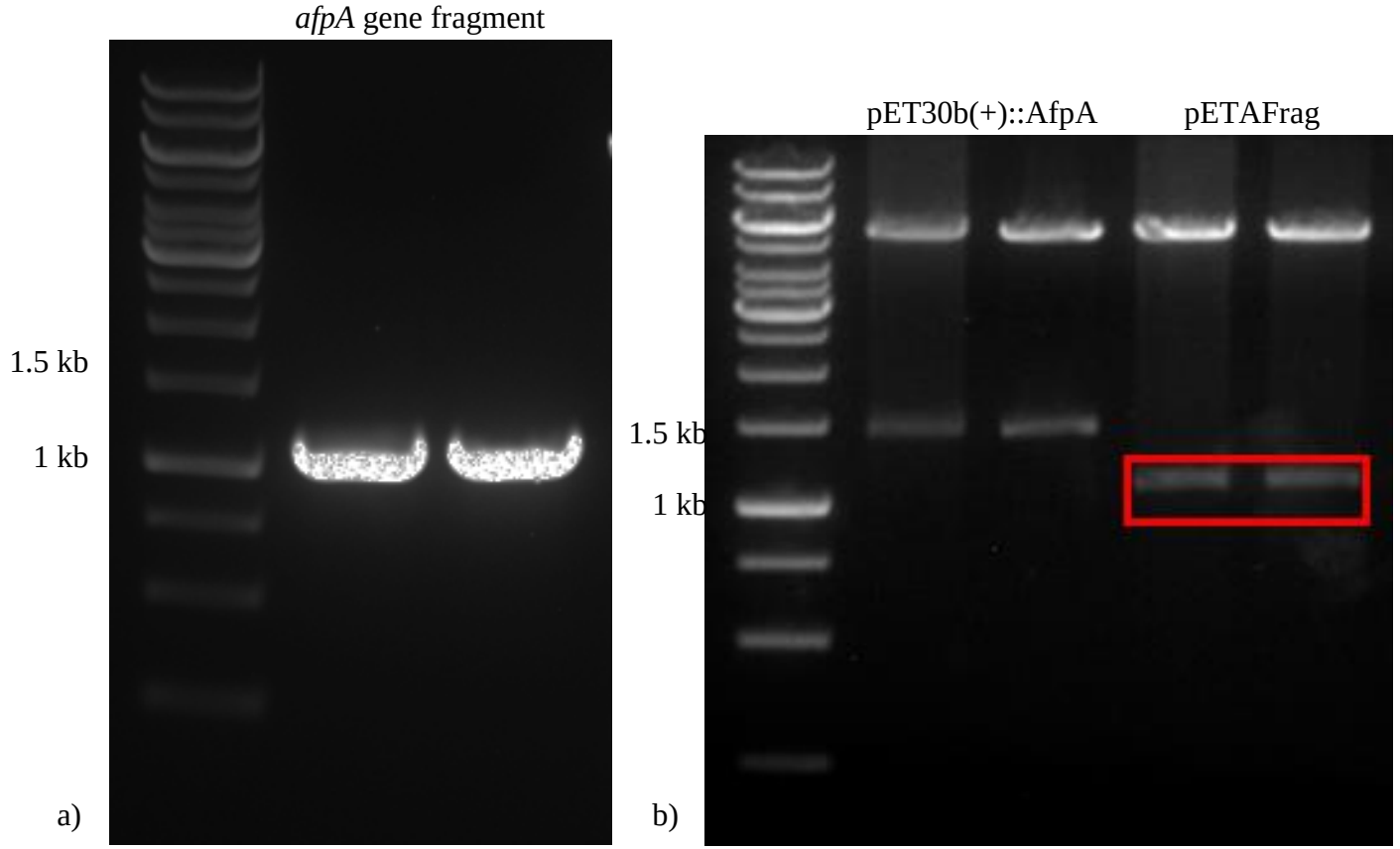


Figure 25. DNA manipulation of the *afpA* coding sequence gene fragment from *P. putida* GR12-2 genomic DNA. Partial *afpA* gene is 1062 bp in length. The DNA ladder used was Fermentas gene ruler 1 kb ladder. (a) Amplification by Hot start KOD polymerase. (b) Confirmation of *afpA* coding sequence fragment insert (highlighted in red) in pET30b(+) vector by double digestion with restriction enzymes NdeI and XhoI. Samples were prepared in duplicates.

cells which were able to express a ~48 kDa protein (Figure 26a). These conditions were then manipulated with respect to growth temperature, induction temperature, induction time and IPTG concentration for induction. Growth at 30°C and induction with 0.4 mM IPTG for 4 hours at room temperature yielded the highest level of soluble protein expression. After cell sonication, the soluble protein was purified with HisPur® resin, which binds to His-tagged recombinant protein in the soluble fraction. The recombinant protein was eluted at three different concentration of imidazole: 20 mM, 50 mM and 100 mM (Figure 26b). To maximize elution, large volumes of a 50 mM imidazole solution were used to continuously elute the recombinant protein. Out of the three conditions, a 20 mM imidazole solution was not used because a portion of the protein remained bound to the resin. A 100 mM imidazole solution was not used to minimize potential effects of imidazole on the recombinant protein when a lower concentration was sufficient for elution.

Since a high volume of elution buffer was used, the dilute protein sample was concentrated using an Amicon® centrifugal filter device with a MWCO of 30 kDa. Even with this MWCO, some proteins smaller than 30 kDa remained in the sample (Figure 26c). These smaller proteins can be subunits of larger multimeric proteins greater than 30 kDa. The recombinant antifreeze protein fragment is similar to the native protein. When high concentrations of the recombinant protein were resolved using SDS-PAGE, the band did not increase in intensity but became a diffuse band suggesting that the recombinant protein has a protein size range, which was investigated using size exclusion chromatography. A concentrated protein sample (~1 mg/ml) was used in downstream experiments.

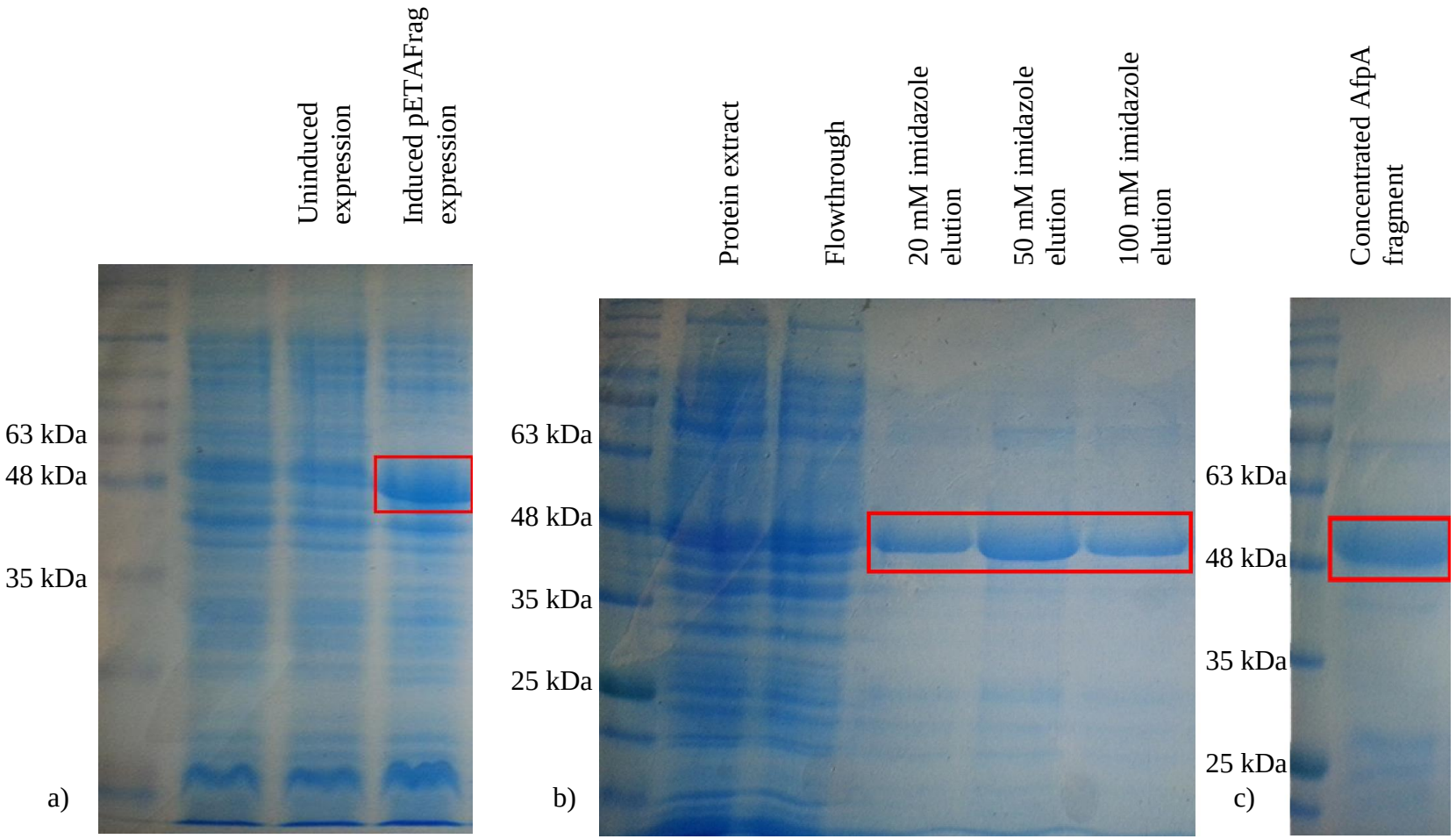


Figure 26. SDS-PAGE of the expression and purification of recombinant AfpA protein fragment. (a) Cellular protein profile before and after recombinant protein expression 0.4 mM IPTG. (b) HisPur® nickel resin purification of recombinant AfpA protein fragment. Recombinant protein was eluted at three different imidazole concentrations. (c) Recombinant AfpA fragment after concentration using amicon® centrifugal devices.

3.8 Size Exclusion Chromatography of Native AfpA and Recombinant AfpA Fragment

Size exclusion chromatography was used in an attempt to determine the native size and further purify both native and recombinant AfpA protein. From SDS-PAGE results, native AfpA was approximately 164 kDa while the recombinant AfpA was around 48 kDa in size (Figure 24 and Figure 26). To determine the relationship between protein size and elution volume, gel filtration standards containing a variety of proteins were separated with size exclusion chromatography under the same conditions as AfpA. A standard curve was created by plotting protein molecular weight against the elution fraction volume on a semi-log graph. This curve was used to determine the native sizes (Figure 27). Blue dextran (2000 kDa) was used to determine the void volume. It eluted in the same fraction as bovine thyroglobulin indicating that proteins eluting in these fractions were eluted at the void volume. This determines that the size exclusion column is unable to separate proteins around 670 kDa or greater.

When 2 ml protein fractions were collected, both native and recombinant AfpA proteins were collected in the 96 ml fraction, the same fraction as bovine thyroglobulin. For both cases, the AfpA proteins eluted at the void volume. Based on the standard curve, this elution volume suggests that both native and recombinant AfpA have a molecular weight of 500 kDa or greater. However, with a denaturing method (SDS-PAGE) both proteins were reported as smaller sizes suggesting that either the protein is a large multimeric structure or it exists as soluble aggregates. During elution of native AfpA, the largest protein peak was determined at the void volume, however, this peak formed a large trailing tail that covered all elution volumes. When random fractions of this trailing tail were resolved using SDS-PAGE, all fractions faintly contained AfpA.

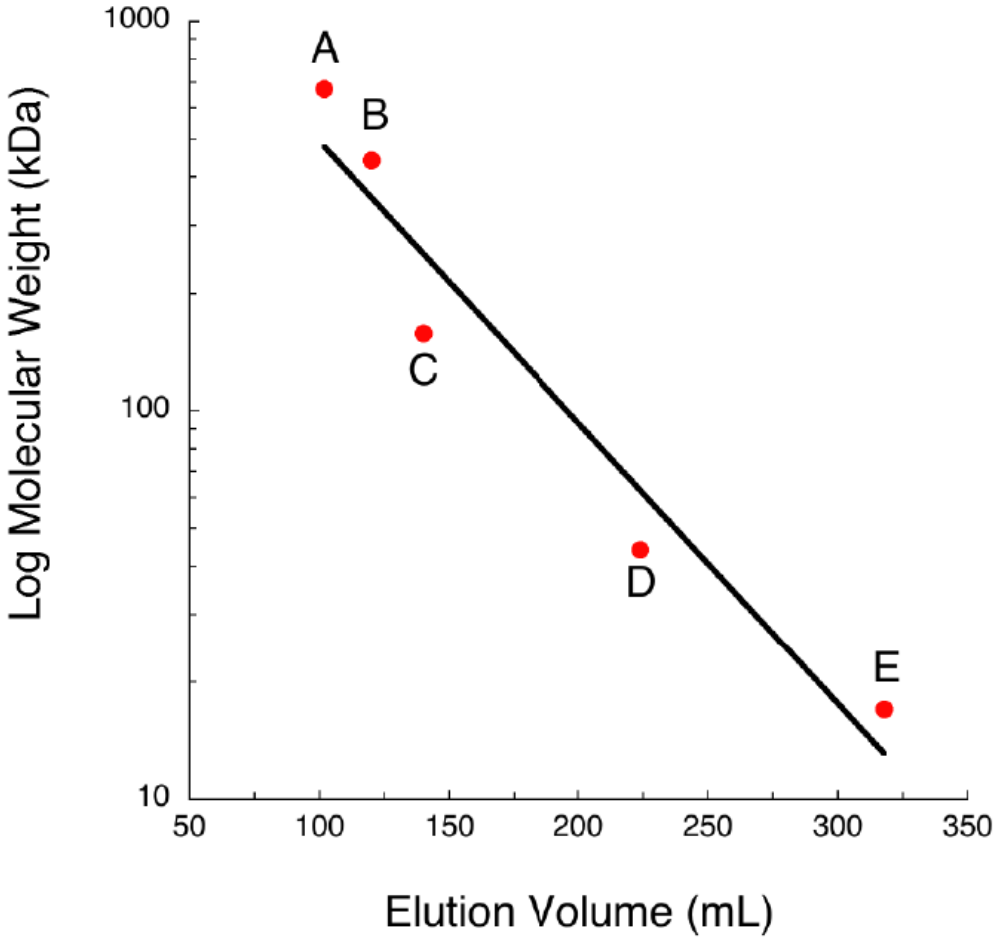


Figure 27. Size exclusion chromatography standard curve of gel filtration protein standards. (A) Bovine thyroglobulin – 670 kDa (B) Horse ferritin – 440 kDa (C) Rabbit aldolase – 158 kDa (D) Chicken ovalbumin – 44 kDa (E) Horse myoglobin – 17 kDa. (Duca 2013)

3.9 Commercial Purification of AfpA

Samples of native AfpA were given to Advanced Electrophoresis Solutions Ltd (Cambridge, ON) in an attempt to further purify the protein. Using high performance capillary electrophoresis, the company was able to determine that a sample of partially purified native AfpA yielded a high protein peak corresponding to a 164 kDa antifreeze protein (Figure 28). However, within this large protein peak, there were three other subpeaks of relatively high intensity. The company suggests that these subpeaks are caused by impurities in the sample (Olivia Li, personal communications, September 5, 2013). Their analysis also indicated that the protein's pI ranged between 4.0 to 5.0 and not 5.6 as previously reported (Xu et al., 1998). After several attempts, they suspected that difficulties in purification arose due to hydrophilic and hydrophobic pockets on the protein (Olivia Li, personal communications, September 18, 2013). However, they were able to purify the protein following denaturation with an 8 M urea solution but refolding it into an active antifreeze protein was unsuccessful.

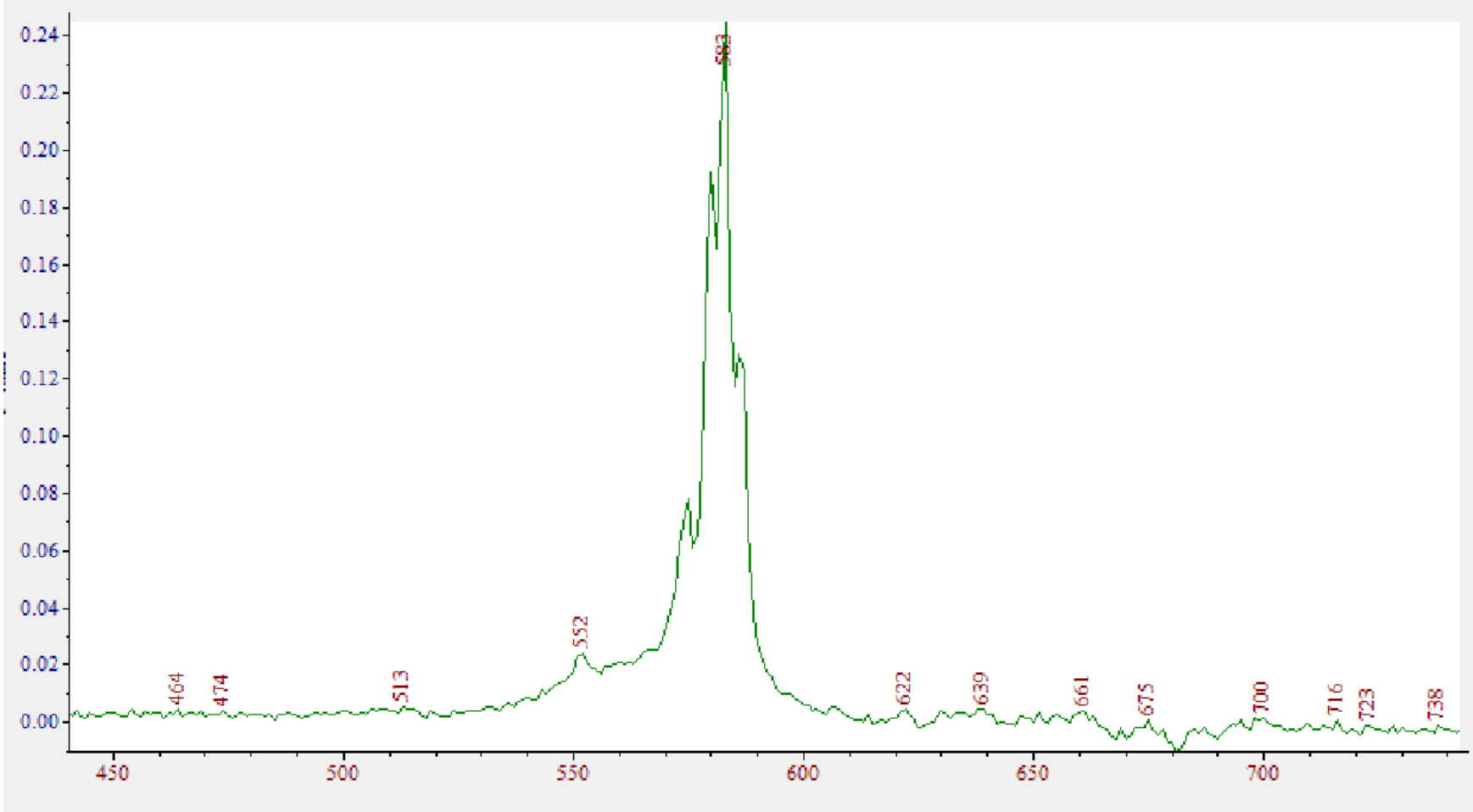


Figure 28. Native AfpA purification using high performance capillary electrophoresis (Advanced Electrophoresis Solutions Ltd, Cambridge, ON). The main peak indicates partially purified AfpA. Smaller peaks indicate impurities in the sample.

3.10 Antifreeze Assays of Native AfpA and Recombinant AfpA Fragment

The antifreeze activity of both native and recombinant AfpA proteins were tested with a thermal hysteresis assay and an ice recrystallization inhibition assay. Thermal hysteresis activity was confirmed by direct observation of ice crystal morphology during changes in temperature. Ice recrystallization inhibition activity was qualitatively confirmed by observing changes in ice crystal size at sustained high subzero temperatures. Native AfpA (~ 1mg/ml) was confirmed to have both antifreeze activities (Figure 29 and Figure 30). To test the effect of calcium on this hemolysin-type calcium binding protein, antifreeze activity was tested when the protein buffer was supplemented with 10 mM $CaCl_2$. In bacterial *Mp*AFP, calcium ions were necessary for antifreeze activity. This necessity was attributed to X-Gly-Thr-Gly-Asn-Asp calcium binding turns required for structural stability (Garnham et al. 2008). Although AfpA has these calcium binding motifs, in this study and a previous study, it has been confirmed that AfpA has antifreeze activity in buffers without calcium ions (Figure 29c and d; Muryoi et al. 2004, Xu et al. 1998). In this study, with 10 mM $CaCl_2$, AfpA antifreeze activity was unaffected (Figure 29e and f). In fact, thermal hysteresis was found to be 0.05°C and ice crystals were observed to form hexagonal or hexagonal bipyramids in buffers with or without calcium. Native AfpA was also confirmed to have ice recrystallization inhibition activity when ice crystal grains were observed to maintain grain size at -6°C for 60 minutes (Figure 30b).

The recombinant AfpA protein fragment was also tested for antifreeze activity. For the thermal hysteresis assay, the recombinant AfpA protein was unable to arrest ice crystal growth nor shape ice crystal morphology (data not shown). However, thermal

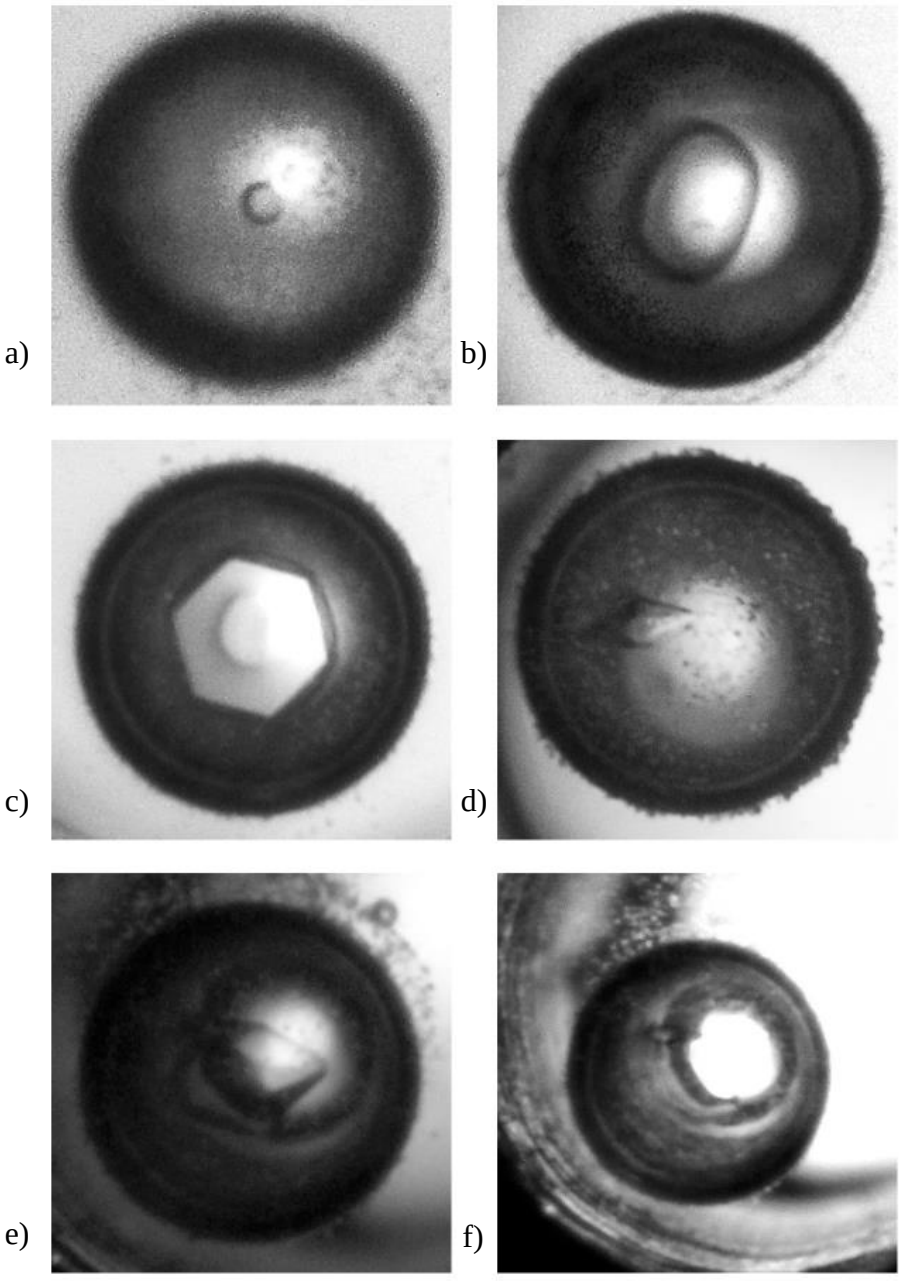


Figure 29. Ice crystal morphology during thermal hysteresis assay. Morphology can vary depending on ice growth stage observed All aliquots of AfpA used were ~1 mg/ml. (a) Negative control (protein buffer): initial isolation of a single ice crystal. (b) Negative control: uniform ice crystal growth. (c) Native AfpA: hexagonal crystal. (d) Native AfpA: bipyramid hexagonal crystal. (e) Native AfpA and 10 mM $CaCl_2$: partially formed bipyramid hexagonal crystal. (f) Native AfpA and 10 mM $CaCl_2$: bipyramid hexagonal crystal.

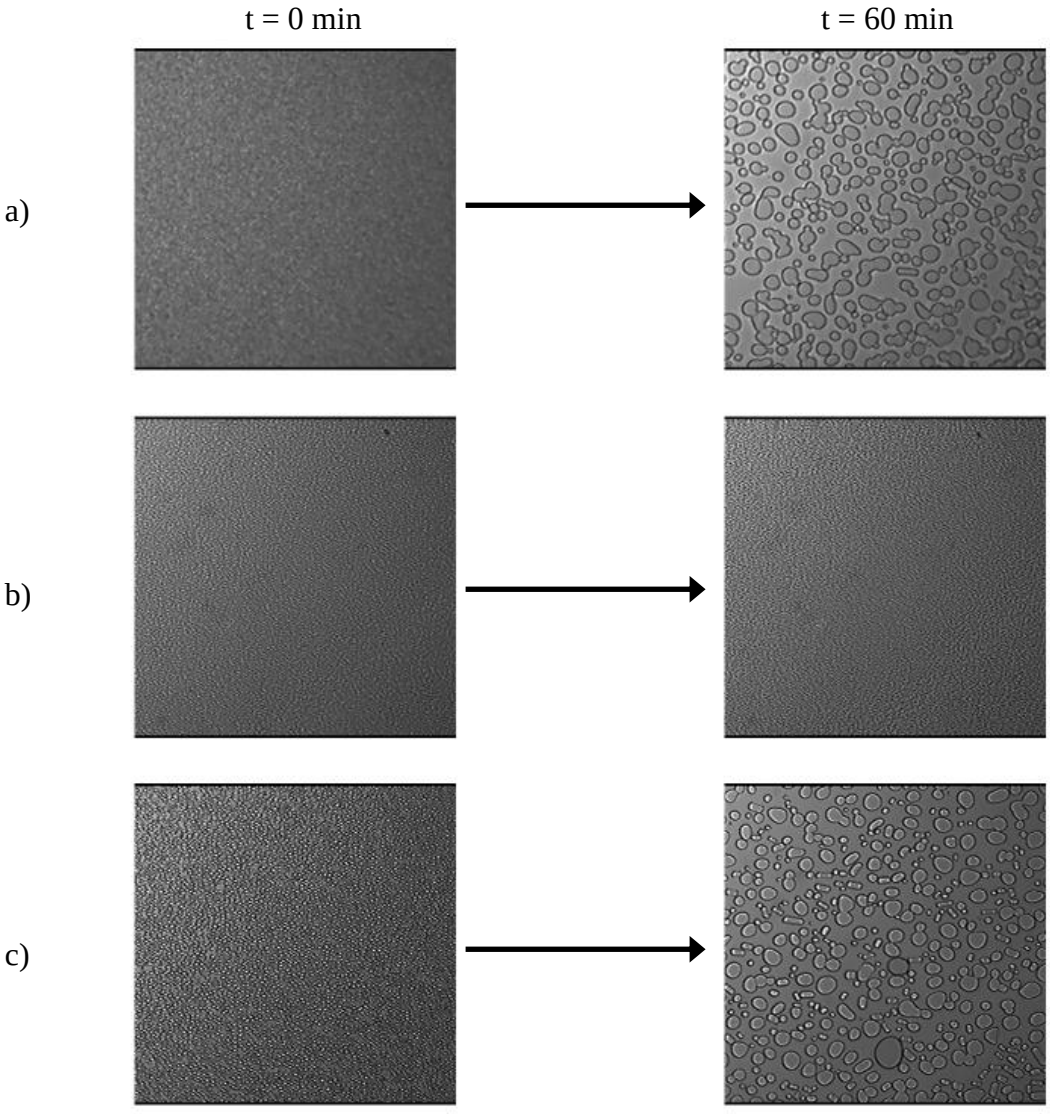


Figure 30. Changes in ice crystal grain morphology during ice recrystallization inhibition assay. (a) Negative control (protein buffer). (b) Native AfpA (~1 mg/ml). (c) Recombinant AfpA protein fragment (~1 mg/ml).

hysteresis activity requires a high concentration of active protein in order to detect activity. There may not have been sufficient active recombinant AfpA fragments for activity detection using a thermal hysteresis assay. To confirm inactivity, this recombinant protein was tested for ice recrystallization inhibition because only a low concentration of 10 μg/ml of active protein is required. At -6°C, ice crystal grains gradually recrystallized into larger ice crystals over time. This assay confirmed the inactivity of the expressed recombinant AfpA fragment.

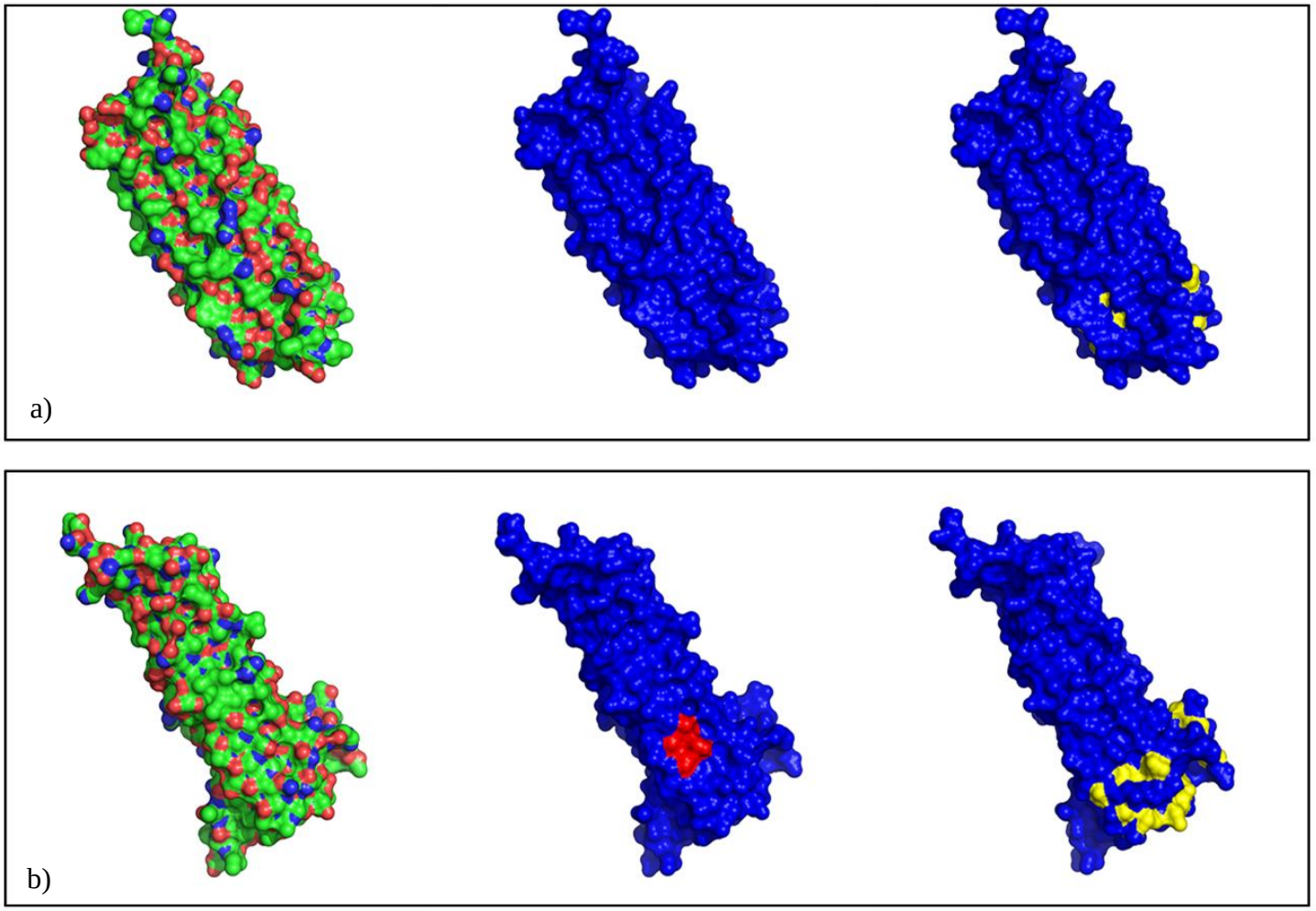
3.11 Protein Structure and Ice Binding Surface Prediction.

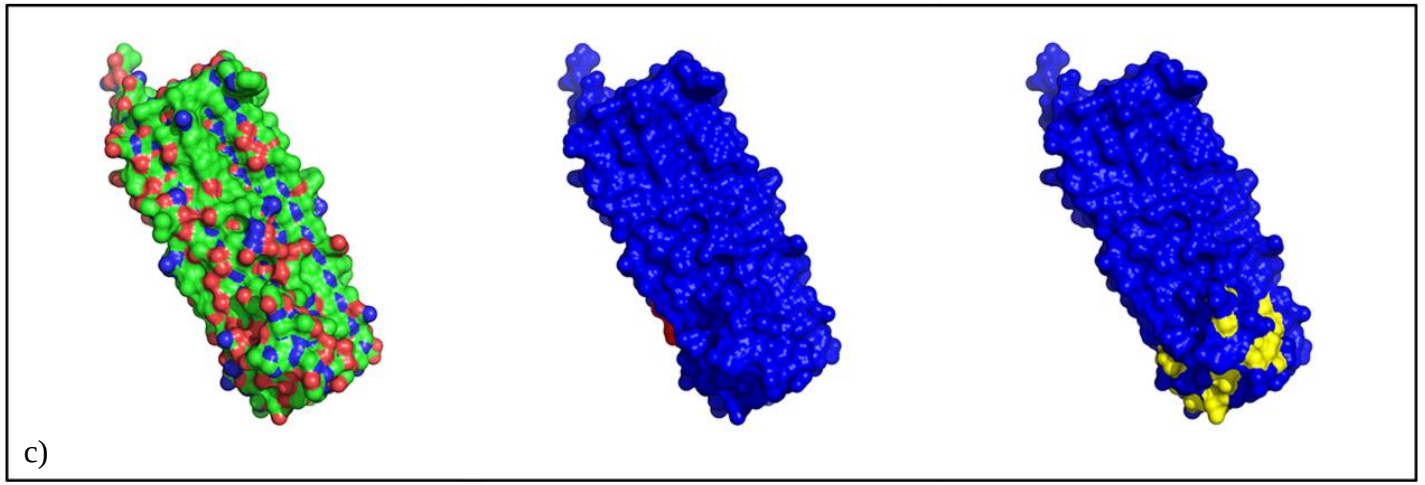
Using the amino acid sequence, AfpA protein structure was predicted using the Phyre2.0 server (Figure 31). Scanning a protein structure database, this program finds protein folds with similar protein patterns as the query sequence using PSI-BLAST. It then threads these areas of similarity together for a probable protein structure. One of the top scoring structures used bacterial *Mp*AFP as the template structure. This template folded AfpA into a β-solenoid structure similar to *Mp*AFP. Using this structure, attempts were made to determine a potential ice binding surface. The structure file was scanned with AFPredictor to find surface areas with ordered surface carbons, solvent accessible residues capable of binding to ice. Amino acid rotamers of the AfpA predicted structure were not altered. Using the recommended solvent accessibility threshold, four ordered surface carbons were detected on one side of AfpA (Figure 32). Five additional ordered surface carbons, surrounding the original four, were detected if the solvent accessibility threshold was decreased (data not shown). A manual analysis of potential ice binding residues did not reveal any distinguishable surface residues commonly found in other antifreeze proteins, i.e., Thr-X-Thr or Thr-X-Asn motifs (Garnham et al., 2008, Middleton et al., 2012). Unlike bacterial *Mp*AFP, the predicted calcium binding motifs of AfpA were not aligned on one side of the protein structure. Due to motif spacing, these motifs were scattered in the N-terminal side of the predicted structure. This predicted structure suggests that these unaligned motifs are not involved in ice binding in AfpA.

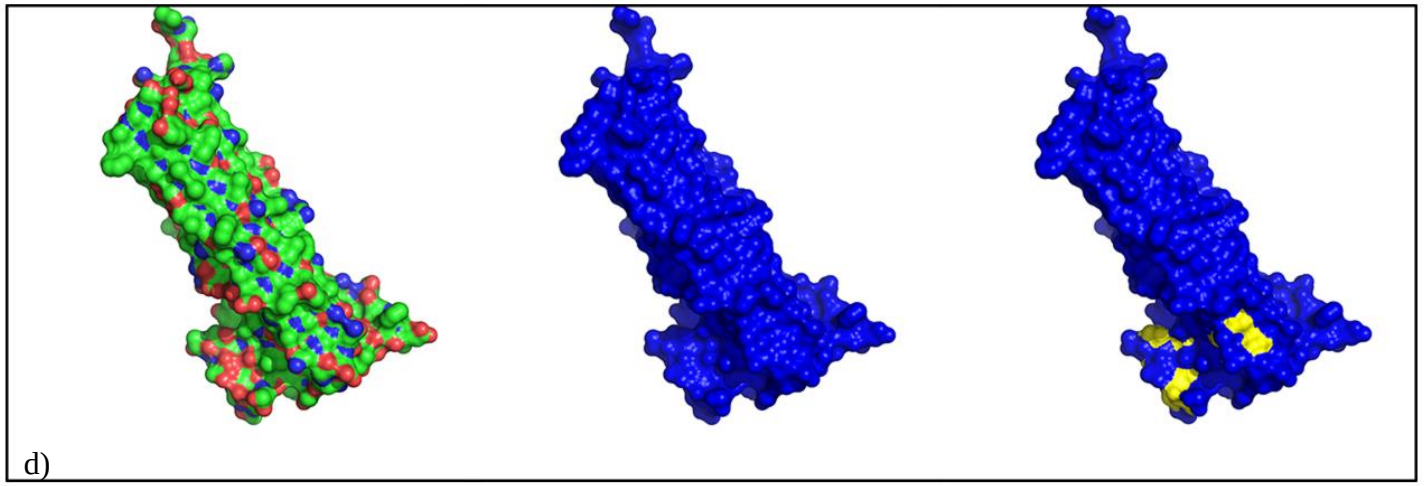
```
  1 MQYDSPITNT EFQTFLTTSS ISDDTAAAIS TLLNLDSADT INLASWDGVN APEIPTGQEG
 61 AADVVIVNVP GAATDLVPVE IPDSLNSAKA FIFDSNASLA VTFDAPVAAE SASLARVAAD
121 TTAGIEFLVT TGAGNDVITV NGDQNSYIDA GNGNDTIVTG NGNNTVVAGA GNNNVTTGTG
181 NDTIILSGTG HTDIVNTGAG YDVVQLDGSA ADYTITAGNS NNVTLTGAQT AAITGAEFLT
241 FADGSSVALA QSEAEASALR LYEGILGRDA DQGGAQNFIA QVEAGTALTD IANSFLNSAE
301 FGGAATEASI DSLYTSLLGR GADTAGSDSW EAIIANGGSL ADVAAGIAGS AEAQEQDQSN
361 GTFVDSLYLN ALGRPSDEAG HDAWVAQLFN GASRAEVAAG IVGSAEAAEK INSDFIDALY
421 LSATGRASDE AGKAGWTEVL ANGGTQADVA IGIVGSQEAI AHNDNVVVLH GAV
```

Figure 31. AfpA amino acid sequence. The amino acid sequence used for protein structure prediction.

Figure 32. The AfpA predicted protein structure surface visualized using pyMOL. Each protein structure subset shows a regular protein surface, a structure with AFPredictor results mapped on the protein surface (red) and a structure with Gly-X-Gly-X(Asp/Asn) locations mapped on the protein surface (yellow). For each structure, C-terminal side is on top; N-terminal side is on the bottom. (a-d) Predicted protein structure is rotating in a counter-clockwise manner.


a)
b)


c)


d)

4.0 Discussion

4.1 Transcriptional Regulation of the afpA Gene

In the past decade, the *afpA* gene sequence and its surrounding nucleotides were sequenced using inverse PCR (Muryoi et al. 2004). Both a Shine Dalgarno sequence and a transcriptional start site were predicted in the region upstream of the *afpA* coding sequence. Upon further analysis, an RpoD promoter was predicted to be approximately 40 nucleotides upstream of the putative transcriptional start site (Figure 6). Both the -35 and -10 promoter elements of this RpoD promoter deviate from the canonical *E. coli* elements TTGACA and TATAAT, respectively. The -35 element in *P. putida* GR12-2 differs by four nucleotides (TTACGA), while the -10 element differs by one nucleotide (TATAAG). Interestingly, the *afpA* promoter is closer to the canonical promoter than the cold regulated *cspA* promoter in *E. coli*, TTGCAT and CTTAAT. Furthermore, unlike *cspA*, an A-T rich UP element was not found directly upstream of the *afpA* -35 box (Mitta et al. 1997; Tanabe et al. 1992). In *E. coli*, the non-canonical *cspA* promoter sequence combined with an upstream UP element creates an strong cold-regulated promoter; at 37°C, the promoter has moderate strength, but when temperature decreases, this promoter becomes exceptionally strong (Tanabe et al. 1992). The lack of an enhancing UP element and a near canonical promoter sequence suggests that *afpA* is not transcriptionally regulated like *cspA*. In fact, based on the *afpA* promoter features alone, transcription of *afpA* is likely active during the exponential growth phase because RpoD is the sigma factor associated with transcription during vegetative cell growth (Saleh and Glick 2001).

Without an identified enhancer or repressor, *afpA* does not appear to be cold inducible at the transcriptional level.

4.1.1 Quantification of afpA Promoter Strength

Using the construct, pQKAfp431, the strength of the *afpA* promoter was indirectly assessed in both host bacteria *P. putida* GR12-2 and *P.* sp. UW4. Under rich growth conditions at 30°C and 24°C, the presence of luciferase expression indicates that transcription occurs at the exponential growth phase under optimal conditions (Figure 12). However, active transcription at these high temperatures contradicts previous evidence of selective AfpA expression at low temperature (Xu et al. 1998). On the other hand, this activity confirms the presence of an active promoter within the promoter region upstream of the *afpA* coding sequence. At 4°C, both tested bacterial strains still had promoter activity but this activity was weaker compared to higher temperatures conditions. Tentatively, this suggests that at a lower temperature the *afpA* promoter becomes less active. At lower temperatures bacterial cells often become stressed. One possibility is that under stress conditions, the RNA polymerase uses the stress associated sigma factor RpoS for promoter recognition, and this leads to an upregulation of stress-related genes with RpoS based promoters and down-regulation of other genes unrelated to stress (Saleh and Glick 2001). Down-regulation of the *afpA* promoter during this cold, stressful condition supports the RpoD promoter prediction and not a stress based promoter. Interestingly, at all temperature conditions, the promoter activity that is observed is around 10-fold lower than the positive control, a putative RpoD promoter for

the *iaaM* gene *P.* sp. UW4 (data not shown). The substantially weaker *afpA* promoter activity suggests that there may be other factors affecting the transcription of this gene.

In an attempt to isolate the predicted RpoD *afpA* promoter and eliminate upstream repressors, the construct pQKAfp131, consisting of a shortened promoter region as the promoter insert, was tested with the luciferase system in both bacteria. Across all conditions, this construct has approximately half the activity of the full promoter region construct. At 4°C, no promoter activity was detected. Since luciferase activity with this construct is consistently lower in all conditions, the lost signal may be attributed to a diminished promoter activity that is not differentiable from background signals rather than a complete loss of activity. A one-way ANOVA on both promoter construct activities at different temperatures suggests that there is a marginally significant (p-value = 0.052) interaction between temperature and promoter construct. This would suggest the difference in temperature activity might depend on the promoter construct. One possible reason for this weaker activity is that an important component of promoter recognition was removed. This component may be an enhancer sequence or may be a part of the *afpA* promoter. Regardless, there appears to be a critical cis-element upstream of the predicted RpoD promoter, which enhances transcription in all conditions.

Bioinformatic strategies were employed to predict the presence of potential transcriptional regulators upstream of the *afpA* gene (Table 4). From these predictions, *afpA* is suggested to be involved in a cold shock response by CspA, cytidine catabolism by CytR, nitrate response by NarL, osmotic response by OmpR or quorum sensing by RhlR (Benkert et al. 2008; Cai and Inouye 2002; Haugo and Watnick 2002; Medina et al. 2003). Aside from the transcriptional factors OmpR and CspA, the other putative

transcriptional factors did not affect promoter activity levels when the system was tested by providing it with exogenous material (Figure 16). This suggests that these putative transcriptional factors were likely false positives and are not involved in the regulation of the *afpA* gene. It is also possible, but not especially likely, that the bacterial test strains are unable to uptake the exogenously provided molecules into the cells.

Osmoregulation was tested with a high salt concentration in the growth medium. Under this condition, the construct pQKAfp431 (full construct) was tested solely in *P.* sp. UW4; *P. putida* GR12-2 was not tested due to contamination problems. The results indicate that under high salt there is a decrease in promoter activity (Figure 17). Similar to a low temperature condition, *P.* sp. UW4 is likely highly stressed under high salt. Both bacteria may transition to a stress based physiology with higher expression of RpoS related genes (Saleh and Glick 2001). In this case, a decrease in RpoD promoter activity is not surprising. This possibility suggests that *afpA* regulation may be tied to growth stage rather than osmoregulation by the transcription factor OmpR.

Test results with a variety of carbon and nitrogen growth sources present a very similar case for both host bacteria (Figure 14 and Figure 15). For both *P. putida* GR12-2 and *P.* sp. UW4, either different carbon or nitrogen sources can minimally affect the promoter strength where some growth media can induce a slightly higher activity than others. These effects are not unusual. Changes in growth media are known to affect ice crystal-controlling protein activity and expression (Chen et al. 2003). For example, in one case, *Pseudomonas putida* KT2440 grown in a glycerol medium triggered upregulation of *cspA* transcription (Kim et al. 2013).

One unusual aspect in this study is that regardless of the tested conditions, *afpA* promoter activity was never upregulated. The measured promoter activity of other conditions did not exceed the measured activity when grown in the rich medium TSB. Different carbon and nitrogen sources solely downregulated promoter activity. These activity levels once again suggest that promoter strength differences may be tied to bacterial growth rather than direct phenotype changes triggered by different growth conditions. Optimal growth with preferred carbon and nitrogen sources lead to activity levels comparable to a rich growth medium. When both bacteria strains were grown with less optimal growth medium, they showed lowered promoter activity levels. Unsurprisingly, when grown with a nitrogen free minimal medium, *P.* sp. UW4 had the lowest promoter activity. These non-optimal conditions can trigger stress physiology with increased expression of RpoS (Saleh and Glick 2001). Once again, this change to stress-associated RpoS would upregulate stress genes and down-regulate RpoD based genes. These decreases in promoter activity further support an RpoD promoter for the *afpA* gene.

From the bioinformatic analysis, one of the highest scoring predicted transcriptional factors is CspA, a cold shock protein. Cold shock proteins are typically expressed during adaptation to a decrease in temperature (Fang et al. 1998; Yamanaka et al. 1999). These chaperone proteins stabilize protein folding at low temperatures. To help the cells adapt to the cold, CspA expression can upregulate the transcription of other cold-related proteins (Jiang et al. 1996). In *E. coli*, CspA is expressed primarily during the cold acclimation phase following low temperature exposure, reaching maximum expression after 45 to 75 minutes (Fang et al. 1998; Tanabe et al. 1992). In this study,

CspA upregulation was tested by exposing the bacterial strains to 4°C for three hours after they had attained an OD around 0.4. This condition also correlates with an entry into the stationary growth phase by stressing the cells with a decrease in temperature.

Unlike *P. putida* GR12-2, the other measured conditions in *P.* sp. UW4 showed a continuous decrease in activity after exposure to 4°C (Figure 13). Comparing the full promoter constructs, *P. putida* GR12-2 was able to maintain stable promoter activity during cold acclimation while *P.* sp. UW4 had decreased activity. This difference between bacterial strains indicates a difference in promoter activation between the two bacteria. Based on the previous luciferase tests, entry into the stationary phase should lead to a decrease in RpoD promoter activity as seen in *P.* sp. UW4. The stability in promoter activity for *P. putida* GR12-2 suggests a different mechanism of promoter activation. One possibility is that the CspA equivalent in *P. putida* GR12-2 is able to recognize the predicted cis-element sequence unlike the *P.* sp. UW4 equivalent (Table 4). Recognition by CspA appears to upregulate *afpA* transcription, however, decreased RpoD promoter activation during the stationary phase can negate this upregulation leading towards a stable promoter activity similar to activity levels prior to exposure. A decrease in the partial promoter constructs was to be expected since the predicted CspA element was at position 174 from the 5' end of the full promoter insert. With the element absent from the partial promoter region, CspA from both bacterial strains should be unable to recognize the sequence to upregulate *afpA* transcription. Unexpectedly, this was not the case for the partial promoter construct in GR12-2 where promoter activity remained stable. The argument against CspA regulation is that this cold acclimation effect is a species-specific effect that does not involve CspA. In this case, promoter activity of either

construct in UW4 would decrease while remaining stable in GR12-2 regardless of whether this CspA regulatory element is present. Overall, the cold acclimation test results tentatively suggest that the *afpA* gene uses an RpoD promoter that decreases during cold acclimation. However, this promoter activity can be potentially recovered by a species dependent regulation mechanism that is not involved with the predicted CspA regulatory element. This upregulation mechanisms is not present in UW4.

The results of all of the luciferase assays suggest that *afpA* is transcriptionally regulated using an RpoD promoter. However, *afpA* transcription also appears to be regulated during cold acclimation depending on the host bacteria. The mode of this regulation has yet to be identified. Whether this transcriptional regulation is due to a cold shock protein based regulator, which is solely recognized in *P. putida* GR12-2, is debatable. Detectable transcription across all conditions with an RpoD promoter contradicts the temperature dependent expression of AfpA at 4°C, especially when the highest expression level was documented to occur during the onset of the stationary growth phase (Xu et al. 1998). Selective expression of AfpA combined with non-selective transcription implies that *afpA* is temperature regulated at an alternate level.

4.2 Posttranscriptional Regulation of the afpA RNA Transcript

The *afpA* gene may also be regulated at the post transcriptional level (Merino and Yanofsky 2005; Yanofsky 1981). In one possible model, *P. putida* GR12-2 expresses a low amount of *afpA* RNA transcripts during optimal growth in preparation for a downshift in temperature. Despite transcription, regulation at the posttranscriptional level can eliminate translation of AfpA until the required temperature conditions are met

(Kortmann and Narberhaus 2012). During a decrease in temperature, the bacterium can then quickly translate the available mRNA to produce AfpA. The ability to rapidly express AfpA in response to changes in temperature can be important during periods of freeze-thaw stress when temperature can rapidly fluctuate (Walker et al. 2006; Wilson et al. 2006). In this case, low but sudden expression of AfpA is sufficient for ice recrystallization inhibition activity, the antifreeze activity that protects against freeze thaw stress (Yu et al. 2010). With this model, the transcriptional and posttranscriptional regulation can provide rapid resistance to freezing stress.

Posttranscriptional regulation depends on translation efficiency. Since alternate modes of translation initiation such as a downstream box were not identified, translation of *afpA* RNA transcripts would be regulated by the accessibility of the ribosome binding site (Giuliodori et al. 2010; Kortmann and Narberhaus 2012). The pQF70K constructs cannot empirically test this regulation model since these constructs contain a native ribosome binding site (Farinha and Kropinski 1990). The luciferase activity would then not be reflective of the translation efficiency of mRNA. To investigate this regulation model, the putative conformations of the *afpA* RNA transcript under various temperature were analyzed (Figure 18 and Figure 19). In all predictions, a putative translation attenuator structure was found upstream of the Shine Dalgarno sequence (Kortmann and Narberhaus 2012). The close proximity of this attenuator suggests that it could readily affect translational efficiency and is a form of posttranscriptional regulation. Since it was predicted to be an attenuator, one would infer that it inhibits translation either through restricting Shine Dalgarno sequence recognition or ribosome binding efficiency. However, the consistent presence of this hairpin loop in all temperature predictions

indicates that it is not temperature dependent suggesting that it may not significantly alter translation efficiency based on temperature. In fact, this putative attenuator can be an equivalent of the upstream box from CspA (Yamanaka et al. 1999). The predicted attenuator is a 16 bp long palindromic sequence located seven nucleotides upstream of the *afpA* Shine Dalgarno sequence. Similarly, the CspA upstream box is a 13 bp palindromic sequence located 11 nucleotides upstream of the *cspA* Shine Dalgarno sequence (Yamanaka et al. 1999). With CspA, the upstream box is complementary to the 16S rRNA in *E. coli* and assists in ribosome binding for greater translation efficiency. In the same manner, the putative attenuator from *P. putida* GR12-2 may affect translation efficiency by facilitating the formation of the downstream ribosome binding hairpin loop (Yamanaka et al. 1999). Formation of this structure places the Shine Dalgarno sequence near the loop end of the hairpin loop allowing recognition and binding to ribosomes (Figure 19c). The similarity between the putative attenuator and the CspA upstream box suggests that the former sequence is a putative upstream box rather than a translational attenuator.

Unlike other cold and posttranscriptionally regulated mRNA, the *afpA* RNA transcript does not have a long 5' UTR (Derzelle et al. 2002; Jiang et al. 1996; Singh et al. 2009). However, this short 5' UTR can still form different secondary structures by interacting with the first 50 nucleotides of the *afpA* coding sequence (Giuliodori et al. 2010). The program MFOLD was able to predict three different conformations at different temperatures depending on the amount of interaction with the coding sequence. Without sufficient coding sequence, 5' UTR folds into the same structure at all temperature conditions (Figure 18). However, the different temperature conformations for

the full RNA transcript support the model that AfpA expression is regulated at the posttranscriptional level (Figure 19). At 37°C, the 5' UTR folds into an elaborate structure where the Shine Dalgarno sequence is paired with a part of the coding sequence (Figure 19a). Not only does this conformation physically reduce ribosome accessibility to the Shine Dalgarno sequence, the self pairing structure also prevents the ribosome from binding the sequence (Yamanaka et al. 1999). At 4°C, the 5' UTR does not interact with the coding sequence (Figure 19c). This less elaborate secondary structure situates the non-bonded Shine Dalgarno sequence on the loop portion of hairpin loop. Such a conformation permits both ribosome accessibility and binding with the sequence (Giuliodori et al. 2010). Translation would then be initiated and AfpA would be expressed at 4°C and repressed at 37°C (Xu et al. 1998).

One problem with this regulation model is that at 24°C, *P. putida* GR12-2's optimal growth temperature, the 5' UTR structure folds in two main conformations. For half the predictions, the Shine Dalgarno sequence is accessible to the ribosome while the other half has an inaccessible conformation. The RNA structure program supports the ribosome accessible conformation. However, empirical evidence where AfpA was not expressed at 24°C supports a ribosome inaccessible conformation (Sun et al. 1995; Xu et al. 1998). The lack of a committed structure indicates structure flexibility at this temperature. This flexibility may even suggest potential AfpA expression at 24°C under the appropriate conditions. The final conformation at 24°C is possibly affected by cotranscriptional folding (Giuliodori et al. 2010). Cotranscriptional folding is when RNA folding is affected by the condition and speed of transcription: the RNA is folded while being transcribed. For CspA, at low temperatures a pseudoknot, permitting ribosome

accessibility, is formed which becomes destabilized during an increase in temperature (Giuliodori et al. 2010). At high temperature, this pseudoknot does not form at all. The *afpA* RNA transcript may follow the same pattern. At low temperatures, RNA structures can become very stable. Thus, when transcribed at low temperature, the initial 5' UTR secondary structures (Figure 19c) formed are stable creating a ribosome accessible form. Whereas, at higher temperatures such as 24°C, these initial structures may be less stable allowing flexible interaction with the coding sequence, forming the ribosome inaccessible structure.

However, cotranscriptional folding does not support the quick AfpA expression model. With cotranscriptional folding, a translation permissive RNA transcript can only be folded at 4°C. Previously transcribed *afpA* RNA transcripts at higher temperatures cannot be translated even after a decrease in temperature. This eliminates the possibility of rapid AfpA expression during a downshift in temperature. On the other hand, transcription appears stable after a decrease in temperature during cold acclimation (Figure 13b). This continual transcription could provide sufficient mRNA at 4°C for AfpA expression, enough to provide ice recrystallization inhibition activity for freezing stress protection (Yu et al. 2010).

An alternate possibility for selective AfpA expression, despite a ribosome accessible conformation at both 24°C and 4°C, is RNA stability (Fang et al. 1998). Although the 5' UTR conformations are similar at both temperatures, some of the folding of the remaining mRNA sections is different. This difference can contribute to an increase or decrease in RNA stability. Even if translation were to occur, an unstable mRNA leads to RNA degradation and ultimately prevents expression (Fang et al. 1998).

Overall, temperature can affect conformation and stability of the *afpA* RNA transcript. These effects suggest that *afpA* can be regulated at the posttranscriptional level.

4.3 Polycistronic Expression of Putative afpA Operon

Besides posttranscriptional regulation, AfpA expression may be regulated as an operon. The arrangement of a hemolysin calcium binding protein (HCBP) gene followed by at least one glycosyltransferase gene in all strains of the phylogeny tree (except *P. putida* SJ3) strongly suggests that these HCBPs are likely regulated as an operon (Figure 23). In some of these strains, a conserved hypothetical protein directly upstream of HCBP genes complicates the situation. Whether expression of this putative operon begins at the conserved hypothetical protein gene or the HCBP gene is uncertain. The conserved nature of this hypothetical protein gene, with similar spacing upstream of the HCBP genes, implies that the operon begins at the hypothetical gene. In fact, this hypothetical protein was only found upstream of a HCBP gene and in *Pseudomonas* strains. This unique arrangement supports the possibility that the operon begins at the hypothetical protein.

However, this hypothetical protein was not found in most strains of the *P. syringae* clade (Figure 23). A weak promoter was even found upstream of the *afpA* gene. This promoter is embedded into the C-terminal end of the conserved hypothetical protein's coding sequence. Within a compact bacterial genome this arrangement is not unusual. The presence of the conserved hypothetical protein may be coincidental and unrelated to the putative operon. Consequently, HCBP genes can be regulated by its own promoters suggesting that expression begins at the HCBP gene in the operon.

Regardless of first gene in the operon, this putative operon may use multiple promoters to regulate polycistronic expression (Yang and Larson 1998). Multiple promoters would allow for differential expression of all three operon genes. A promoter upstream of the conserved hypothetical protein gene would regulate the transcription of the full polycistronic mRNA containing all genes while internal promoters can fine tune the expression of downstream genes or regulate differential expression under various conditions. In *E. coli*, the *glpEGR* genes of the *glp* operon are regulated with multiple promoters (Yang and Larson 1998). These genes can be cotranscribed together as a polycistronic mRNA by the *glpE* promoter, but a moderately strong promoter was also found upstream of *glpG* with two weak promoters upstream of *glpR*. The weak strength of the tested *afpA* promoter suggests that it may be an internal promoter (Figure 33). The predicted glycosyltransferase promoter may also be another internal promoter. This promoter can be posttranscriptionally regulated by the transcriptional terminator and anti-terminator region of the *afpA* gene (Merino and Yanofsky 2005; Yanofsky 1981). Under correct conditions, the terminator region folds into a conformation where the glycosyltransferase promoter is accessible for further transcription (Figure 7). In this permissible condition, the polycistronic RNA transcript produced has both the *afpA* and glycosyltransferase coding sequence for translation. A multiple promoter model for *afpA* in *P. putida* GR12-2 includes a main promoter upstream of the putative hypothetical protein gene with two internal promoters upstream of the *afpA* and glycosyltransferase genes respectively.

Figure 33. Potential multiple promoter model of the putative *afpA* operon in *P. putida* GR12-2. Each promoter is represented in purple and regulates the respective downstream gene. Hypo protein – hypothetical protein; MP – main promoter; IP1 – internal promoter one; IP2 – internal promoter two.

4.3.1 Conserved Hypothetical Protein in the afpA Operon

An important aspect of the multiple promoter model is the conserved hypothetical gene directly upstream of *afpA*. One large assumption is that this gene is regulated or involved with *afpA*, but a lack of empirical data makes this relation very uncertain. Furthermore, there is insufficient available DNA sequence upstream of *afpA* to determine if there is a putative gene. However, one interesting aspect is that this gene appears in ten pseudomonads and belongs mostly to the other *Pseudomonas* cluster on the phylogeny tree (Figure 23). From this tree, the lack of this gene in the other clade suggests that this gene and its promoter are lost in *P. syringae*. In *P. syringae*, the operon may begin at HCBP gene whereas the operon begins at the hypothetical protein in other *Pseudomonas* strains. *P. syringae* are largely isolated from plants whereas the other pseudomonads usually originate from soil and wastewater samples suggesting that the environment may play a role in this gene loss (Feil et al. 2005; Li et al. 2012b; Nelson et al. 2002).

4.3.2 Putative Glycosyltransferase in afpA Operon

In nearly all pseudomonads investigated, one or two glycosyltransferase genes were found downstream of the respective HCBP. The consistency of the presence of one or more glycosyltransferase genes in all clades suggests it is involved with the HCBP.

Since in the putative *afpA* operon this glycosyltransferase gene is likely regulated along with *afpA*, one may speculate that the glycosyltransferase is involved in posttranslational glycosylation of the HCBP (El-Battari et al. 2003). However, this may not be the case for AfpA. The downstream glycosyltransferase in the *afpA* operon is only a partial protein making its function unclear. The truncation of this protein could have resulted in a loss of glycosylation activity.

Based on Blast results, the glycosyltransferase in *P. putida* GR12-2 is a family 2 glycosyltransferase using the sequence-based classification system. The returned hits in other *Pseudomonas* strains either belonged to family 2 or were unclassified. This classification provides no indication of a putative function for *P. putida* GR12-2's glycosyltransferase. In fact, this protein family is polyspecific in activity and includes proteins such as cellulose synthase, chitin synthase and more (Breton et al. 2006; Coutinho et al. 2003). Furthermore, the partial glycosyltransferase protein lacks a characteristic GT-A fold in family 2 glycosyltransferase proteins due to a premature stop codon. The middle section of *P. putida* GR12-2's glycosyltransferase aligns poorly with the remaining glycosyltransferases and may reflect a large mutation event, perhaps a recombination event since there were no large gaps during alignment (Figure 22). This difference likely affects the folding of the partial protein and ultimately its function. Based on the alignment and sequence identity, the partial protein is a putative unclassified glycosyltransferase. Until an active site can be identified, whether the protein has glycosylation function is uncertain.

4.4 Further Characterization of AfpA

The native AfpA protein was re-isolated from an old stock of *P. putida* GR12-2. The isolated AfpA protein was slightly different than previous studies. The thermal hysteresis activity of a 1 mg/ml solution of AfpA was confirmed to be 0.05°C, about half the previously reported value of 0.13°C (Sun et al. 1995; Xu et al. 1998). The pI value reported by the company Advanced Electrophoresis Solutions Ltd was lower at around 4.0 to 5.0 instead of 5.6 (Xu et al. 1998). These differences potentially are a result of an altered isolation method, i.e., less active AfpA was isolated. However, sufficient amounts of AfpA were isolated for further characterization of this antifreeze protein.

In this study, AfpA has the second antifreeze activity, ice recrystallization inhibition activity. Since the ice recrystallization inhibition assay was a qualitative assay, the degree and strength of this activity could not be compared to other antifreeze proteins. Using size exclusion chromatography, the ~164 kDa antifreeze protein (previously determined on the basis of SDS-PAGE) was determined to be greater than 500 kDa in size (Figure 27). This large size is not unusual for a bacterial antifreeze protein (Garnham et al. 2008). However, unlike *Mp*AFP, this size may be due to aggregation of AfpA. The antifreeze protein does not appear to aggregate into a predetermined multimeric structure since AfpA was observed to aggregate into a white precipitate after several months of storage. When the fractions beyond the elution peak (trailing tail) were tested, these fractions were found to contain the antifreeze protein. AfpA in these fractions indicate that there are AfpA aggregates smaller than 500 kDa. These smaller sizes support the idea that AfpA is slowly and continuously aggregating, producing proteins of many different sizes. The slow AfpA aggregation is likely due to the interaction between hydrophobic pockets found in the protein. This aggregation is potentially an important characteristic of

AfpA. It can influence protein function allowing AfpA to have both ice nucleation and antifreeze activity (Du et al. 2006). Aggregated AfpA would express an ice nucleation activity while monomeric or smaller aggregates of AfpA would express antifreeze activities. Since partially purified sample appears to have both AfpA aggregates and monomers, aggregation is a plausible explanation for this antifreeze protein to have both activities.

4.4.1 Putative Ice Binding Surface

Based on the amino acid sequence, AfpA was modeled into several structures. Most of these structures were composed of a combination of α-helices and β-sheets. When protein surfaces of these structures were examined, there were few flat surfaces that could potentially contain an ice binding site (Kawahara 2008). However, there was one structure of interest, a β-solenoid structure similar to other antifreeze and ice nucleation proteins (Garnham et al. 2008; Garnham et al. 2011b; Middleton et al. 2012). This structure was considered a putative AfpA structure because well characterized bacterial antifreeze and ice nucleation proteins, with RTX toxin-like features, were modeled as β-solenoid structures (Garnham et al. 2008; Garnham et al. 2011b; Graether and Jia 2001). One problem with this model is that it relies on tandem amino acid repeats in the protein sequence. These repeats are align to create an organized ice binding surface on one side of the protein. This is problematic since AfpA lacks distinctive sequence repeats besides the calcium binding motif, an RTX toxin-like feature. However, some irregularities in repetition are tolerable

Although the predicted β-solenoid structure used *Mp*AFP as a template, one essential difference between the structures is that the calcium binding motifs in AfpA were not aligned along one side of the protein (Figure 32). In fact, with around four amino acid residues separating these putative motifs, they were clustered near the N-terminus of the predicted partial protein structure. The lack of the motif alignment suggests that these motifs are not directly involved in ice binding unlike *Mp*AFP (Garnham et al. 2011a). Experimentally, calcium ions were confirmed to not affect AfpA activity (Figure 29) (Kawahara et al. 2007). This evidence supports that these motifs are not involved in ice binding affecting antifreeze activity (Garnham et al. 2008).

A putative ice binding surface was predicted in the AfpA protein structure (Doxey et al. 2006). Encompassing a small surface area, this ice binding surface is located on folding turns and is near the N-terminus of the protein structure (Figure 32). This putative ice binding site composes a small percentage of the structure surface, not including the non modeled portions of the protein. Compared to the *Mp*AFP antifreeze domain with an ice binding site spanning the length of the entire structure, this small surface area may explain the low protein activity of AfpA (Figure 4). However, posttranslational modifications needs to be considered and can affect antifreeze activity. Four N-glycosylation sites were predicted within this putative ice binding region of the protein and can likely interfere with the antifreeze activity (Muryoi et al. 2004). In some cases, glycans can form the ice binding site when aligned on one side of the protein (Bouvet and Ben 2003). If glycosylation and lipidation are considered, the putative ice binding surface may be a false prediction.

4.4.2 AfpA Secretion

In a previous study AfpA was speculated to be secreted using either a hemolysin-like secretion system or a type V autotransporter (Muryoi et al. 2004). One putative hemolysin based secretion system for AfpA is the type I secretion system. This secretion system is known to secrete RTX toxins and speculated to secrete *Mp*AFP (Guo et al. 2013; Lecher et al. 2012). Identified as RTX-like features by Blast, the calcium binding motifs in AfpA may be involved in secretion using the type I system. This system does not have a specific conserved signal sequence however, three weakly identified signal subtypes were found near the C-terminus of proteins (Holland et al. 2005). These signals are not removed during transportation. Some of these type I transported proteins also have calcium binding repeats upstream of the C-terminal signal sequence. These repeat sequences tend to be glycine and asparate rich. For the RTX protein, HlyA in *E. coli*, the consensus sequence for these repeats is Gly-Gly-X-Gly-X-Asp-X-U-X, where U represents a hydrophobic residue (Lecher et al. 2012). Interestingly, the putative calcium binding motifs in AfpA have a general repeating consensus of [Ala/Thr]-Gly-X-Gly-X-[Asp/Asn]-X-U-X (Figure 20). Furthermore, AfpA has four motifs of unknown function near the C-terminus of the protein (Muryoi et al. 2004). Both the calcium binding motifs and motifs with unknown function are well-conserved among the HCBP found in other *Pseudomonas* strains (Figure 20). When combined these two features support a type I secretion system for AfpA where the motif near the C-terminus acts as the signal while the calcium binding motif assists in the secretion.

One difficulty with the idea of type I secretion of AfpA is the posttranslational modification and folding of the protein. Although there are some cases of eukaryotic glycosyltransferases secreted outside the cell, most documented cases of bacterial

glycosyltransferases operate in the periplasm (Szymanski and Wren 2005). This protein localization suggests that AfpA may need to be folded and modified prior to secretion. Although it has been suggested that this type I system can transport posttranslationally modified proteins, the similarity between AfpA and HlyA features support secretion as an unfolded form using calcium to facilitate protein folding (Holland et al. 2005). Furthermore, *Mp*AFP is speculated to be secreted in a similar manner. Another problem with this system is the use of calcium for folding. TSB medium lacks known free calcium ions, but when *P. putida* GR12-2 was grown in this medium it still expressed and secreted AfpA. This difficulty is less problematic in arctic soil since there is a higher concentration of calcium in areas rich in plant species (Van der Welle et al. 2003). As a plant growth promoting bacterium, it is likely for *P. putida* GR12-2 to thrive in these soils.

5.0 Conclusion

To combat freezing stress, many bacteria have been documented to express ice crystal controlling proteins (Lorv et al. 2014). At 5°C, the plant growth promoting bacterium *P. putida* GR12-2 expresses AfpA, an ice crystal controlling protein (Sun et al. 1995). AfpA has both antifreeze activities, thermal hysteresis and ice recrystallization inhibition activity. To better understand the mechanism(s) regulating these activities in AfpA, the protein was further analyzed. Based on computational analysis, a small putative ice binding site was identified on one side the protein surface and situated on the β-sheet turns. Although this site may be used for ice binding, aggregation of AfpA, by hydrophilic and hydrophobic pockets, likely determines the protein's ice crystal controlling activities. For these antifreeze activities to benefit the bacterium, AfpA is potentially secreted using a type I secretion system.

Since the protein is beneficial at low temperatures, the *afpA* regulation system was investigated to determine how AfpA expression is limited to these low temperature conditions. Using a luciferase reporter assay, *afpA* was determined to be transcribed using a weak RpoD promoter that is active under vegetative cell growth. Under cell stress conditions such as entry into the stationary growth phase, the promoter is less active. However, this stress based decrease in transcription can be briefly negated by upregulation with CspA during the cold acclimation phase. One potential explanation for the weakness of the *afpA* RpoD promoter is that it is an internal promoter within an operon consisting of an upstream conserved hypothetical protein, *afpA* and a partial glycosyltransferase gene. Beyond transcription, AfpA can be further regulated at the posttranscriptional level. The *afpA* RNA transcript is capable of folding into either a

ribosome binding site accessible or inaccessible structure based on the interaction of the 5’ UTR with the coding sequence. From these results, *afpA* is likely regulated at multiple different levels to restrict expression at higher temperatures.

Reference List

Abreu-Goodger, C., Ontiveros-Palacios, N., Ciria, R., and Merino, E. 2004. Conserved regulatory motifs in bacteria: riboswitches and beyond. Trends Genet. **20**(10): 475-479.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**(3): 403-410.

Benkert, B., Quack, N., Schreiber, K., Jaensch, L., Jahn, D., and Schobert, M. 2008. Nitrate-responsive NarX-NarL represses arginine-mediated induction of the *Pseudomonas aeruginosa* arginine fermentation *arcDABC* operon. Microbiology **154**(Pt 10): 3053-3060.

Bouvet, V., and Ben, R.N. 2003. Antifreeze glycoproteins - Structure, conformation, and biological applications. Cell Biochem. Biophys. **39**(2): 133-144.

Breton, C., Šnajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. 2006. Structures and mechanisms of glycosyltransferases. Glycobiology **16**(2): 29R-37R.

Cai, S.J., and Inouye, M. 2002. EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. J. Biol. Chem. **277**(27): 24155-24161.

Cambours, M.A., Nejad, P., Granhall, U., and Ramstedt, M. 2005. Frost-related dieback of willows. Comparison of epiphytically and endophytically isolated bacteria from different *Salix* clones, with emphasis on ice nucleation activity, pathogenic properties and seasonal variation. Biomass Bioenergy **28**(1): 15-27.

Cao, S., and Chen, S. 2006. Predicting RNA pseudoknot folding thermodynamics. Nucleic Acids Res. **34**(9): 2634-2652.

Celik, Y., Graham, L.A., Mok, Y., Bar, M., Davies, P.L., and Braslavsky, I. 2010. Superheating of ice crystals in antifreeze protein solutions. Proc. Natl. Acad. Sci. U. S. A. **107**(12): 5423-5428.

Chen, M.L., Chiou, T.K., Tsao, C.Y., and Jiang, S.T. 2003. Enhancement of the expression of ice-nucleation activity of *Pseudomonas fluorescens* MACK-4 isolated from mackerel. Fisheries Science **69**(1): 195-203.

Cheng, Z., Park, E., and Glick, B.R. 2007. 1-Aminocyclopropane-1-carboxylate deaminase from *Pseudomonas putida* UW4 facilitates the growth of canola in the presence of salt. Can. J. Microbiol. **53**(7): 912-918.

Choi, K.H., Kumar, A., and Schweizer, H.P. 2006. A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. J. Microbiol. Methods **64**(3): 391-397.

Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. 2003. An evolving hierarchical family classification for glycosyltransferases. J. Mol. Biol. **328**(2): 307-317.

Derzelle, S., Hallet, B., Ferain, T., Delcour, J., and Hols, P. 2002. Cold shock induction of the *cspL* gene in *Lactobacillus plantarum* involves transcriptional regulation. J. Bacteriol. **184**(19): 5518-5523.

Doucet, D., Tyshenko, M.G., Kuiper, M.J., Graether, S.P., Sykes, B.D., Daugulis, A.J., Davies, P.L., and Walker, V.K. 2000. Structure-function relationships in spruce budworm antifreeze protein revealed by isoform diversity. Eur. J. Biochem. **267**(19): 6082-6088.

Doxey, A.C., Yaish, M.W., Griffith, M., and McConkey, B.J. 2006. Ordered surface carbons distinguish antifreeze proteins and their ice-binding regions. Nat. Biotech. **24**(7): 852-855.

Du, N., Liu, X.Y., and Hew, C.L. 2006. Aggregation of antifreeze protein and impact on antifreeze activity. J. Phys. Chem. B **110**(41): 20562-20567.

Duan, J., Jiang, W., Cheng, Z., Heikkila, J.J., and Glick, B.R. 2013. The complete genome sequence of the plant growth-promoting bacterium *Pseudomonas* sp. UW4. PLoS ONE **8**(3): e58640.

Duca, D. 2013. Characterizing the indole-3-acetic acid biosynthesis pathways in *Pseudomonas Putida* UW4. Master of Science thesis, University of Waterloo, Waterloo.

Duman, J.G., and Olsen, T.M. 1993. Thermal hysteresis protein activity in bacteria, fungi, and phylogenetically diverse plants. Cryobiology **30**(3): 322-328.

El-Battari, A., Prorok, M., Angata, K., Mathieu, S., Zerfaoui, M., Ong, E., Suzuki, M., Lombardo, D., and Fukuda, M. 2003. Different glycosyltransferases are differentially processed for secretion, dimerization, and autoglycosylation. Glycobiology **13**(12): 941-953.

Etchegaray, J.P., Jones, P.G., and Inouye, M. 1996. Differential thermoregulation of two highly homologous cold-shock genes, *cspA* and *cspB*, of *Escherichia coli*. Genes Cells **1**(2): 171-178.

Fang, L., Hou, Y., and Inouye, M. 1998. Role of the cold-box region in the 5' untranslated region of the *cspA* mRNA in its transient expression at low temperature in *Escherichia coli*. J. Bacteriol. **180**(1): 90-95.

Fang, L., Xia, B., and Inouye, M. 1999. Transcription of *cspA*, the gene for the major cold-shock protein of *Escherichia coli*, is negatively regulated at 37°C by the 5'-untranslated region of its mRNA. FEMS Microbiol. Lett. **176**(1): 39-43.

Farinha, M.A., and Kropinski, A.M. 1990. Construction of broad-host-range plasmid vectors for easy visible selection and analysis of promoters. J. Bacteriol. **172**(6): 3496-3499.

Feil, H., Feil, W.S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., Thiel, J., Malfatti, S., Loper, J.E., Lapidus, A., Detter, J.C., Land, M., Richardson, P.M., Kyrpides, N.C., Ivanova, N., and Lindow, S.E. 2005. Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U. S. A. **102**(31): 11064-11069.

Garnham, C.P., Gilbert, J.A., Hartman, C.P., Campbell, R.L., Laybourn-Parry, J., and Davies, P.L. 2008. A Ca^2+-dependent bacterial antifreeze protein domain has a novel β-helical ice-binding fold. Biochem. J. **411**: 171-180.

Garnham, C.P., Campbell, R.L., and Davies, P.L. 2011a. Anchored clathrate waters bind antifreeze proteins to ice. Proc. Natl. Acad. Sci. U. S. A. **108**(18): 7363–7367.

Garnham, C.P., Campbell, R.L., Walker, V.K., and Davies, P.L. 2011b. Novel dimeric beta-helical model of an ice nucleation protein with bridged active sites. BMC Struct. Biol. **11**: 36.

Garnham, C.P., Nishimiya, Y., Tsuda, S., and Davies, P.L. 2012. Engineering a naturally inactive isoform of type III antifreeze protein into one that can stop the growth of ice. FEBS Lett. **586**(21): 3876-3881.

Gilbert, J.A., Hill, P.J., Dodd, C.E.R., and Laybourn-Parry, J. 2004. Demonstration of antifreeze protein activity in Antarctic lake bacteria. Microbiology **150**: 171-180.

Gilbert, J.A., Davies, P.L., and Laybourn-Parry, J. 2005. A hyperactive, Ca^2+-dependent antifreeze protein in an Antarctic bacterium. FEMS Microbiol. Lett. **245**(1): 67-72.

Giuliodori, A.M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., Gualerzi, C.O., and Pon, C.L. 2010. The *cspA* mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. Mol. Cell **37**(1): 21-33.

Govindarajan, A.G., and Lindow, S.E. 1988. Phospholipid requirement for expression of ice nuclei in *Pseudomonas syringae* and *in vitro*. J. Biol. Chem. **263**(19): 9333-9338.

Graether, S.P., and Jia, Z.C. 2001. Modeling *Pseudomonas syringae* ice-nucleation protein as a beta-helical protein. Biophys. J. **80**(3): 1169–1173.

Grant, S.G., Jessee, J., Bloom, F.R., and Hanahan, D. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc. Natl. Acad. Sci. U. S. A. **87**(12): 4645-4649.

Griffith, M., and Ewart, K.V. 1995. Antifreeze proteins and their potential use in frozen foods. Biotechnol. Adv. **13**(3): 375-402.

Guo, S., Garnham, C.P., Whitney, J.C., Graham, L.A., and Davies, P.L. 2012. Re-evaluation of a bacterial antifreeze protein as an adhesin with ice-binding activity. PLoS ONE **7**(11): e48805.

Guo, S., Garnham, C.P., Karunan Partha, S., Campbell, R.L., Allingham, J.S., and Davies, P.L. 2013. Role of Ca^{2}+ in folding the tandem β-sandwich extender domains of a bacterial ice-binding adhesin. FEBS J. **280**(22): 5919-5932.

Hakim, A., Nguyen, J.B., Basu, K., Zhu, D.F., Thakral, D., Davies, P.L., Isaacs, F.J., Modis, Y., and Meng, W. 2013. Crystal structure of an insect antifreeze protein and its implications for ice binding. J. Biol. Chem. **288**(17): 12295–12304.

Haugo, A.J., and Watnick, P.I. 2002. *Vibrio cholerae* CytR is a repressor of biofilm development. Mol. Microbiol. **45**(2): 471-483.

Hirano, S.S., and Upper, C.D. 2000. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*—a pathogen, ice nucleus, and epiphyte. Microbiol. Mol. Biol. Rev. **64**(3): 624-653.

Hobbs, R.S., Shears, M.A., Graham, L.A., Davies, P.L., and Fletcher, G.L. 2011. Isolation and characterization of type I antifreeze proteins from cunner, *Tautogolabrus adspersus*, order Perciformes. FEBS J. **278**(19): 3699–3710.

Holland, I.B., Schmitt, L., and Young, J. 2005. Type I protein secretion in bacteria, the ABC-transporter dependent pathway (review). Mol. Membr. Biol. **22**(1-2): 29-39.

Jia, Z.C., and Davies, P.L. 2002. Antifreeze proteins: an unusual receptor-ligand interaction. Trends Biochem. Sci. **27**(2): 101-106.

Jiang, W., Fang, L., and Inouye, M. 1996. The role of the 5'-end untranslated region of the mRNA for CspA, the major cold-shock protein of *Escherichia coli*, in cold-shock adaptation. J. Bacteriol. **178**(16): 4919-4925.

Junge, K., and Swanson, B.D. 2008. High-resolution ice nucleation spectra of sea-ice bacteria: implications for cloud formation and life in frozen environments. Biogeosciences **5**(3): 865-873.

Kajava, A.V., and Lindow, S.E. 1993. A model of the 3-dimensional structure of ice nucleation proteins. J. Mol. Biol. **232**(3): 709-717.

Kawahara, H. 2008. Cryoprotectants and ice-binding proteins. *In* Psychrophiles: from biodiversity to biotechnology. *Edited by* R. Margesin, F. Schinner, J. Marx and C. Gerday. Springer-Verlag Berlin Heidelberg, Heidelberg, Germany. pp. 229-246.

Kawahara, H., Mano, Y., and Obata, H. 1993. Purification and characterization of extracellular ice-nucleating matter from *Erwinia uredovora* KUIN-3. Biosci. Biotechnol. Biochem. **57**(9): 1429-1432.

Kawahara, H., Li, J.P., Griffith, M., and Glick, B.R. 2001. Relationship between antifreeze protein and freezing resistance in *Pseudomonas putida* GR12-2. Curr. Microbiol. **43**(5): 365-370.

Kawahara, H., Nakano, Y., Omiya, K., Muryoi, N., Nishikawa, J., and Obata, H. 2004. Production of two types of ice crystal-controlling proteins in Antarctic bacterium. J. Biosci. Bioeng. **98**(3): 220-223.

Kawahara, H., Iwanaka, Y., Higa, S., Muryoi, N., Sato, M., Honda, M., Omura, H., and Obata, H. 2007. A novel, intracellular antifreeze protein in an Antarctic bacterium, *Flavobacterium xanthum*. Cryoletters **28**(1): 39-49.

Kelley, L.A., and Sternberg, M.J.E. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protocols **4**(3): 363-371.

Kim, J., Oliveros, J.C., Nikel, P.I., de Lorenzo, V., and Silva-Rocha, R. 2013. Transcriptomic fingerprinting of *Pseudomonas putida* under alternative physiological regimes. Environ. Microbiol. Rep. **5**(6): 883-891.

Kortmann, J., and Narberhaus, F. 2012. Bacterial RNA thermometers: molecular zippers and switches. Nat. Rev. Microbiol. **10**(4): 255-265.

Kozloff, L.M., Turner, M.A., Arellano, F., and Lute, M. 1991. Phosphatidylinositol, a phospholipid of ice-nucleating bacteria. J. Bacteriol. **173**(6): 2053-2060.

Kumaki, Y., Kawano, K., Hikichi, K., Matsumoto, T., and Matsushima, N. 2008. A circular loop of the 16-residue repeating unit in ice nucleation protein. Biochem. Biophys. Res. Commun. **371**(1): 5-9.

Lecher, J., Schwarz, C.K., Stoldt, M., Smits, S.H., Willbold, D., and Schmitt, L. 2012. An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. Structure **20**(10): 1778-1787.

Li, Q., Yan, Q., Chen, J., He, Y., Wang, J., Zhang, H., Yu, Z., and Li, L. 2012a. Molecular characterization of an ice nucleation protein variant (InaQ) from *Pseudomonas syringae* and the analysis of its transmembrane transport activity in *Escherichia coli*. Int. J. Biol. Sci. **8**(8): 1097-108.

Li, S., Zhao, H., Li, Y., Niu, S., and Cai, B. 2012b. Complete genome sequence of the naphthalene-degrading *Pseudomonas putida* strain ND6. J. Bacteriol. **194**(18): 5154-5155.

Lin, F., Davies, P.L., and Graham, L.A. 2011. The thr- and ala-rich hyperactive antifreeze protein from inchworm folds as a flat silk-like beta-helix. Biochemistry **50**(21): 4467–4478.

Lin, F., Sun, T., Fletcher, G.L., and Davies, P.L. 2012. Thermolabile antifreeze protein produced in *Escherichia coli* for structural analysis. Protein Expr. Purif. **82**(1): 75-82.

Livernois, A.M., Hnatchuk, D.J., Findlater, E.E., and Graether, S.P. 2009. Obtaining highly purified intrinsically disordered protein by boiling lysis and single step ion exchange. Anal. Biochem. **392**(1): 70-76.

Lorv, J.S.H., Rose, D.R., and Glick, B.R. 2014. Bacterial ice crystal controlling proteins. Scientifica **2014**: 20.

Maki, L.R., Galyan, E.L., Chang-Chien, M., and Caldwell, D.R. 1974. Ice nucleation induced by *Pseudomonas syringae*. Appl. Microbiol. **28**(3): 456-459.

Margesin, R., and Miteva, V. 2011. Diversity and ecology of psychrophilic microorganisms. Res. Microbiol. **162**(3): 346-361.

Medina, G., Juárez, K., Valderrama, B., and Soberón-Chávez, G. 2003. Mechanism of *Pseudomonas aeruginosa* RhlR transcriptional regulation of the *rhlAB* promoter. J. Bacteriol. **185**(20): 5976-5983.

Merino, E., and Yanofsky, C. 2005. Transcription attenuation: a highly conserved regulatory strategy used by bacteria. Trends Genet. **21**(5): 260-264.

Middleton, A.J., Marshall, C.B., Faucher, F., Bar-Dolev, M., Braslavsky, I., Campbell, R.L., Walker, V.K., and Davies, P.L. 2012. Antifreeze protein from freeze-tolerant grass has a beta-roll fold with an irregularly structured ice-binding site. J. Mol. Biol. **416**(5): 713-724.

Mitta, M., Fang, L., and Inouye, M. 1997. Deletion analysis of *cspA* of *Escherichia coli*: requirement of the AT-rich UP element for *cspA* transcription and the downstream box in the coding region for its cold shock induction. Mol. Microbiol. **26**(2): 321-335.

Modig, K., Qvist, J., Marshall, C.B., Davies, P.L., and Halle, B. 2010. High water mobility on the ice-binding surface of a hyperactive antifreeze protein. Phys. Chem. Chem. Phys. **12**(35): 10189–10197.

Mok, Y., Lin, F., Graham, L.A., Celik, Y., Braslavsky, I., and Davies, P.L. 2010. Structural basis for the superior activity of the large isoform of snow flea antifreeze protein. Biochemistry **49**(11): 2593–2603.

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., and Jahn, D. 2005. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics **21**(22): 4187-4189.

Muryoi, N., Matsukawa, K., Yamade, K., Kawahara, H., and Obata, H. 2003. Purification and properties of an ice-nucleating protein from an ice-nucleating bacterium, *Pantoea ananatis* KUIN-3. J. Biosci. Bioeng. **95**(2): 157-163.

Muryoi, N., Sato, M., Kaneko, S., Kawahara, H., Obata, H., Yaish, M.W.F., Griffith, M., and Glick, B.R. 2004. Cloning and expression of *afpA*, a gene encoding an antifreeze protein from the arctic plant growth-promoting rhizobacterium *Pseudomonas putida* GR12-2. J. Bacteriol. **186**(17): 5661–5671.

Nada, H., Zepeda, S., Miura, H., and Furukawa, Y. 2010. Significant alterations in anisotropic ice growth rate induced by the ice nucleation-active bacteria *Xanthomonas campestris*. Chem. Phys. Let. **498**(1–3): 101-106.

Near, T.J., Dornburg, A., Kuhn, K.L., Eastman, J.T., Pennington, J.N., Patarnello, T., Zane, L., Fernández, D.A., and Jones, C.D. 2012. Ancient climate change, antifreeze, and the evolutionary diversification of Antarctic fishes. Proc. Natl Acad. Sci. **109**(9): 3434–3439 .

Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Dusterhoft, A., Tummler, B., and Fraser, C.M. 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ. Microbiol. **4**(12): 799-808.

Nemecek-Marshall, M., LaDuca, R., and Fall, R. 1993. High-level expression of ice nuclei in a *Pseudomonas syringae* strain is induced by nutrient limitation and low temperature. J. Bacteriol. **175**(13): 4062-4070.

Obata, H., Tanaka, T., Kawahara, H., and Tokuyama, T. 1993. Properties of cell-free ice nuclei from ice nucleation-active *Pseudomonas fluorescens* KUIN-1. J. Ferment. Bioeng. **76**(1): 19-24.

O'Brien, R.D., and Lindow, S.E. 1988. Effect of plant species and environmental conditions on ice nucleation activity of *Pseudomonas syringae* on Leaves. Appl. Environ. Microbiol. **54**(9): 2281-2286.

Pertaya, N., Marshall, C.B., DiPrinzio, C.L., Wilen, L., Thomson, E.S., Wettlaufer, J.S., Davies, P.L., and Braslavsky, I. 2007. Fluorescence microscopy evidence for quasi-

permanent attachment of antifreeze proteins to ice surfaces. Biophys. J. **92**(10): 3663-3673.

Phelps, P., Giddings, T.H., Prochoda, M., and Fall, R. 1986. Release of cell-free ice nuclei by *Erwinia herbicola*. J. Bacteriol. **167**(2): 496-502.

Ramløv, H. 2010. Measuring antifreeze activity. *In* Biochemistry and function of antifreeze proteins. *Edited by* S. Graether. Nova Science Publishers, Incorporated pp. 7-41.

Raymond, J.A., and DeVries, A.L. 1977. Adsorption inhibition as a mechanism of freezing resistance in polar fishes. Proc. Natl. Acad. Sci. U. S. A. **74**(6): 2589-2593.

Reuter, J.S., and Mathews, D.H. 2010. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics **11**: 129-2105-11-129.

Saleh, S.S., and Glick, B.R. 2001. Involvement of *gacS* and *rpoS* in enhancement of the plant growth-promoting capabilities of *Enterobacter cloacae* CAL2 and UW4. Can. J. Microbiol. **47**(8): 698-705.

Sarron, E., Cochet, N., and Gadonna-Widehem, P. 2013. Effects of aqueous ozone on *Pseudomonas syringae* viability and ice nucleating activity. Process Biochem. **48**(7): 1004-1009.

Scotter, A.J., Marshall, C.B., Graham, L.A., Gilbert, J.A., Garnham, C.P., and Davies, P.L. 2006. The basis for hyperactivity of antifreeze proteins. Cryobiology **53**(2): 229-239.

Shivaji, S., and Prakash, J.S. 2010. How do bacteria sense and respond to low temperature? Arch. Microbiol. **192**(2): 85-95.

Sidebottom, C., Buckley, S., Pudney, P., Twigg, S., Jarman, C., Holt, C., Telford, J., McArthur, A., Worrall, D., Hubbard, R., and Lillford, P. 2000. Heat-stable antifreeze protein from grass. Nature **406**(6793): 256.

Singh, A.K., Pindi, P.K., Dube, S., Sundareswaran, V.R., and Shivaji, S. 2009. Importance of *trmE* for growth of the psychrophile *Pseudomonas syringae* at low temperatures. Appl. Environ. Microbiol. **75**(13): 4419-4426.

Smolin, N., and Daggett, V. 2008. Formation of ice-like water structure on the surface of an antifreeze protein. J. Phys. Chem. B **112**(19): 6193-6202.

Solovyev, V.V., and Salamov, A. 2011. Automatic annotation of microbial genomes and metagenomic sequences. *In* Metagenomics and its applications in agriculture, biomedicine and environmental studies. *Edited by* R.W. Li. Nova Science Publishers, New York. pp. 61-78.

Studier, F.W., and Moffatt, B.A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189**(1): 113-130.

Sun, X., Griffith, M., Pasternak, J.J., and Glick, B.R. 1995. Low temperature growth, freezing survival, and production of antifreeze protein by the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2. Can. J. Microbiol. **41**(9): 776-784.

Szymanski, C.M., and Wren, B.W. 2005. Protein glycosylation in bacterial mucosal pathogens. Nat. Rev. Microbiol. **3**(3): 225-237.

Tanabe, H., Goldstein, J., Yang, M., and Inouye, M. 1992. Identification of the promoter region of the *Escherichia coli* major cold shock gene, *cspA*. J. Bacteriol. **174**(12): 3867-3873.

Turner, M.A., Arellano, F., and Kozloff, L.M. 1991. Components of ice nucleation structures of bacteria. J. Bacteriol. **173**(20): 6515-6527.

Van der Welle, M., Vermeulen, P.J., Shaver, G.R., and Berendse, F. 2003. Factors determining plant species richness in Alaskan arctic tundra. J. Veg. Sci. **14**(5): 711-720.

Vance, T.D., Olijve, L.L., Campbell, R.L., Voets, I.K., Davies, P.L., and Guo, S. 2014. Ca^2+-stabilized adhesin helps an Antarctic bacterium reach out and bind ice. Biosci. Rep. **34**(4): e00121.

Walker, V.K., Palmer, G.R., and Voordouw, G. 2006. Freeze-thaw tolerance and clues to the winter survival of a soil community. Appl. Environ. Microbiol. **72**(3): 1784-1792.

Wilson, S.L., and Walker, V.K. 2010. Selection of low-temperature resistance in bacteria and potential applications. Environ. Technol. **31**(8-9): 943–956.

Wilson, S.L., Kelley, D.L., and Walker, V.K. 2006. Ice-active characteristics of soil bacteria selected by ice-affinity. Environ. Microbiol. **8**(10): 1816–1824.

Wilson, S.L., Grogan, P., and Walker, V.K. 2012. Prospecting for ice association: characterization of freeze-thaw selected enrichment cultures from latitudinally distant soils. Can. J. Microbiol. **58**(4): 402–412.

Wu, Z., Kan, F.W.K., She, Y., and Walker, V.K. 2012. Biofilm, ice recrystallization inhibition and freeze-thaw protection in an epiphyte community. Appl. Biochem. Microbiol. **48**(4): 363–370.

Xu, H., Griffith, M., Patten, C., and Glick, B. 1998. Isolation and characterization of an antifreeze protein with ice nucleation activity from the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2. Can. J. Microbiol. **44**(1): 64-73.

Yamanaka, K., and Inouye, M. 2001. Induction of CspA, an *E. coli* major cold-shock protein, upon nutritional upshift at 37°C. Genes Cells **6**(4): 279-290.

Yamanaka, K., Mitta, M., and Inouye, M. 1999. Mutation analysis of the 5′ untranslated region of the cold shock *cspA* mRNA of *Escherichia coli*. J. Bacteriol. **181**(20): 6284-6291.

Yamashita, Y., Nakamura, N., Omiya, K., Nishikawa, J., Kawahara, H., and Obata, H. 2002. Identification of an antifreeze lipoprotein from *Moraxella* sp of Antarctic origin. Biosci. Biotech. Biochem. **66**(2): 239-247.

Yang, B., and Larson, T.J. 1998. Multiple promoters are responsible for transcription of the *glpEGR* operon of *Escherichia coli* K-12. Biochim. Biophys. Acta **1396**(1): 114-126.

Yanofsky, C. 1981. Attenuation in the control of expression of bacterial operons. Nature **289**(5800): 751-758.

Yu, F., Liu, X., Tao, Y., and Zhu, K. 2013. High saturated fatty acids proportion in *Escherichia coli* enhances the activity of ice-nucleation protein from *Pantoea ananatis*. FEMS Microbiol. Lett. **345**(2): 141-146.

Yu, S.O., Brown, A., Middleton, A.J., Tomczak, M.M., Walker, V.K., and Davies, P.L. 2010. Ice restructuring inhibition activities in antifreeze proteins with distinct differences in thermal hysteresis. Cryobiology **61**(3): 327–334.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31**(13): 3406-3415.