4-2021

Expression and Purification of Mitochondrial Group 3 LEA Protein, AfrLEA3m_29, from Diapause Embryos of Artemia franciscana

Ivan Dinh Nguyen

Expression and Purification of Mitochondrial Group 3 LEA Protein, AfrLEA3m_29, from Diapause Embryos of *Artemia franciscana*

Ivan Dinh Nguyen

Undergraduate honors thesis under the direction of

Dr. Steven C. Hand

Department of Biological Sciences

LSU Roger Hadfield Ogden Honors College

The Upper Division Honors Program

April 2021

Louisiana State University & Agricultural and Mechanical College

Baton Rouge, Louisiana

ABSTRACT

Embryos from *Artemia franciscana* have a demonstrated ability to survive extreme water stress. Their anhydrobiotic capabilities are linked to the expression of late embryogenesis abundant (LEA) proteins, for these proteins serve as molecular stabilizers that aid in desiccation tolerance. Three variants of AfrLEA3m have been discovered in the embryos of *Artemia franciscana,* and they all are targeted to the mitochondrion. The purpose of multiple variants of LEA proteins within the same organelle is not understood. Thus, this experiment focuses on expression and purification of the smallest variant of AfrLEA3m, AfrLEA3m_29. First, *Afrlea3m_29* was obtained from the pENTR/D-TOPO sequencing vector and inserted into the pET-30a expression vector. Next, the expression vector was transformed into Rosetta 2(DE3) Singles Competent *Escherichia coli* cells and induced with IPTG. To purify AfrLEA3m_29, His-Tag and anion exchange chromatography were performed. Western blotting confirmed the successful expression and purification of recombinant AfrLEA3m_29. The presence of a possible homodimer was identified, but a significant amount of breakdown product for AfrLEA3m_29 was also visible in the final purified sample. Densitometry revealed that purified AfrLEA3m_29 consisted of 37% contaminants, the majority of which was degradation product as identified by cross reactivity to primary antibody. These results will hopefully aid in clarifying the functional importance for expressing multiple LEA proteins with a given subcellular compartment and facilitate the purification of the remaining variants, AfrLEA3m_43 and AfrLEA3m_47.

INTRODUCTION

Anhydrobiosis

Anhydrobiosis is an extreme state in which organisms can survive with minimal water, and some of these organisms can tolerate as low as 1% tissue water (Crowe et al., 1992). This dehydrated state is normally associated with plant seeds and fungal spores, but animals such nematodes, tardigrades, and brine shrimp embryos also possess such tolerance to severe water stress (Browne et al., 2002; Hengherr et al., 2008; Hand et al., 2011). One characteristic associated with anhydrobiosis in some species (but not all) is elevated concentrations of trehalose in the dehydrated organism (Crowe, 2002). In addition to being a fuel source for organisms, trehalose prevents protein aggregation, slows enzyme inactivation, and preserves lipid membranes after incidences of dehydration and freezing (Yancey, 2005; Tapia and Koshland, 2014; Hand et al., 2018). During dehydration, trehalose replaces water by binding to membranes and macromolecules to form a sugar glass that maintains cellular structures (Crowe et al., 1998a; Crowe et al., 1998b; Yancey, 2005). Membrane leakage is decreased upon rehydration since trehalose prevents the phase transition of membranes from gel to liquid-crystalline (Crowe et al., 1989). Furthermore, anhydrobiosis has also been linked with the presence of late embryogenesis abundant (LEA) proteins that stabilize the trehalose sugar glasses along with various other mechanisms of preservation (Tunnacliffe and Wise, 2007; Hand et al., 2011; Hand and Menze, 2015). Originally discovered in desiccation resistant seeds (Dure et al., 1981), LEA proteins play a crucial role in stabilization of cells under desiccating conditions, and they have been extensively studied in the embryos of the brine shrimp, *Artemia franciscana* (Hand et al., 2011; Warner et al., 2012; Toxopeus et al., 2014; Hand and Menze, 2015; Hand et al., 2018). This thesis focuses on the expression and purification of a LEA protein variant from

A. franciscana, AfrLEA3m_29 (Boswell et al., 2014b). Thus far, the gene sequence has been reported, but the recombinant protein has not been expressed. Accordingly, the nucleic acid sequence for *Afrlea3m_29* was inserted into an appropriate expression vector, and bacterial cells were transformed. After induction with IPTG, AfrLEA3m_29 was purified from cell extracts with His-Tag chromatography and anion exchange chromatography. With Western blotting, I was successful in identifying AfrLEA3m_29 as the expressed protein, which displayed an estimated mass of 29 kDa.

Artemia franciscana

Thriving in the Great Salt Lake of Utah, adult *A. franciscana* do not express LEA proteins, for these protective proteins are only found in the diapause and post-diapause embryos of the brine shrimp (Podrabsky and Hand, 2015). Diapause is a developmentally programmed state of developmental and metabolic arrest, which is entered preceding the onset of adverse environmental conditions (Clegg et al., 1996; Denlinger, 2002; Hahn and Denlinger, 2011; Hand et al., 2016). Diapause may be hormonally regulated and occur in response to signaling factors received before the adverse period (Denlinger et al., 2012; Hand et al., 2016). In the case of *A. franciscana,* nauplius larvae are released during the warmer months, but adults switch to oviparous reproduction during the winter and release encysted diapause embryos (Clegg and Conte, 1980; Patil et al., 2013). These overwintering embryos float to the surface of the lake and sometimes wash ashore where they are exposed to cold and dehydration (Patil et al., 2013). Diapause embryos at the gastrula stage remain dormant until conditions become more temperate in the spring (Clegg and Conte, 1980; Qiu and MacRae, 2007; Podrabsky and Hand, 2015).

In diapause embryos of *A. franciscana*, mitochondrial metabolism is reduced to less than 1% compared to that of non-diapause embryos, and prior to the larval stage, the main energy

source for *A. franciscana* is trehalose (Patil et al., 2013). The presence of both LEA proteins and trehalose contribute to the ability of embryos to withstand freezing and water stress (Tunnacliffe and Wise, 2007; Hand et al., 2011). Surprisingly, these diapause embryos survive up to several years of desiccation, and upon rehydration, the diapause embryos resume metabolism and development (Nambu et al., 2008; Hand et al., 2018). The expression of LEA proteins and the ability of the brine shrimp to withstand water stress disappear once it enters the larval stage (Qiu and MacRae, 2007; Podrabsky and Hand, 2015; Hand et al., 2018). It is important to note that not only one, but a multitude of different LEA proteins are expressed in the diapause embryos (Warner et al., 2012).

Distribution and properties of LEA proteins

LEA proteins are grouped based on shared amino acid motifs, and LEA proteins found in animals typically are categorized as Group 3 LEA proteins (Wise, 2003). The Group 3 LEA proteins in the diapause embryos of *A. franciscana* include AfrLEA2 and AfrLEA3m (Boswell et al., 2014b). Additionally, *A. franciscana* has been found to also contain Group 1 and Group 6 LEA proteins, which have never been seen in animals previously (Hand et al., 2007; LeBlanc et al., 2019). Structurally, LEA proteins are extremely hydrophilic and are intrinsically disordered (Tunnacliffe and Wise, 2007). They exist in random coils when in the hydrated state, but secondary structure, especially alpha helix, increases when LEA proteins are gradually dehydrated (Hand et al., 2011; Boswell et al., 2014a). In AfrLEA2, gain of protective function has been associated with increased alpha helix structure (LeBlanc and Hand, 2021). Nevertheless, various functions of LEA proteins have been reported in both hydrated and dry states; for example, human cells expressing Group 3 LEA proteins displayed a reduction in protein aggregation when fully hydrated (Chakrabortee et al., 2007).

The purpose of multiple groups of LEA protein in *A. franciscana* might be explained in part by their localization to different areas in the cell (LeBlanc and Hand, 2020). For example, AfrLEA2 targets the nuclei and cytoplasm while AfrLEA3m targets the mitochondrion (Boswell and Hand, 2014; Boswell et al., 2014a). In terms of Group 1 LEA proteins, they target both the mitochondrion and the cytoplasm (Marunde et al., 2013). AfrLEA6, a Group 6 LEA protein, localizes in the cytoplasm (LeBlanc and Hand, 2020) and exist at cellular concentrations ten times lower than Group 3 LEA proteins in *A. franciscana* (Boswell et al., 2014b; LeBlanc et al., 2019). Despite their specific subcellular localizations, there appears to be no advantage in protection when cytoplasmic proteins are protected by cytoplasmic-localized LEA proteins compared to mitochondrial-localized LEA proteins and *vice versa* (Hand and Menze, 2015; LeBlanc and Hand, 2021). Functionally, both AfrLEA2 and AfrLEA3m can slow the loss of enzymatic activity and protect lipid membranes from damage caused by desiccation or freezing (Boswell et al., 2014a; Moore and Hand, 2016; LeBlanc and Hand, 2021). Similarly, Group 1 LEA proteins also provide cells with resistance against water stress and freezing (Toxopeus et al., 2014). Finally, AfrLEA6 protects proteins through the formation of membraneless organelles; it is predicted that these protein assemblies function to increase viscosity and create protective compartments for sensitive proteins during drying (Belott et al., 2020; LeBlanc and Hand, 2021). Overall, the central role of LEA proteins lies in their ability to protect against desiccation-induced damage (Hand and Menze, 2015).

Protein Degradation

Intrinsically disordered proteins, such as LEA proteins, are unstructured in the naturally hydrated state of cells, but they apparently avoid protein degradation that normally targets misfolded and unstable proteins (Suskiewicz et al., 2011). Present in all organisms, protein

chaperones play an essential role in detecting and stabilizing disordered proteins (Ellis, 1990). Chaperones function by recognizing hydrophobic portions of misfolded proteins, and failure to stabilize the protein results in degradation of the protein by proteasomes (Liberek et al., 2008). The protein-folding and protein-degradation systems are integral components in ensuring removal of malfunctioning proteins and prevention of protein aggregation (Marques et al., 2006; Liberek et al., 2008). Due to their unstructured regions, intrinsically disordered proteins are subjected to degradation by default by uncapped 20s proteasomes; however, nanny proteins can prevent degradation from 20s proteasomes by masking the unstructured regions on newly synthesized intrinsically disordered proteins (Tsvetkov et al., 2009). Once bound to a nanny protein, degradation of intrinsically disordered proteins is only possible through the ubiquitination pathway using the 26s proteasome (Tsvetkov et al., 2009). Additionally, intermolecular interactions with other intrinsically disordered proteins also provides protection against proteasomes (Suskiewicz et al., 2011). The comprehensive suite of mechanisms by which LEA proteins avoid protein degradation has not been fully identified.

Variants of AfrLEA3m

During the cloning of AfrLEA3m from *A. franciscana*, Boswell et al. (2014) discovered three new variants of this mitochondrial-targeted LEA protein. Using chicken antibodies raised against AfrLEA3m, the Western blot revealed four differently sized bands, which indicated the presence of separate proteins. The variants were described as AfrLEA3m_47, AfrLEA3m_43, and AfrLEA3m_29 based on the mass of each protein. Next, reverse transcription was used on the extracted RNA sequences, and the cDNA was transformed into *Escherichia coli* cells. From the transformed cell colonies, plasmids were isolated and sequenced (Boswell et al., 2014b). It is important to note that the protein variants appeared larger than originally expected since

intrinsically disordered proteins have decreased binding and migration on SDS gels (Tompa, 2002). The apparent sizes of AfrLEA3m_47, AfrLEA3m_43, and AfrLEA3m_29 were 62.3 ± 0.4 kDa, 48.6 ± 1.4 kDa, and 32.4 ± 1.1 kDa, respectively (Boswell et al., 2014b). These variants also possessed multiple single-nucleotide differences among their sequences, and as a result, it was concluded that each variant is coded from an independent gene rather than being splice variants (Boswell et al., 2014b).

Little is known about these new variants of LEA protein except for their localization to the mitochondrion. Furthermore, it is still a mystery as to why an organism would need to express so many different types of LEA proteins. To uncover the differences between these protein variants and explore the purpose they serve in the diapause embryos, this thesis will focus on the expression and purification of AfrLEA3m_29.

GENERAL OBJECTIVE

The objective of this thesis is to express and purify recombinant AfrLEA3m_29. These methods should be equally effective on the remaining AfrLEA3m protein variants, AfrLEA3m_47 and AfrLEA3m_43. After isolation of the targeted proteins, future experiments hopefully will deduce the role of each LEA protein variant in desiccation tolerance.

SPECIFIC AIMS:

A. Construction of the pET-30a/*Afrlea3m_29* recombinant plasmid and transformation into Rosetta™ 2(DE3) Singles™ Competent cells.
B. Induction AfrLEA3m_29 and purification with His-Tag chromatography and anion exchange chromatography.
C. Analysis of AfrLEA3m_29 with SDS-PAGE and Western Blotting.

METHODS

Construction of the recombinant plasmid and transformation

In order to express *Afrlea3m_29*, the sequence was transferred from the sequencing vector and ligated into an expression vector, which was designed to add a His-Tag to the expressed protein to facilitate purification. The original target sequence was obtained from preexisting One Shot TOP10® Chemically Competent *E. coli* cells (Invitrogen) containing the pENTR/D-TOPO/*Afrlea3M_29* recombinant plasmid (Boswell et al., 2014b). First, the cells were incubated overnight in Luria-Bertani (LB) broth at 37ºC and 300 RPM. To foster only kanamycin-resistant bacterial growth, 51.5 µM of kanamycin was added to the broth. After growth, the harvested cells were used for isolating the *Afrlea3m_29* plasmid with QIAprep96 Turbo Miniprep Kit (Qiagen). The isolated plasmid was used for amplification of the *Afrlea3m_29* sequence by the polymerase chain reaction (PCR) using AfrLEA3m_29 primers containing added restriction sites (**Table 1)** and the GeneAmp PCR System 9700. The PCR reaction mixture consisted of 2 µl of extracted plasmid DNA, 2 µl of primers, and 44 µl of master mix.

QIAquick PCR purification kit (Qiagen) was used to clean the PCR product before its confirmation on a 1.5% agarose gel with 0.5 µg/ml ethidium bromide. Gel electrophoresis was run for 90 min at 120 v, and the Bio-Rad ChemiDoc XRS Imager was used to visualize the length of the isolated plasmid. For sequence verification, the plasmids were sent to the LSU Genomics Facility for Sanger sequencing with the M13 forward and M13 reverse primers.

To prepare *Afrlea3m_29* for ligation, the PCR product and pET-30a expression vector were both digested using NdeI and XhoI restriction enzymes (New England Biolabs) at 37°C for 2 h. Afterwards, these enzymes were heat inactivated at 65ºC, and to ensure removal of digestion

Table 1. Primers and restriction enzymes used in PCR amplification and digestion of *Afrlea3m_29*.

	Restriction Enzyme	Primer Sequence
Forward Primer	Nde1	5’ ATTAGTCA*TATGTTGTCCAAGCGTTTAATT ‘3
Reverse Primer	Xho1	5’ ACCACCC*TCGAGTCTTTCATGAGCTCCAGA ‘3

[a] The underlined portion of the primer sequence indicates the restriction site of the respective restriction enzyme.

[b] The asterisks (*) indicates the exact location of cleavage for the restriction enzyme.

enzymes, the digested *Afrlea3m_29* gene and digested pET-30a expression vector were electrophoresed on agarose gel and purified with QIAquick Gel Extraction Kit (Qiagen). Ligation followed next using a mixture of 2 µl Ligase Buffer, 100 mM DTT, 10 mM ATP, 1 µl T4 DNA ligase, 9 µl DNA insert, and 5 µl of pET-30a vector. The ligation incubation was performed at 16ºC for 16 h.

The pET-30a/*Afrlea3m_29* recombinant plasmid was then inserted into Rosetta™ 2(DE3) Singles™ Competent Cells via heat shock transformation at 42°C for 1 min. The cells were thawed on ice before heat shock and returned to ice immediately after heat shock. Subsequently, they were incubated (37ºC, 250 RPM) for 3 h in Super Optimal broth with Catabolite repression (SOC broth), and the transformed Rosetta™ 2(DE3) Singles™ Competent Cells were plated on 1.5% agarose and LB broth containing 51.5 µM kanamycin and 105 µM chloramphenicol. Incubation of transformed cells occurred at 37ºC and continued until colonies were visible (typically, 36 to 48 h). In order to verify transformation of the Rosetta™ 2(DE3) Singles™ Competent Cells with the pET-30a/*Afrlea3m_29* recombinant plasmid, eight resulting colonies were randomly selected and plasmids isolated by miniprep using the QIAprep96 Turbo Miniprep Kit (Qiagen). The extracted plasmids were sent to the LSU Genomics Facility for Sanger sequencing with the T7 forward and T7 reverse primers. Glycerol stocks of the Rosetta™ 2(DE3) Singles™ Competent Cells that contained the pET-30a/*Afrlea3m_29* recombinant plasmid were prepared in LB broth containing 51.5 µM kanamycin and 105 µM chloramphenicol and transferred into autoclaved glycerol (50% final concentration). The glycerol stocks of transformed cells were stored at -80ºC.

Induction and purification of AfrLEA3m_29

From a 50% glycerol stock, Rosetta™ 2(DE3) competent cells containing the pET-30a/*Afrlea3m_29* recombinant plasmid were grown in six separate 25 ml tubes of LB broth overnight at 37ºC and 300 RPM. The broth contained 51.5 µM kanamycin and 105 µM chloramphenicol. Each overnight culture was transferred into 1 L of LB broth for further incubation at 37ºC and 150 RPM. Incubation continued until the log growth phase was achieved, which was indicated by a spectrometer reading of 0.6 OD measured at 600 nm. The cell broths were then induced with 1mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and incubated for an additional 2 h at 37°C and 150 RPM.

After induction of AfrLEA3m_29, the cells were centrifuged (5,000×*g*, 15 min, 4°C) into pellets and resuspended with 60 ml of HisTrap binding buffer (20 mM sodium phosphate, 20 mM imidazole, 0.5 M NaCl, pH 7.4) with two tablets of cOmplete, EDTA-free protease inhibitors (Roche). Next, the cells were sonicated at 70% amplitude for ten bursts of 10 s with 20 s of rest in between bursts. The lysed cells were then centrifuged (20,000×*g*, 30 min) at 4°C to remove cellular debris, and the resulting supernatant was retained and heat-treated for 20 min at 95°C before another cycle of centrifugation (20,000×*g*, 30 min) at 4°C. The heat treatment precipitated significant quantities of non-LEA proteins. Next, the supernatant was retained and processed through a 0.45 µM syringe filter.

Using fast protein liquid chromatography (FPLC), the AfrLEA3m_29 protein solution was initially purified with His-Tag chromatography. The protein was bound to a 5 ml HisTrap™ FF crude column with a 5 ml/min flow rate of HisTrap binding buffer. The proteins were then eluted at 1 ml/min with HisTrap elution buffer (20 mM sodium phosphate, 0.5 M imidazole, 0.5 M NaCl, pH 7.4). With the Primeview Evaluation program, purified protein fractions were

visualized at 280 nm and pooled together for overnight buffer dialysis at 4°C into anion exchange binding buffer (20 mM triethanolamine, 10 mM NaCl, pH 7).

The second purification step was anion exchange chromatography, which separated proteins on a 5 mL HiTrap™ Q FF column. The flow rate was 2 ml/min, during sample loading, and the elution flow rate was 1 ml/min using anion exchange elution buffer (20 mM triethanolamine, 1 M NaCl, pH 7). The eluted AfrLEA3m_29 was then detected using Primeview Evaluation and transferred into TEN buffer (100 µM EDTA, 10 mM TES, 50 mM NaCl, 0.02% sodium azide) using an Amicon Ultra-15 10K Centrifugal Filter at 23°C (5,000×*g*, 30 min). The concentration of purified proteins was determined using a modified Lowry Assay (Peterson, 1977).

Analysis of AfrLEA3m_29 with SDS-PAGE and Western blotting

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used to confirm the presence and estimate the purity of AfrLEA3m_29. A 4% stacking gel and 11% resolving gel were used for electrophoresis of the samples at 175 mV for 70 min using the Bio-Rad mini-Protean 3 Cell. After electrophoresis, bands were fixed in the gel using a 95% ethanol and 10% acetic acid solution. Then, the gel was stained with Coomassie Blue for 1 h and washed in solution containing 50% methanol and 10% acetic acid for 1 h. Furthermore, the gel was stored overnight in 25% methanol and 5% acetic acid to ensure removal of the stained background. The Bio-Rad ChemiDoc XRS Imager was used for high resolution visualization of protein bands, and density was analyzed with the Bio-Rad Imager Lab software.

For Western Blotting, electrophoresis was performed as mentioned above, but afterwards, protein samples were transferred from the gel to a Protran nitrocellulose membrane (Amersham) using a Bio-Rad Mini Trans-Blot Cell at 100 mV for 90 min. Protein transfer was performed in

Towbin's buffer (25 mM Tris-base, 192 mM glycine, 20% methanol, 0.025% SDS). Subsequently, to prevent unwanted protein binding, the membrane was blocked in 5% dry milk and TBST (20 mM Tris-base, 136 mM NaCl ,0.1% tween, pH 7.6). The primary antibody used to target AfrLEA3m was obtained from chickens raised by Aves Lab (chicken IgY). Incubation in primary antibody was performed at 4ºC overnight with gentle shaking. After washing with TBST, the membrane was incubated for 1 h in Goat anti-Chicken IgY that was coupled with horseradish peroxidase. The membrane was washed with TBST for 2 h with transfer into fresh solution every 30 min. Finally, the membrane was bathed in luminol and peroxide from the ECL^{TM} Prime Western Blotting Detection Reagent kit (Amersham), and bands were visualized with a Bio-Rad ChemiDoc XRS Imager using the Western blot auto-imaging protocol.

RESULTS

Transfer of Afrlea3m_29 *to expression vector and transformation of competent* E. coli *cells*

The pENTR/D-TOPO sequencing vector from One Shot® TOP10 Chemically competent *E. coli* cells was analyzed for possession of the *Afrlea3m_29* sequence (**Fig. 1)**. The size of the empty pENTR/D-TOPO sequencing vector was 2580 bp, while the size for *Afrlea3m_29* was 777 bp. Therefore, the expected size for sequencing vector with the insert gene was 3357 bp; however, supercoiled DNA tends to migrate quicker during electrophoresis on agarose gel (Aaij and Borst, 1972; Oppenheim, 1981). The pENTR/D-TOPO sequencing vector with *Afrlea3m_29* showed a band slightly under 3.0 kb (**Fig. 1)**.

For verification of *Afrlea3m_29* after PCR amplification using *Afrlea3m29* specific primers (**Table 1**) and the pENTR/D-TOPO sequencing vector as template, the PCR product was analyzed with electrophoresis. The PCR product for *Afrlea3m_29* showed a band near its expected size of 777 bp (**Fig. 2)**. With a size of 5422 bp, the mobility of an empty pET-30a expression vector was also analyzed. The uncut expression vector was expected to migrate slightly quicker during electrophoresis due to supercoiling. The band for the empty pET-30a expression vector was visible between 6.0 kb and 5.0 kb (**Fig. 2)**, but whether it migrated faster than expected was not readily resolvable. Sanger sequencing showed that the isolated *Afrlea3m_29* had 776 bp out 777 bp identical to the original sequence reported by Boswell et al. (2014b).

After digestion of *Afrlea3m_29* and the pET-30a expression vector with Nde1 and Xho1 (**Table 1)**, gel electrophoresis was used to remove restriction enzymes in preparation for ligation. Sticky ends were formed as a result of the digestion. The linearized expression vector and insert were identified near 5.0 kb and 0.8 kb, respectively **(Fig. 3)**, and both bands were extracted from the gel for ligation, which was performed with T4 DNA ligase as described earlier. After

transformation of the pET-30a/*Afrlea3m_29* plasmid into Rosetta™ 2(DE3) Singles™ Competent cells, transformed colonies were picked and plasmids isolated by miniprep. Sanger sequencing detected 774 bp of which 773 bp were identical to the original *Afrlea3m_29*, and the mismatched base pair was consistent with the mismatch found in the isolated *Afrlea3m_29* from the pENTR/D-TOPO sequencing vector.

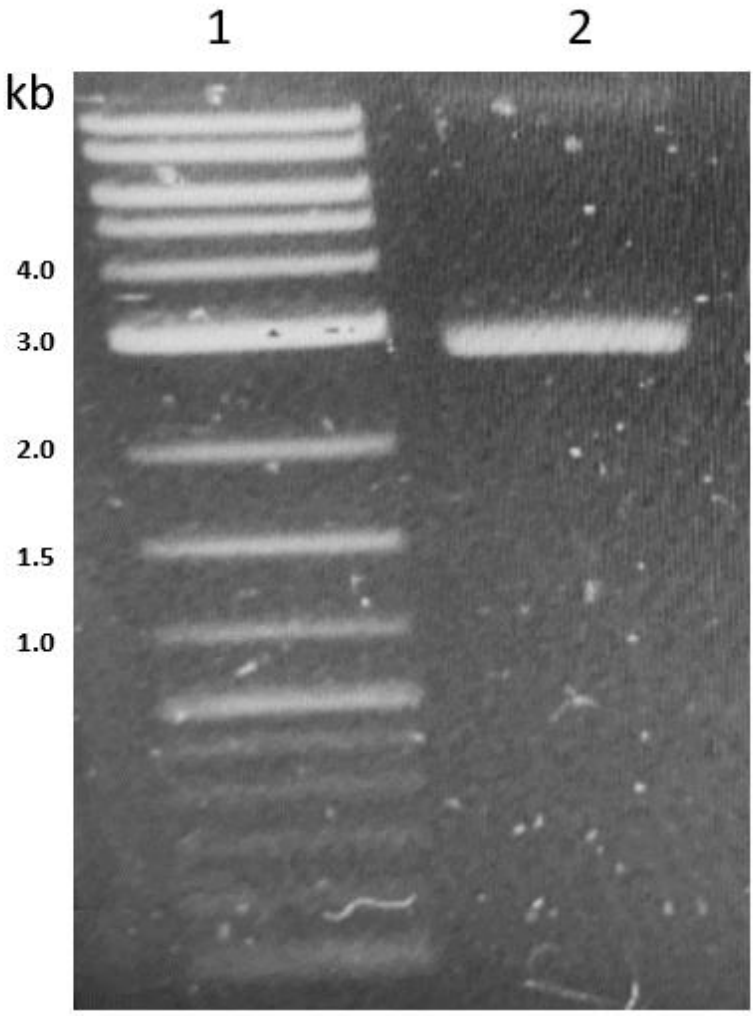

Figure 1. Agarose gel confirming the presence of an insert in the original pENTR/D-TOPO sequencing vector. Gel was performed with isolated plasmids after miniprep of One Shot® TOP10 Chemically competent *Escherichia coli* cells (Invitrogen). Size of the empty vector is 2580 bp; size for *Afrlea3m_29* is 777 bp. Lane 1 was loaded with a DNA ladder (kb).

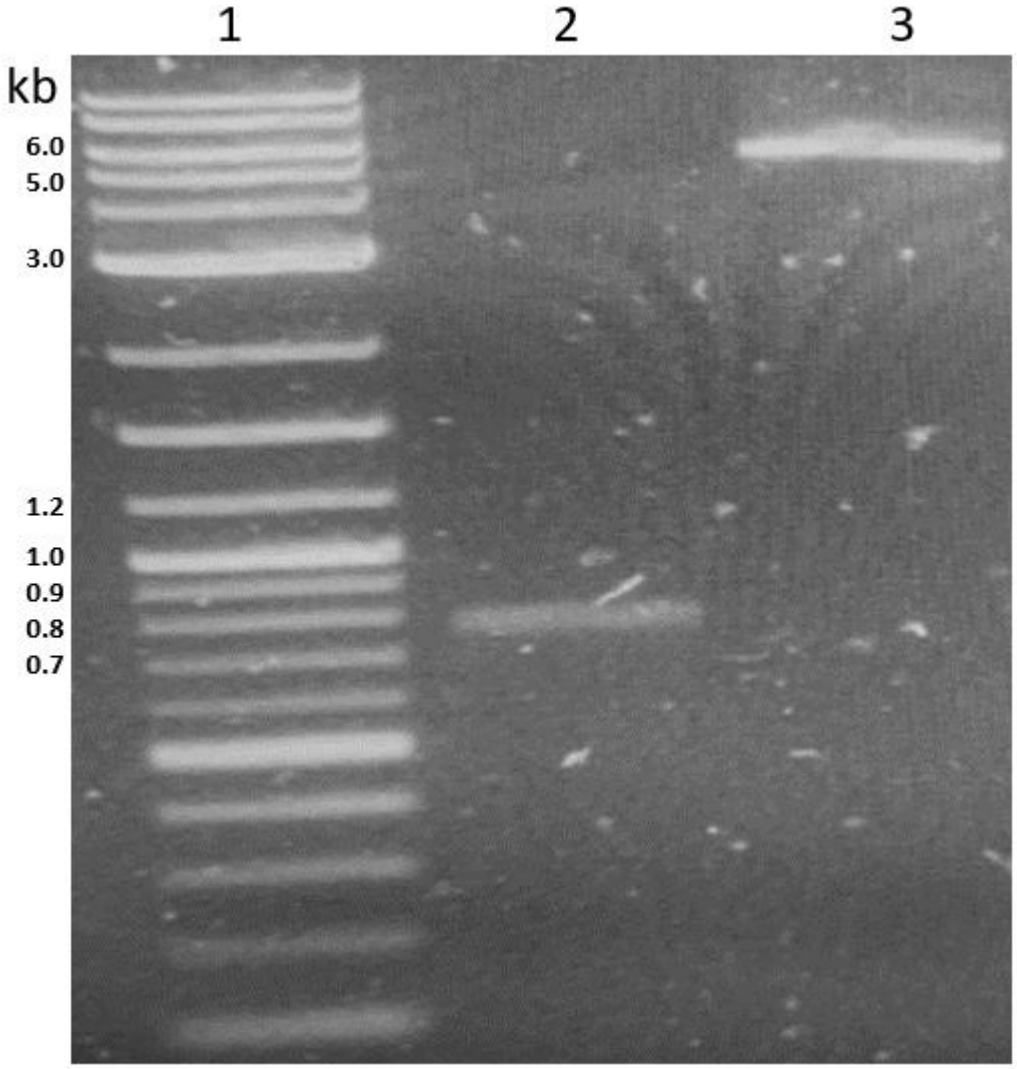

Figure 2. Agarose gel of PCR product that demonstrates the size of the insert in the original pENTR/D-TOPO sequencing vector matches that of *Afrlea3m_29* (lane 2) generated with *Afrlea3m_29* specific primers. The correct size of *Afrlea3m_29* is 777 bp. Lane 3 shows the mobility of the empty pET-30a expression vector (size = 5422 bp) that will be used in the next step for ligation with *Afrlea3m_29*. Lane 1 was loaded with a DNA ladder (kb).

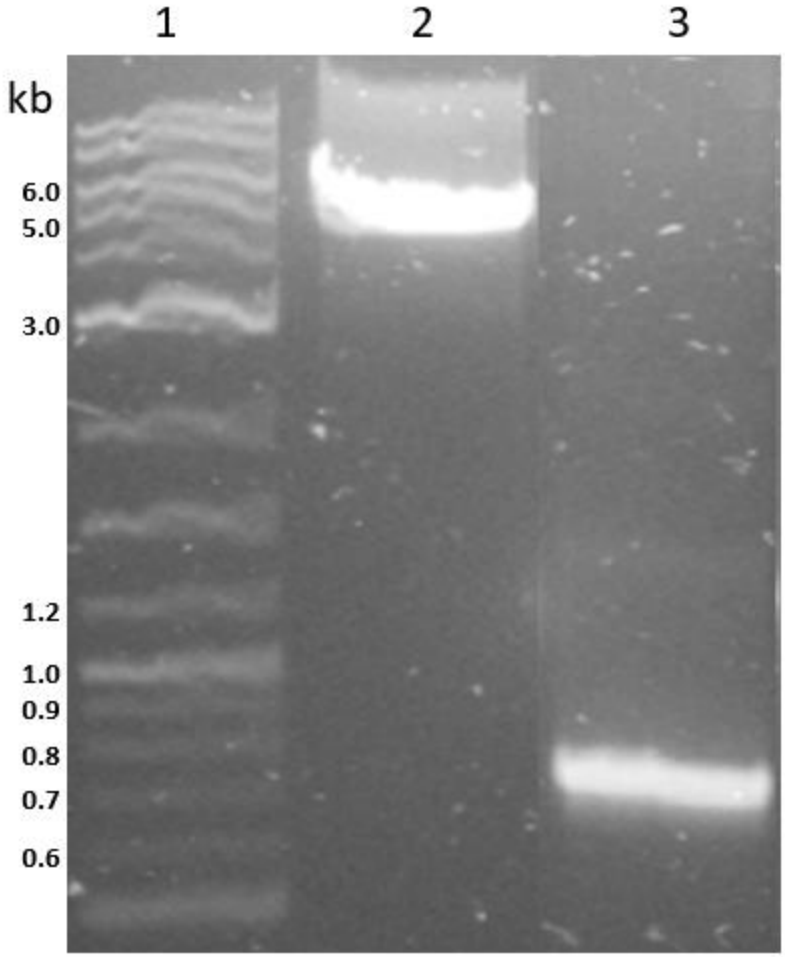

Figure 3. Agarose gel showing the pET-30a expression vector after digestion with Nde1 and Xho1 to prepare for ligation (lane 2), and the *Afrlea3m_29* insert digested with Nde1 and Xho1 to generate sticky ends (lane 3). These products were then cut and extracted from the gel. Lane 1, DNA ladder (kb).

Purification of AfrLEA3m_29 protein

After sonication and heat treatment of IPTG-induced Rosetta™ 2(DE3) Singles™ Competent cells containing the *Afrlea3m_29* gene, His-Tag chromatography was performed to purify recombinant AfrLEA3m_29. The extract of bacterial cells was loaded onto the HisTrap™ FF 5 ml column, and absorbance was monitored at 280 nm **(Fig. 4A)**. The void volume containing unbound proteins eluted in fractions 2-16 with a peak absorbance of 2600 mAu. The column was washed extensively with His-Trap binding buffer to allow absorbance to return to baseline. AfrLEA3m_29 was then released from the column with elution buffer that contained a high concentration of imidazole **(Fig. 4B)**. The eluted peak displayed a maximal absorbance of 1800 mAu **(Fig. 4B)**; fractions 34-42 were pooled and retained for anion exchange chromatography.

Anion exchange chromatography was performed with a HiTrap™ Q FF column to remove further impurities from the AfrLEA3m_29 sample obtained from the His-Trap column. Previous work has shown that this anion exchange step is useful for reducing the contamination from breakdown products of LEA proteins (Boswell et al., 2014b; LeBlanc et al., 2019). The sample was loaded in anion exchange binding buffer and rinsed thoroughly to removed unbound contaminants **(Fig. 5A)**; the peak absorbance for the void volume (215 mAu) was much lower than seen for the His-Trap column, which was as expected due to the higher purity of the starting sample. AfrLEA3m_29 was released from the column (peak absorbance, 2100 mAu) with a gradient of NaCl **(Fig. 5B)**. Fractions 32-35 were retained and exchanged into storage buffer.

Table 2 provides a summary of the purification steps and the final yield of AfrLEA3m_29 after anion exchange chromatography. The yield of AfrLEA3m_29 was 30 mg

of protein, which represented approximately 2% of the total starting protein present in the supernatant from the lysed bacterial cells.

The purity of recombinant AfrLEA3m_29 was estimated with Coomassie blue staining of a gel after SDS-PAGE **(Fig. 6A)**, and contaminants of AfrLEA3m_29 were analyzed with densitometry **(Fig. 6B)**. A dense band representing recombinant AfrLEA3m_29 was visible across all lanes and migrated at about 32 kDa, which as previously mentioned, is where AfrLEA3m_29 runs on these 11% gels (Boswell et al., 2014b). A light band around 64 kDa was also present for His-Tag, anion exchange, and buffer exchange samples, which likely represents a dimer of AfLEA3m_29 (see below). Densitometry indicates the final preparation is about 62-63% pure **(Fig. 6B)**.

A Western blot was performed to confirm the identity of recombinant AfrLEA3m_29 across each purification step, as well as potential breakdown products that also cross-react with the primary antibody **(Fig. 7)**. A dark band was visible near 32 kDa for all samples (sonication, His-Tag, anion exchange, buffer exchange). Additionally, a faint band was also visible at about 64 kDa in all lanes, which supports the suggestion of a dimer offered above. Dimers of LEA proteins from *A. franciscana* have been observed previously (Boswell et al., 2014b). The Western blot also revealed a breakdown product(s) of AfrLEA3m_29 at approximately 20 kDa in each lane. The anion exchange chromatography visually reduced the amount of breakdown product in the final sample (**Fig. 7**; compare lane 3 with lane 5).

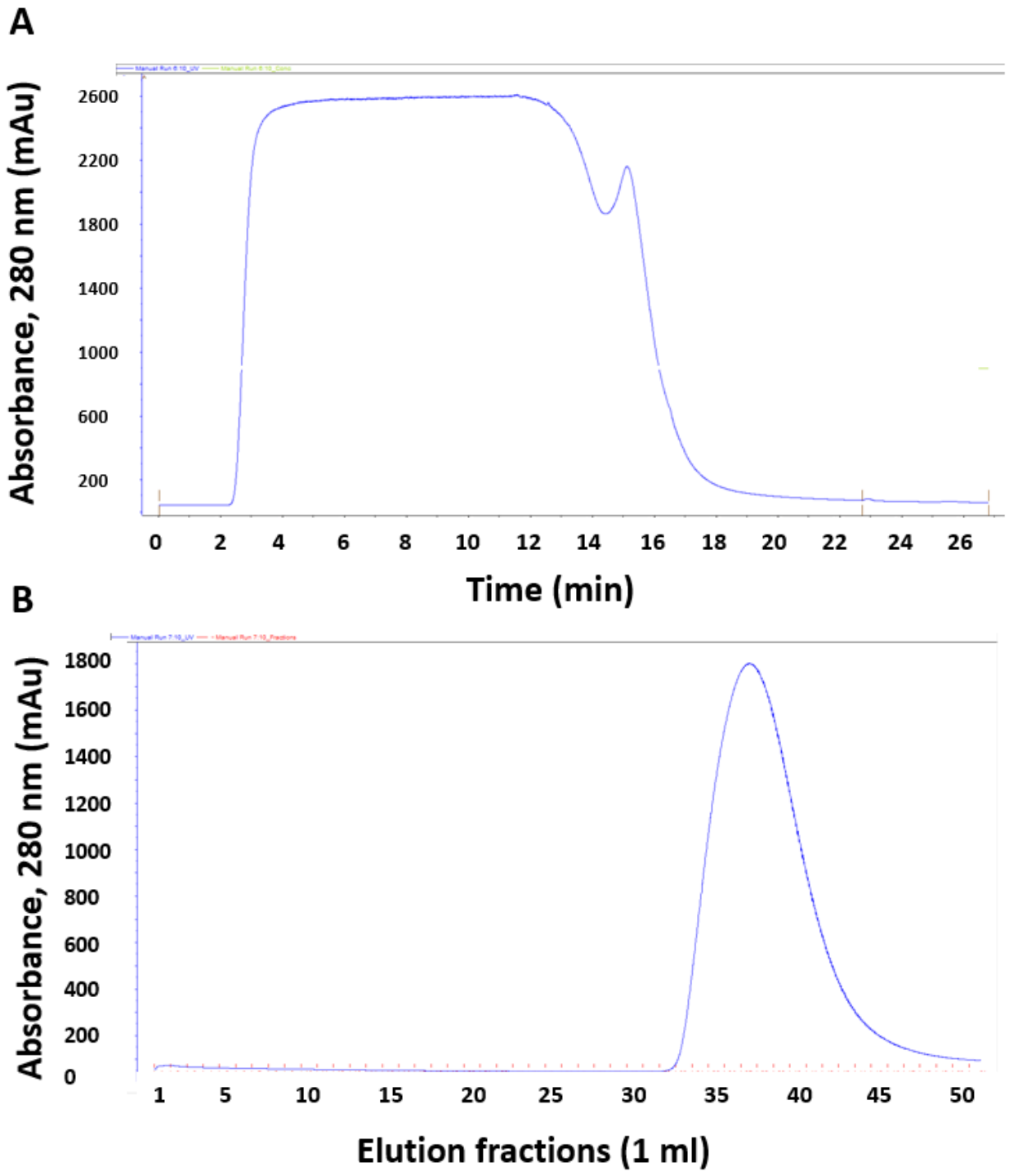

Figure 4. His-Tag chromatography of the supernatant from extracts of expression cells containing AfrLEA3m_29. **A**. Profile of the void volume eluting from the His-Tag column. **B**. Elution of AfrLEA3m_29 from the His-Tag column. Absorbance was measured at 280 nm, and the flow rate for loading was 5 ml//min without collection of individual fractions. For elution, 1 ml fractions were collected, and fractions 34-42 were retained and pooled. Graphs were generated by the Primeview Evaluation program.

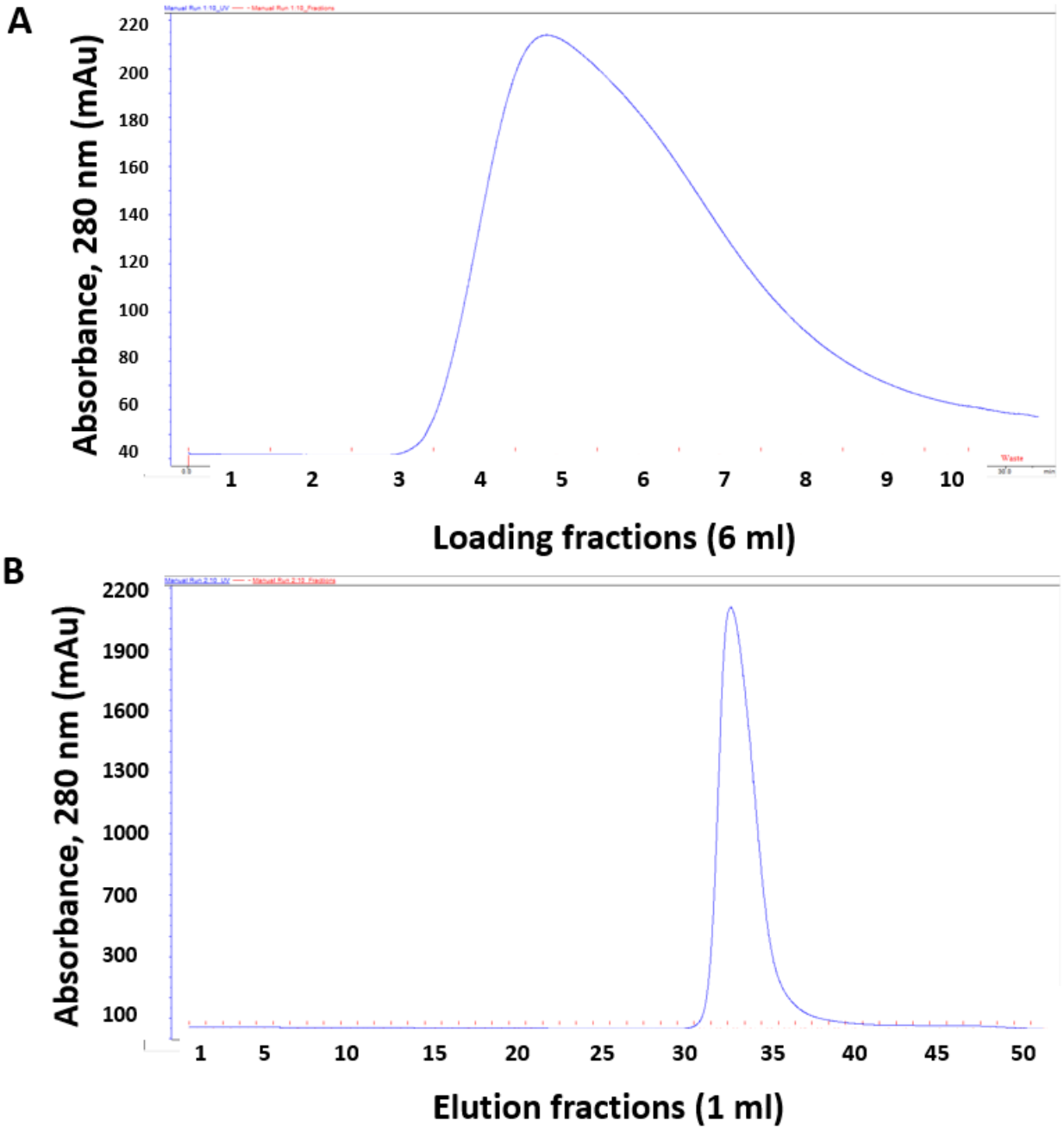

Figure 5. Anion exchange chromatography of a sample collected from the His-Tag column. **A**. Profile of the void volume eluting from the anion exchange column. **B**. Elution of AfrLEA3m_29 from the anion exchange column. Absorbance was measured at 280 nm, and the eluate was collected in 6 ml fractions (A) and 1 ml fractions (B). Fractions 32-35 that contained enriched AfrLEA3m_29 were retained and pooled. Graphs were generated by the Primeview Evaluation program.

Table 2. Purification of recombinant AfrLEA3m_29 from induced Rosetta 2 (DE3) expression cells.

Purification Step	Total Volume (ml)[b]	Protein Concentration (mg/ml)	Total Protein (mg)	Yield (% of total protein)[c]
Sonication[a]	65	20.3	1320	100
His-Tag	9	7.07	63.6	4.82
Anion Exchange	4	7.08	36.9	2.80
Buffer Exchange	0.2	151	30.2	2.29

[a] The sonication step refers to the supernatant collected after centrifugation of the sonicated Rosetta™ 2(DE3) Singles™ Competent cells.

[b] Volumes from His-Tag and anion exchange include the pooled fractions of purified proteins after the respective step.

[c] Yield refers to the total protein mass after each step relative to the total protein measured after sonication.

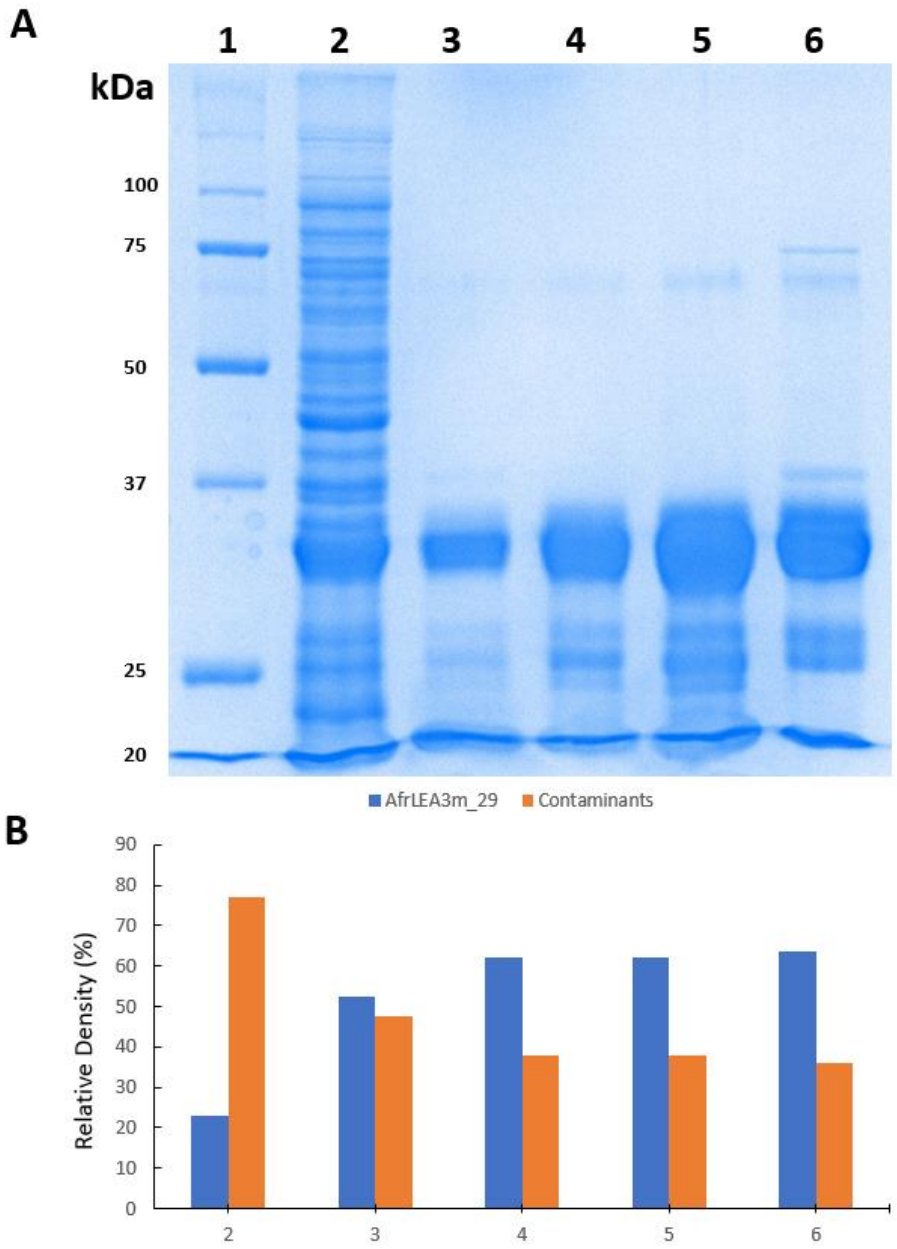

Figure 6. Assessment of purity of AfrLEA3m_29 across the purification steps. **A**. Coomassie Blue staining of gel after SDS-PAGE showing the purity of AfrLEA3m_29. **B**. Densitometry measurements of the bands found in lanes 2-6 in 'A' showing the relative amounts of AfrLEA3m_29 (blue) versus contaminants (orange). Lane 1 shows the kDa ladder. Each purification step is shown as follows: (lane 2) supernatant after centrifugation of sonicated expression cells, (lane 3) pooled fractions from His-Tag column, (lane 4) pooled fractions from anion exchange column, (lane 5) sample after buffer exchange, (lane 6) sample after buffer exchange from a separate/earlier purification AfrLEA3m_29. The Bio-Rad Image Lab software was used to compare the relative density of contaminants to AfrLEA3m_29 at 32 kDa.

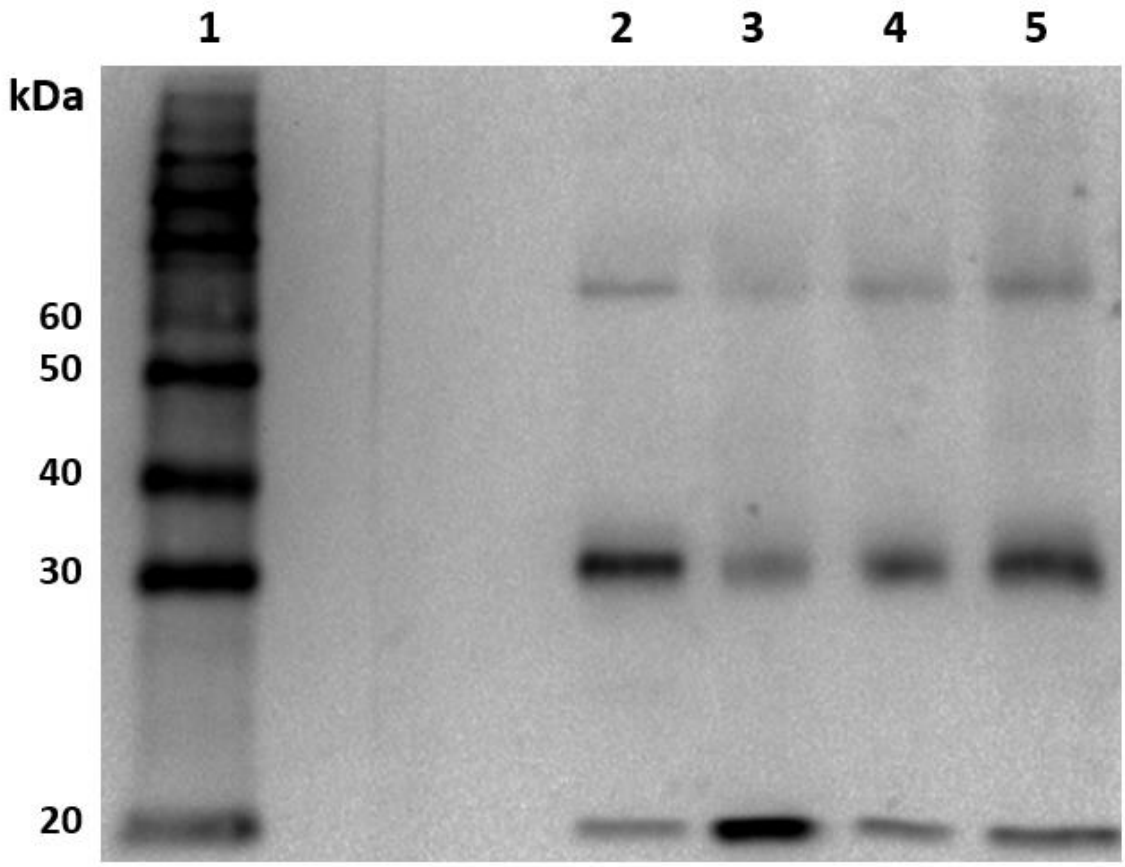

Figure 7. Western blot confirming the identity of AfrLEA3m_29, possible homodimer and breakdown products. (lane 1) kDa ladder, (lane 2) supernatant after centrifugation of sonicated expression cells, (lane 3) pooled fractions from His-Tag column, (lane 4) pooled fractions from anion exchange anion column, (lane 5) sample after buffer exchange. Protein band for AfrLEA3m_29 is visible at 32 kDa, and protein degradation product(s) visible near 20 kDa. Possible homodimer visible at 64 kDa.

DISCUSSION

For this thesis, I successfully expressed and purified recombinant AfrLEA3m_29 from *E. coli* cells. First, the original gene for *Afrlea3m_29* from the pENTR/D-TOPO sequencing vector was transferred into a pET-30a expression vector. The expression vector was then transformed into Rosetta™ 2(DE3) Singles™ Competent cells and successfully induced to express AfrLEA3m_29. Using His-Tag and anion exchange chromatography, I effectively purified AfrLEA3m_29, but interestingly, SDS-PAGE and Western blotting revealed a possible dimer and a substantial amount of breakdown products. With AfrLEA3m_29 successfully expressed and purified, future experimentation can be done to study the biological functions of AfrLEA3m_29. Additionally, the methods performed in this thesis should be applicable for purification of AfrLEA3m_43 and AfrLEA3m_47.

Cloning and expression of Afrlea3m_29

The primers used for amplification and preparing the sequence for expression in these experiments were similar to those in the original study that identified all variants of *Afrlea3m* (**Table 1)** (Boswell et al., 2014b). Due to bending of the DNA ladder (kb) in the top portion of the gel **(Fig. 3)**, the length of the digested pET-30a expression vector was not precisely estimated, but the uncut vector clearly was between 5.0 and 6.0 kDA (**Fig. 2**). Finally, sequencing the insert contained in the plasmids isolated from resulting from colonies of the transformed Rosetta 2(DE3) Competent cells confirmed the presence of *bona fide* coding sequence for *Afrlea3m_29*.

Both kanamycin and chloramphenicol were present in the growth medium for transformed Rosetta 2(DE3) competent cells containing the pET-30a/*Afrlea3m_29* expression vector. Presence of kanamycin ensured that growth was only possible for cells containing the

ligated pET-30a expression vector, which expressed kanamycin resistance. The chloramphenicol resistance for this strain of Rosetta 2(DE3) cells arose from the presence of a chloramphenicol-resistant plasmid that also supplied tRNAs for seven codons rarely used by *E. coli*, but which are important for enhancing the expression of eukaryotic proteins. Bacterial contaminants transformed with the ligated pET-30a expression vector or Rosetta 2(DE3) Competent cells without the expression vector were unable to form colonies since these would possess resistance to only one of the antibiotics.

Induction of the expression of AfrLEA3m_29 with IPTG was possible due to presence of the T7-lac system in the pET-30a expression vector. The T7 lac promoter, located above the insert gene in the expression vector, can only be recognized by the T7 RNA polymerase; however, both the T7 lac promoter and T7 RNA polymerase are inhibited by *lacI* in Rosetta 2(DE3) Competent *E. coli* cells (Rosano and Ceccarelli, 2014). Therefore, IPTG was required to remove the inhibitory effect of *lacI* to allow for expression of AfrLEA3m_29.

Purification and analysis of AfrLEA3m_29

Both Coomassie blue staining and Western blotting revealed that AfrLEA3m_29 was successfully purified and migrated with an estimated size of 32 kDa (**Fig. 6A, 7)**. A previous study also confirmed that AfrLEA3m_29 migrated similarly (Boswell et al., 2014b). As mentioned earlier, intrinsically disordered proteins tend to move slower than predicted based on mass on SDS gels (Tompa, 2002). Although Boswell et al. (2014b) only reported the presence of dimers for AfrLEA2, this thesis showed that a possible homodimer for AfrLEA3m_29 was present near 64 kDa (**Fig. 6A, 7)**. It is worth noting that the dye front migrated on top of the contaminants observed during SDS-PAGE and may have affected the densitometry results (**Fig. 6)**. Nevertheless, substantial degradation of the purified AfrLEA3m_29 occurred during

purification. Breakdown also was observed for AfrLEA3m by Boswell et al (2014b), but the problem appeared to be worse for AfrLEA3m_29. Thus far, anion exchange is the only procedure identified to eliminate these breakdown products. It might be possible to carefully evaluate the purity of AfrLEA3m_29, fraction by fraction, as it elutes from the anion exchange column, since previous work with AfrLEA6 has shown that detailed analysis of contaminants across the eluted peak helps pinpoint the best fractions to pool, thereby increasing the final purity (LeBlanc et al, 2019). Another possibility would be to adjust the steepness of the NaCl gradient used for elution from the anion exchange column, which could more effectively separate contaminating proteins.

Overall, the successfully expressed AfrLEA3m_29 can now be used for studies of protein structure and function, including analysis with circular dichroism to investigate the degree of intrinsic disorder of the hydrated protein and its gain of structure during drying. Being the smallest variant of AfrLEA3m, AfrLEA3m_29 has already been shown to target the mitochondrion (Boswell et al., 2014b), and similar to previous experiments on Group 3 LEA proteins, it should be tested for synergy with trehalose and ability to protect enzymatic activity during water stress (LeBlanc and Hand, 2021). It is also necessary to express and purify AfrLEA3m_43 and AfrLEA3m_47, and it would be interesting to see if high amounts of breakdown products are visible in these variants as well. By studying AfrLEA3m_29, a better understanding of LEA proteins in *A. franciscana* may be gained, especially why so many variants are necessary in their diapause embryos.

ACKNOWLEDGEMENTS

I would like to thank Dr. Steven C. Hand for the guidance and support that he provided me throughout my entire thesis project. I would also like to thank Elbert Hoang and Daniel Arabe for teaching me the necessary lab techniques to be successful. Appreciation is extended to Alex Landry for supporting me in my experiments. I would like to acknowledge Dr. Scott Herke (LSU Genomics Center) for his work on DNA sequencing. Thank you to Dr. Justin Ragains and Dr. Huangen Ding for consenting to serve on my thesis defense committee. Finally, I am indebted to Louisiana State University for giving me the opportunity to participate in scientific research.

LITERATURE CITED

Aaij, C., and Borst, P. (1972). The gel electrophoresis of DNA. Biochim Biophys Acta **269,** 192-200.

Belott, C., Janis, B., and Menze, M.A. (2020). Liquid-liquid phase separation promotes animal desiccation tolerance. Proc Natl Acad Sci U S A **117,** 27676-27684.

Boswell, L.C., and Hand, S.C. (2014). Intracellular localization of Group 3 LEA proteins in embryos of *Artemia franciscana*. Tissue Cell **46,** 514-519.

Boswell, L.C., Menze, M.A., and Hand, S.C. (2014a). Group 3 late embryogenesis abundant proteins from embryos of *Artemia franciscana*: structural properties and protective abilities during desiccation. Physiol Biochem Zool **87,** 640-651.

Boswell, L.C., Moore, D.S., and Hand, S.C. (2014b). Quantification of cellular protein expression and molecular features of Group 3 LEA proteins from embryos of *Artemia franciscana*. Cell Stress Chaperones **19,** 329-341.

Browne, J., Tunnacliffe, A., and Burnell, A. (2002). Anhydrobiosis: plant desiccation gene found in a nematode. Nature **416,** 38.

Chakrabortee, S., Boschetti, C., Walton, L.J., Sarkar, S., Rubinsztein, D.C., and Tunnacliffe, A. (2007). Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. Proc Natl Acad Sci U S A **104,** 18073-18078.

Clegg, J.S., and Conte, F.P. (1980). A review of the cellular and developmental biology of *Artemia*. In The Brine Shrimp *Artemia* (ed. G. Persoone, P. Sorgeloos, O. Roels and E. Jaspers) **Wetteren, Belgium: Universa Press,** 11-54.

Clegg, J.S., Drinkwater, L.E., and Sorgeloos, P. (1996). The metabolic status of diapause embryos of *Artemia franciscana* (SFB). Physiol Zool **69,** 49-66.

Crowe, J.H., Crowe, L.M., and Hoekstra, F.A. (1989). Phase transitions and permeability changes in dry membranes during rehydration. J Bioenerg Biomembr **21,** 77-91.

Crowe, J.H., Hoekstra, F.A., and Crowe, L.M. (1992). Anhydrobiosis. Annu Rev Physiol **54,** 579-599.

Crowe, J.H., Carpenter, J.F., and Crowe, L.M. (1998a). The role of vitrification in anhydrobiosis. Annu Rev Physiol **60,** 73-103.

Crowe, J.H., Clegg, J.S., and Crowe, L.M. (1998b). Anhydrobiosis: the water replacement hypothesis. In The Properties of Water in Foods ISOPOW 6, D.S. Reid, ed (Boston, MA: Springer US), pp. 440-455.

Crowe, L.M. (2002). Lessons from nature: the role of sugars in anhydrobiosis. Comp Biochem Physiol A Mol Integr Physiol **131,** 505-513.

Denlinger, D.L. (2002). Regulation of diapause. Annu Rev Entomol **47,** 93-122.

Denlinger, D.L., Yocum, G.D., and Rinehart, J.P. (2012). Hormonal control of diapause. insect Endocrinology**,** 430-463.

Dure, L., Greenway, S.C., and Galau, G.A. (1981). Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. Biochemistry **20,** 4162–4168.

Ellis, R.J. (1990). The molecular chaperone concept. Semin Cell Biol **1,** 1-9.

Hahn, D.A., and Denlinger, D.L. (2011). Energetics of insect diapause. Annu Rev Entomol **56,** 103-121.

Hand, S.C., and Menze, M.A. (2015). Molecular approaches for improving desiccation tolerance: insights from the brine shrimp *Artemia franciscana*. Planta **242,** 379-388.

Hand, S.C., Moore, D.S., and Patil, Y. (2018). Challenges during diapause and anhydrobiosis: Mitochondrial bioenergetics and desiccation tolerance. IUBMB Life **70,** 1251-1259.

Hand, S.C., Jones, D., Menze, M.A., and Witt, T.L. (2007). Life without water: expression of plant LEA genes by an anhydrobiotic arthropod. J Exp Zool A Ecol Genet Physiol **307,** 62-66.

Hand, S.C., Denlinger, D.L., Podrabsky, J.E., and Roy, R. (2016). Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish. Am J Physiol Regul Integr Comp Physiol **310,** R1193-1211.

Hand, S.C., Menze, M.A., Toner, M., Boswell, L., and Moore, D. (2011). LEA proteins during water stress: not just for plants anymore. Annu Rev Physiol **73,** 115-134.

Hengherr, S., Heyer, A.G., Kohler, H.R., and Schill, R.O. (2008). Trehalose and anhydrobiosis in tardigrades--evidence for divergence in responses to dehydration. FEBS J **275,** 281-288.

LeBlanc, B.M., and Hand, S.C. (2020). A novel Group 6 LEA protein from diapause embryos of *Artemia franciscana* is cytoplasmically localized. Tissue Cell **67,** 101410.

LeBlanc, B.M., and Hand, S.C. (2021). Target enzymes are stabilized by AfrLEA6 and a gain of alpha-helix coincides with protection by a Group 3 LEA protein during incremental drying. Biochim Biophys Acta Proteins Proteom **1869,** 140642.

LeBlanc, B.M., Le, M.T., Janis, B., Menze, M.A., and Hand, S.C. (2019). Structural properties and cellular expression of AfrLEA6, a Group 6 late embryogenesis abundant protein from embryos of *Artemia franciscana*. Cell Stress Chaperones **24,** 979-990.

Liberek, K., Lewandowska, A., and Zietkiewicz, S. (2008). Chaperones in control of protein disaggregation. EMBO J **27,** 328-335.

Marques, C., Guo, W., Pereira, P., Taylor, A., Patterson, C., Evans, P.C., and Shang, F. (2006). The triage of damaged proteins: degradation by the ubiquitin-proteasome pathway or repair by molecular chaperones. Faseb J **20,** 741-743.

Marunde, M.R., Samarajeewa, D.A., Anderson, J., Li, S., Hand, S.C., and Menze, M.A. (2013). Improved tolerance to salt and water stress in *Drosophila melanogaster* cells conferred by late embryogenesis abundant protein. J Insect Physiol **59,** 377-386.

Moore, D.S., and Hand, S.C. (2016). Cryopreservation of lipid bilayers by LEA proteins from *Artemia franciscana* and trehalose. Cryobiology **73,** 240-247.

Nambu, Z., Tanaka, S., Nambu, F., and Nakano, M. (2008). Influence of temperature and darkness on embryonic diapause termination in dormant *Artemia* cysts that have never been desiccated. J Exp Zool A Ecol Genet Physiol **309,** 17-24.

Oppenheim, A. (1981) Separation of closed circular DNA from linear DNA by electrophoresis in two dimensions in agarose gels. Nucleic Acids Res. **9(24)**: 6805-6812.

Patil, Y.N., Marden, B., Brand, M.D., and Hand, S.C. (2013). Metabolic downregulation and inhibition of carbohydrate catabolism during diapause in embryos of *Artemia franciscana*. Physiol Biochem Zool **86,** 106-118.

Peterson, G.L. (1977). A simplification of the protein assay method of Lowry et al. which is more generally applicable. Anal Biochem **83,** 346-356.

Podrabsky, J.E., and Hand, S.C. (2015). Physiological strategies during animal diapause: lessons from brine shrimp and annual killifish. J Exp Biol **218,** 1897-1906.

Qiu, Z., and MacRae, T.H. (2007). Developmentally regulated synthesis of p8, a stress-associated transcription cofactor, in diapause-destined embryos of *Artemia franciscana*. Cell Stress Chaperones **12,** 255-264.

Rosano, G.L., and Ceccarelli, E.A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol **5,** 172.

Suskiewicz, M.J., Sussman, J.L., Silman, I., and Shaul, Y. (2011). Context-dependent resistance to proteolysis of intrinsically disordered proteins. Protein Sci **20,** 1285-1297.

Tapia, H., and Koshland, D.E. (2014). Trehalose is a versatile and long-lived chaperone for desiccation tolerance. Curr. Biol. **24,** 2758–2766.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem Sci **27,** 527-533.

Toxopeus, J., Warner, A.H., and MacRae, T.H. (2014). Group 1 LEA proteins contribute to the desiccation and freeze tolerance of *Artemia franciscana* embryos during diapause. Cell Stress Chaperones **19,** 939-948.

Tsvetkov, P., Reuven, N., and Shaul, Y. (2009). The nanny model for IDPs. Nat Chem Biol **5,** 778-781.

Tunnacliffe, A., and Wise, M.J. (2007). The continuing conundrum of the LEA proteins. Naturwissenschaften **94,** 791-812.

Warner, A.H., Chakrabortee, S., Tunnacliffe, A., and Clegg, J.S. (2012). Complexity of the heat-soluble LEA proteome in *Artemia* species. Comp Biochem Physiol Part D Genomics Proteomics **7,** 260-267.

Wise, M.J. (2003). LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. BMC Bioinformatics **4,** 52.

Yancey, P.H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J Exp Biol **208,** 2819-2830.