Expression of Artemia LEA Proteins in Drosophila melanogaster Cells Using Multicistronic Vector Constructs

Kazi Nazrul Islam  
*Eastern Illinois University*

This research is a product of the graduate program in **Biological Sciences** at Eastern Illinois University. Find out more about the program.

---

**Recommended Citation**

https://thekeep.eiu.edu/theses/2454

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.
Thesis Maintenance and Reproduction Certificate

FOR: Graduate Candidates Completing Theses in Partial Fulfillment of the Degree
Graduate Faculty Advisors Directing the Theses

RE: Preservation, Reproduction, and Distribution of Thesis Research

Preserving, reproducing, and distributing thesis research is an important part of Booth Library’s responsibility to provide access to scholarship. In order to further this goal, Booth Library makes all graduate theses completed as part of a degree program at Eastern Illinois University available for personal study, research, and other not-for-profit educational purposes. Under 17 U.S.C. § 108, the library may reproduce and distribute a copy without infringing on copyright; however, professional courtesy dictates that permission be requested from the author before doing so.

Your signatures affirm the following:

- The graduate candidate is the author of this thesis.
- The graduate candidate retains the copyright and intellectual property rights associated with the original research, creative activity, and intellectual or artistic content of the thesis.
- The graduate candidate certifies her/his compliance with federal copyright law (Title 17 of the U. S. Code) and her/his right to authorize reproduction and distribution of all copyrighted materials included in this thesis.
- The graduate candidate in consultation with the faculty advisor grants Booth Library the non-exclusive, perpetual right to make copies of the thesis freely and publicly available without restriction, by means of any current or successive technology, including by not limited to photocopying, microfilm, digitization, or internet.
- The graduate candidate acknowledges that by depositing her/his thesis with Booth Library, her/his work is available for viewing by the public and may be borrowed through the library’s circulation and interlibrary loan departments, or accessed electronically.
- The graduate candidate waives the confidentiality provisions of the Family Educational Rights and Privacy Act (FERPA) (20 U. S. C. § 1232g; 34 CFR Part 99) with respect to the contents of the thesis and with respect to information concerning authorship of the thesis, including name and status as a student at Eastern Illinois University.

I have conferred with my graduate faculty advisor. My signature below indicates that I have read and agree with the above statements, and hereby give my permission to allow Booth Library to reproduce and distribute my thesis. My adviser’s signature indicates concurrence to reproduce and distribute the thesis.

Graduate Candidate Signature ____________________ Faculty Adviser Signature ____________________
Kazi Nazrul Islam Michael A. Menze
Printed Name Printed Name
Biological Sciences
Graduate Degree Program

Please submit in duplicate.
Expression of Artemia LEA proteins in Drosophila melanogaster cells

using multicistronic vector constructs

(TITLE)

BY
Kazi Nazrul Islam

THESIS
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Master of Science in Biological Sciences
IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

2016
YEAR

I HEREBY RECOMMEND THAT THIS THESIS BE ACCEPTED AS FULFILLING
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

THESIS COMMITTEE CHAIR      DATE

DEPARTMENT/SCHOOL CHAIR      DATE
OR CHAIR'S DESIGNEE

THESIS COMMITTEE MEMBER      DATE

THESIS COMMITTEE MEMBER      DATE

THESIS COMMITTEE MEMBER      DATE

THESIS COMMITTEE MEMBER      DATE
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>PAGE NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>3</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES AND FIGURES</td>
<td>5</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>9</td>
</tr>
<tr>
<td>MATERIALS AND METHODS</td>
<td>23</td>
</tr>
<tr>
<td>RESULTS</td>
<td>39</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>65</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>72</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>73</td>
</tr>
</tbody>
</table>
ABSTRACT

In nature several organisms exhibit anhydrobiosis, the outstanding feature to survive in extreme desiccation by entering into a state of dormancy known as diapause. The cyst of the brine shrimp *Artemia franciscana* shows anhydrobiosis by entering into a diapause phase. Previous studies showed a correlation between anhydrobiosis and expression of highly hydrophilic polypeptides termed late embryogenesis abundant (LEA) proteins. However, the precise molecular mechanisms of LEA proteins are still unknown. The presence of multiple LEA proteins in *Artemia* suggests that some of them might work together. Here, I aimed to express different combination of two LEA proteins from *Artemia franciscana* in the *Drosophila melanogaster* cell lines Kc167 and S2R+ by using the multicistronic vector pAc5-STABLE2-Neo. Immunoblot confirmed concurrent expression of both mCherry-LEA3m and GFP-LEA6 proteins in the Kc167 cells transfected with LEA3m+LEA6 construct. However, in three other Kc167 clones, although Western blot verified expression of mCherry tagged LEA proteins transcribed at first position of the vector, GFP tagged LEA protein cloned at second position of the vector was not detected. Another assumption that consensus ribosome recognition sequence of *Drosophila* would improve expression of LEA proteins in *Drosophila* cells was supported by the images of fluorescence microscopy. The final goal, simultaneous expression of two LEA proteins without the fluorescent reporters, was partially successful as immunoblot identified only DDK tagged first LEA proteins but not 6X His tagged second LEA proteins. Nonetheless, our results showed that expression of two LEA proteins concurrently in the *Drosophila* cells is possible by using the multicistronic vector instead of conventional two vectors system.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Michael A. Menze for his continuous support in my graduate study and research, as well as his perseverance and motivation. His guidance has made my research and writing of this thesis possible. Before joining the lab, I did not have much knowledge in molecular biology but his excellent guidance and depth knowledge made me competent in laboratory techniques as well as theoretical knowledge. I could not have imagined having a better advisor for my graduate study.

Besides my advisor, I would like to thank Dr. Gary A. Bulla and Dr. Britto P. Nathan as members of my thesis committee. Their encouragement in my current, as well as future graduate studies helped me to finish writing this thesis. Their helpfulness and insightful questions helped me to produce the most viable results.

I also thank my fellow lab mates including Brett Janis, Clinton Belott, Daniel Webster, Robert Skolik, David Grimm, David Constantinescu, Fazlul Karim, and Shankar Gurung for their assistance and inspiring discussions and all the fun we have had in the last two years. As an international student, it would be difficult to study and concentrate in research without their cheerful presence.
**LIST OF TABLES AND FIGURES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>LEA proteins found in the brine shrimp <em>Artemia franciscana</em></td>
<td>11</td>
</tr>
<tr>
<td>Table 2</td>
<td>The primers used for cloning of fluorescent marked LEA proteins in the STABLE-2 vector</td>
<td>25</td>
</tr>
<tr>
<td>Table 3</td>
<td>The primers used for cloning of mCherry marked LEA1.3 proteins in the STABLE-2 vector with and without ribosome recognition sequence</td>
<td>26</td>
</tr>
<tr>
<td>Table 4</td>
<td>The primers used for PCR amplification and cloning of LEA proteins in the STABLE-2 vector with 6X His and DDK tags</td>
<td>27</td>
</tr>
<tr>
<td>Table 5</td>
<td>The reagents used to amplify DNA through PCR</td>
<td>28</td>
</tr>
<tr>
<td>Table 6</td>
<td>The reagents used to set up reactions for restriction enzyme digestion</td>
<td>29</td>
</tr>
<tr>
<td>Table 7</td>
<td>The reagents used to set up ligation reactions</td>
<td>30</td>
</tr>
<tr>
<td>Table 8</td>
<td>Optical density of standard BSA samples in Bradford reagent at 595 nm</td>
<td>35</td>
</tr>
<tr>
<td>Figure 1</td>
<td>Schematic representation of the pAc5-STABLE2-Neo vector</td>
<td>23</td>
</tr>
<tr>
<td>Figure 2</td>
<td>PCR amplification of LEA3m DNA (921 bases) using isolated plasmids from bacterial colonies as template</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Restriction enzyme digestion of the plasmids that yielded LEA3m DNA in PCR</td>
<td>40</td>
</tr>
<tr>
<td>Figure 4</td>
<td>PCR products of LEA6 (771 bases) using isolated plasmids as template DNA</td>
<td>41</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Figure 5</td>
<td>Single digestion of the isolated plasmids that showed positive PCR result for gene LEA6</td>
<td></td>
</tr>
<tr>
<td>Figure 6</td>
<td>Amplification of LEA1 (1074 bases) and LEA2 (1092 bases) through PCR using isolated plasmids as template</td>
<td></td>
</tr>
<tr>
<td>Figure 7</td>
<td>Linearized plasmids that showed positive PCR result for LEA1 and LEA2</td>
<td></td>
</tr>
<tr>
<td>Figure 8</td>
<td>Single digestion of the plasmids purified from the bacteria cloned with LEA1.3+LEA6</td>
<td></td>
</tr>
<tr>
<td>Figure 9</td>
<td>PCR products of LEA 1.3 (600 bases) using isolated plasmids as template DNA</td>
<td></td>
</tr>
<tr>
<td>Figure 10</td>
<td>Expression of mCherry (red, column 1) and GFP (green, column 2) tagged LEA proteins in Kc167 cell lines</td>
<td></td>
</tr>
<tr>
<td>Figure 11</td>
<td>Expression of mCherry (red, column 1) and GFP (green, column 2) labeled LEA proteins in S2R+ cell lines</td>
<td></td>
</tr>
<tr>
<td>Figure 12</td>
<td>Immunoblotting for identification of fluorescent tagged LEA proteins in the <em>Drosophila</em> Kc167 cell lines</td>
<td></td>
</tr>
<tr>
<td>Figure 13</td>
<td>Immunoblotting to detect mCherry tagged LEA1.3 and LEA2 proteins in the <em>Drosophila</em> Kc167 cell lines</td>
<td></td>
</tr>
<tr>
<td>Figure 14</td>
<td>Immunoblotting for identification of GFP tagged LEA protein in the <em>Drosophila</em> Kc167 cell lines</td>
<td></td>
</tr>
<tr>
<td>Figure 15</td>
<td>PCR products of LEA1.1 (546 bases) using purified plasmids as template DNA</td>
<td></td>
</tr>
<tr>
<td>Figure 16</td>
<td>Single digestion of the plasmids that showed positive PCR result for LEA1.1</td>
<td></td>
</tr>
<tr>
<td>Figure 17</td>
<td>Amplification of LEA1.1 DNA (546 bases) through PCR using isolated plasmids as templated DNA</td>
<td>53</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Single digestion of the plasmids that showed positive result in PCR for LEA1.1 gene</td>
<td>53</td>
</tr>
<tr>
<td>Figure 19</td>
<td>PCR products of LEA1.3 (600 bases) using isolated plasmid DNA as template</td>
<td>55</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Digestion of the plasmids that showed PCR products of LEA1.3</td>
<td>55</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Expression of mCherry tagged LEA1.3 with and without ribosome recognition sequences (red, column 1) and GFP fused LEA6 proteins (green, column 2)</td>
<td>56</td>
</tr>
<tr>
<td>Figure 22</td>
<td>PCR amplification of LEA6-His (771 bases) using isolated plasmids as template</td>
<td>57</td>
</tr>
<tr>
<td>Figure 23</td>
<td>Linearized empty vector and plasmids that yielded DNA of LEA6-His</td>
<td>58</td>
</tr>
<tr>
<td>Figure 24</td>
<td>PCR amplification of LEA1-DDK (1074 bases) using isolated plasmids as template</td>
<td>59</td>
</tr>
<tr>
<td>Figure 25</td>
<td>Single digestion of the plasmids that showed positive result in PCR for LEA1-DDK</td>
<td>59</td>
</tr>
<tr>
<td>Figure 26</td>
<td>PCR amplification LEA2-DDK (1092 bases) using isolated plasmids as template</td>
<td>60</td>
</tr>
<tr>
<td>Figure 27</td>
<td>PCR product of LEA3m-DDK from the isolated plasmids</td>
<td>61</td>
</tr>
<tr>
<td>Figure 28</td>
<td>Single and double digestion of the plasmids (# 5) that showed positive result in PCR for LEA2 and LEA3m</td>
<td>61</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>PCR products of LEA1.3-DDK (600 bases) from the isolated plasmids</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Double digestion of the three isolated plasmids (# 1-3) that showed positive PCR result for LEA1.3-DDK</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Immunoblotting to detect DDK tagged LEA3m and LEA2 proteins in <em>Drosophila Kc167</em> cell lines</td>
<td></td>
</tr>
</tbody>
</table>


INTRODUCTION

Anhydrobiosis

The term anhydrobiosis literally translates to ‘life without water’ and denotes to the unique ametabolic state that enables an organism to maintain viability even after losing more than 97% of its body water (Gusev et al. 2014). David Wharton (2015) defined anhydrobiosis as an extraordinary capacity of an organism to survive the loss of almost all body water and to enter into a state of dormancy in which metabolism comes to a standstill. Anhydrobiotic organisms can endure loss of almost all cellular water, surviving dry conditions for prolonged periods, even at subzero temperatures (Toxopeus et al. 2014). Anhydrobiosis has been reported in cyanobacteria, yeast, lichens, algae, mosses, some plant seeds, and resurrection plants. Animals that exhibit anhydrobiosis are small and relatively simple invertebrates including insects, nematodes, rotifers, tardigrades, and the crustacean Artemia franciscana (Wharton 2015).

Anhydrobiosis in the brine shrimp

The brine shrimp Artemia franciscana is a primitive arthropod that lives in saline water (Kim et al. 2015) and undergoes either oviparous or ovoviviparous development (MacRae 2003). When environmental conditions are favorable brine shrimp develop ovoviviparously by yielding free swimming larvae (nauplii), however, under adverse conditions oviparously developing embryos arrest at gastrulation and are released from females as cysts before entering diapause (MacRae 2016). Diapause is a genetically programmed arrest of development that can occur at the embryonic, larval, pupal, or adult stage, depending on the species (Danilevskii 1965, Tauber et al. 1986). Entering diapause promotes survival of some organisms during exposure to temperature
fluctuation, desiccation, and hypoxia (Robbins et al. 2010). The cyst of the brine shrimp is able to tolerate complete desiccation, long-term anoxia, and low temperatures without loss in viability (Clegg 2000). The cysts synthesize various proteins including late embryogenesis abundant (LEA) proteins, small heat shock proteins (sHSP), and artemin before the onset of water stress (Kim et al. 2015). LEA proteins are believed to be critical for desiccation tolerance since an organism’s expression levels of LEA protein and mRNA are closely related to its capacity for water loss (Browne et al. 2004, Menze et al. 2009). Indeed, knockout of Group1 LEA proteins reduce survival of *Artemia franciscana* embryos after desiccation and freezing (Toxopeus et al. 2014). In addition to LEA proteins, trehalose which is a non-reducing disaccharide, contributes to the extreme desiccation tolerance in this animal (Clegg 1965, Crowe et al. 1977, Hand et al. 2011, MacRae 2016).

**Late embryogenesis abundant (LEA) proteins**

LEA proteins are hydrophilic, intrinsically disordered, flexible proteins (MacRae 2016), but during desiccation many LEA proteins assume their native conformation (Hand et al. 2011). LEA proteins were originally discovered in the late stages of embryo development in cotton seeds (Dure et al. 1981). In plants, most LEA proteins accumulate to high concentrations in embryonic tissues during the last stages of seed development when desiccation occurs (Ingram 1996). As orthodox seeds acquire the ability to withstand severe dehydration at this stage, LEA proteins have been associated with desiccation tolerance (Dure et al. 1981). In addition to plants, LEA proteins have been identified in some microorganisms, fungi, protozoa, rotifers, nematodes, insects, and the crustacean *Artemia franciscana* (Amara et al. 2014). LEA proteins found in *Artemia* cyst are similar to those in seeds and other anhydrobiotic organisms, and they protect proteins and membranes during desiccation (MacRae
2016). However, the precise molecular mechanisms of how LEA proteins work in anhydrobiotic organisms yet to elucidate.

**Classification of LEA proteins**

LEA proteins were initially classified into six groups or families based on specific domains and sequence motifs in the amino acid composition (Dure et al. 1989). Recently, Amara et al. (2014) organized LEA proteins into seven distinctive groups based on specific motifs and domains, and considering all available sequence information from different organisms. However, groups 1, 2, and 3 are considered the major families containing most of the LEA proteins (Amara et al. 2014). *Artemia franciscana* is the only known animal species that expresses LEA proteins from groups 1, 3, and 6 (Hand and Menze 2015).

Table 1: LEA proteins found in the brine shrimp *Artemia franciscana*

<table>
<thead>
<tr>
<th>Protein</th>
<th>Group</th>
<th>Location</th>
<th>Number of amino acids</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfrLEA1.1</td>
<td>1</td>
<td>Cytoplasm</td>
<td>182</td>
<td>Sharon et al. 2009</td>
</tr>
<tr>
<td>AfrLEA1.3</td>
<td>1</td>
<td>Mitochondria</td>
<td>197</td>
<td>Warner et al. 2010</td>
</tr>
<tr>
<td>AfrLEA1</td>
<td>3</td>
<td>Cytoplasm</td>
<td>357</td>
<td>Hand et al. 2007</td>
</tr>
<tr>
<td>AfrLEA2</td>
<td>3</td>
<td>Cytoplasm</td>
<td>364</td>
<td>Hand et al. 2007</td>
</tr>
<tr>
<td>AfrLEA3m</td>
<td>3</td>
<td>Mitochondria</td>
<td>307</td>
<td>Menze et al. 2009</td>
</tr>
<tr>
<td>AfrLEA6</td>
<td>6</td>
<td>Cytoplasm</td>
<td>257</td>
<td>Wu et al. 2011</td>
</tr>
</tbody>
</table>

Group 1 LEA proteins were originally represented by the D-19 and D-132 proteins from cotton seeds and contain an internal 20-mer sequence that may be present in several copies arranged in tandem, from one to four in plants, and up to eight in other
organisms (Amara et al. 2014). Several variants of a group 1 protein are found in *A. franciscana*. LEA-1a, LEA-1b, and LEA-1c, were the first of this group found in an animal. Group 1 proteins tend to be acidic and very hydrophilic (Sharon et al. 2009). Several other members of group 1 LEA proteins have been described later such as the AfrLEA1.1 and AfrLEA1.3 (Marunde et al. 2013).

Group 3 LEA proteins are characterized by a repeating motif of 11 amino acids. Differences in the molecular weight in this group of proteins are usually a consequence of the number of repetitions of this 11-mer motif (Dure 1993). Several members of group 3 LEA proteins have been described in *A. franciscana*. The two proteins AfrLEA1 and AfrLEA2 are cytosolic and hydrophilic proteins of approximately 39 kDa with repeating motifs of 32 and 14 amino acid residues respectively (Hand et al. 2007, Boswell et al. 2014). AfrLEA1 and AfrLEA2 are prominent in quiescent, diapause, and post-diapause embryos of *Artemia*, but not presence in larvae. Another group 3 LEA protein, AfrLEA3m, is enriched in α-helices, contains repeated motifs, and has a mitochondrial pre-sequence (Menze et al. 2009, Boswell et al. 2014). The mRNA of AfrLEA3m is more abundant in diapause-destined embryos than in swimming larvae and adult (Menze et al. 2009).

Group 6 LEA proteins are also termed seed maturation proteins (SMP) and have the potential to restore cellular functions following desiccation by dissociating protein aggregation during rehydration (Boucher et al. 2010). A LEA6 protein from *Artemia franciscana* (AfrLEA6) that has recently been identified, exhibits strong sequence homologies to SMP in plants (Wu et al. 2011, Hand and Menze 2015). AfrLEA6 is less hydrophilic than groups 1 and 3 LEA proteins, which is a characteristic of SMP like MtPM25 (Boucher et al. 2010). It is assumed that AfrLEA6 has an important role in
improving long-term desiccation tolerance in animal cells as suggested for SMP’s in plants (Chatelain et al. 2012).

**Subcellular localization of LEA proteins**

The presence of multiple LEA proteins in a single organism suggests different subcellular targets of LEA proteins to protect vital cellular components from damage exerted by desiccation. Localization of LEA proteins to both the cytoplasm and subcellular organelles has now been documented in several plant (Amara et al. 2014) and animal species (Hand et al. 2011). In plants, LEA proteins are localized to the cytoplasm, nucleus, mitochondrion, chloroplast, endoplasmic reticulum, vacuole, peroxisome, and the plasma membrane (Tunnacliffe and Wise 2007). The group 3 LEA protein, AfrLEA3m from *A. franciscana* is the first protein from an animal species reported to be targeted to the mitochondria (Menze et al. 2009). This group 3 LEA protein is composed of 307 amino acids and contains a 29 amino acid pre-sequence at the N-terminus. Group 1 LEA proteins are found in the cytosol and mitochondria of *Artemia* cyst (Warner et al. 2010). Confocal microscopy revealed that a construct composed of green fluorescent protein (GFP) and AfrLEA1.3 accumulated in the mitochondria, while AfrLEA1.1-GFP was found in the cytoplasm (Marunde et al. 2013). The LEA proteins from the bdelloid rotifer *Adineta ricciae*, ArLEA1A and ArLEA1B, have a hydrophobic N-terminal region as well as a putative endoplasmic reticulum retention signal (the amino acid sequence ATEL) at the C-terminus. Therefore, both proteins are likely targeted to, or transported through, the endoplasmic reticulum, although this conclusion has yet to be supported by experimental evidence (Pouchkina-Stantcheva et al. 2007).
Structural and biochemical properties of LEA proteins

Most of the biochemical properties of LEA proteins have been proposed based on their amino acid composition and hydrophilic nature (Amara et al. 2014). Although significant similarities have not been found between the members of the different LEA groups, a unifying and outstanding feature of most of these proteins is their high hydrophilicity and glycine content, and a lack or underrepresentation of cysteine and tryptophan residues (Baker et al. 1988). The high hydrophilicity of LEA proteins might be the reason for their lack of defined secondary structure in the hydrated state. In aqueous solution, most of the LEA proteins mainly exist as randomly coiled proteins. Although structure modeling and structure prediction programs suggest that at least some LEA proteins contain defined secondary structure (Close 1996), all hydrophilic LEA proteins studied experimentally have revealed a high degree of random coil structure in solution. This classifies them as intrinsically unstructured proteins (Kushwaha et al. 2013). Similarly, the ability of LEA proteins to remain soluble at elevated temperatures can be attributed to their hydrophilic nature. Heat-induce aggregation of proteins results from partial denaturation and association through exposed hydrophobic regions, something that cannot occur in hydrophilic and natively unfolded proteins (Tunncliffe and Wise 2007).

The molecular size of LEA proteins ranges from 5 to 77 kDa among most groups. They can be acidic, neutral or basic. Group 1 proteins are acidic to neutral, group 2 proteins comprise with different isoelectric points, and groups 3 are neutral to basic (Shih et al. 2008). The anomalous movement of LEA proteins in SDS-PAGE is probably caused by reduced interactions between SDS and charged amino acid residues (Gentile et al. 2002). LEA proteins from groups 1, 2, and 3 are predicted to be at least
50% unfolded (Tompa 2002). Therefore, futile attempts to crystallize purified LEA proteins for X-ray crystallography are not surprising (McCubbin 1985).

Nevertheless, some LEA proteins do show some secondary structure motifs and structural elements are in equilibrium with unstructured states. Many natively unfolded proteins are known to undergo increased folding under some conditions, usually when they bind to a partner molecule or cation (Uversky et al. 2000). Environmental conditions can also affect folding and several LEA proteins become more structured when water is removed (Tunnacliffe and Wise 2007). For animal LEA proteins, Tunnacliffe’s group demonstrated this phenomenon by using Fourier-transform infrared (FTIR) spectroscopy. FTIR spectroscopy allows for the assessment of protein secondary structure in the dry state by using the profile of the amide-I band, which provides information on the relative contributions of α-helix, β-sheet, and turn structures (Goyal et al. 2003). Hand et al. found that desiccation of AfrLEA2, a member of group 3 proteins from *Artemia franciscana*, caused an increase in α-helix content from 4% in solution to 46% in the dried state. Similarly, AfrLEA3m which was predominantly disordered in solution, adopted a more α-helical structure after drying. However, AfrLEA3m possessed a greater percentage of β-sheet in the dry state compared to AfrLEA2, which could explain the lower α-helix content in AfrLEA3m (Boswell et al. 2014, Hand and Menze 2015). The propensity of some LEA proteins to gain structure under some conditions may be a general property of these proteins, and may have important functional implications in their physiological roles (Olvera-Carrillo et al. 2011).
Functions of LEA proteins

Several studies, both in vitro and in vivo, showed a correlation between expression of LEA proteins and stress resistance (Hand et al. 2011). Many studies demonstrated that introduction of LEA proteins into plants and microorganisms results in an enhanced stress tolerance (Shih et al. 2008). Transgenic approaches have shown that overexpression of LEA proteins from different species in Arabidopsis, tobacco, rice, wheat, maize, lettuce, or cabbage conferred improved abiotic stress resistance (Amara et al. 2014). However, the exact molecular functions of LEA proteins are still unclear and LEA proteins have been suggested to act as protein and membrane protectants, cell membrane stabilizers, hydration buffers, antioxidants, organic glass formers and ion chelators (Tunnacliffe and Wise 2007).

Protein-protein interactions

LEA proteins have the potential to protect target proteins from inactivation and aggregation during water stress. A role in protein stabilization is supported by the fact that some LEA proteins preserve enzyme activity in vitro during water stress (Reyes et al. 2005). Many proteins, including the enzymes citrate synthase and lactate dehydrogenase, form insoluble aggregates when dried or frozen, but aggregation is reduced in the presence of LEA proteins from groups 1, 2, and 3 (Amara et al. 2014). Group 2 proteins also prevent protein aggregation from heat stress (Kovacs et al. 2008). Due to their hydrophilic, unstructured nature, LEA proteins themselves are not vulnerable to aggregation on desiccation, freezing, or boiling (Tunnacliffe et al. 2010). When the enzyme phosphofructokinase was dried in the presence of AfrLEA2 plus 100 mM trehalose, 98% of control (non-dried) activity was preserved, and 103% of control activity remained intact in the presence of AfrLEA3m plus 100 mM trehalose (Boswell
et al. 2014). A group 1 protein from the brine shrimp, AfrLEA1.3 preserved mitochondrial function and improved viability of transgenic Drosophila melanogaster Kc167 cells during freeze-thawing, drying, and hyperosmotic stress. The protection conferred by AfrLEA1.3 is interesting because it worked during moderate water stress, condition in which cellular water content is above 20% and LEA protein usually does not form α-helical structure. This provides an example that folding is not a prerequisite for activity of LEA proteins (Marunde et al. 2013, MacRae 2016).

It has been proposed that some LEA proteins may exhibit a “molecular shield” activity. In the increasing crowded environment of the dehydrating cytoplasm, LEA proteins could decrease the interaction between partially denatured polypeptides and avoid their aggregation. The shield proteins might also have a space-filling role and help to prevent the cell from collapsing as water is lost (Tunnacliffe et al. 2005). Another functional hypothesis is the chaperone activity (Kovacs et al. 2008). The anti-aggregation activity of LEA proteins resembles a molecular chaperone with “holding” properties, which function in the cell would be to stabilize protein species in a partially unfolded state, preventing aggregation while the stress lasts. In contrary to classical chaperones which require ATP, LEA proteins resemble “holding” chaperone activity without the requirement for ATP. However, LEA proteins are unique in that they lack defined secondary structure and do not form transient complexes with their client proteins through hydrophobic surfaces, as they are hydrophilic (Reyes et al. 2005).

**LEA-membrane interaction**

Maintaining integrity of the cell and organelles is crucial during desiccation. In organisms that accumulate compatible osmolytes such as trehalose, some LEA proteins may contribute to the H-bond network and protect membranes in the dry state (Hoekstra
et al. 2001). Since LEA proteins are highly hydrophilic, interactions with cellular membranes under hydrated condition are not expected, but interactions cannot be excluded through amphipathic α-helices, motifs that containing both hydrophilic and hydrophobic regions, in a dehydrating cell (Amara et al. 2014). Hand et al. observed that cells loaded with trehalose and expressing AfrLEA2 or AfrLEA3m showed 98% membrane integrity compared with 0% intact membranes for control cells without LEA proteins or trehalose. Even without intracellular trehalose, AfrLEA3m conferred 94% protection based on membrane integrity (Li et al. 2012). By using FTIR, Pouchkina-Stantcheva et al. (2007) showed that a group 3 LEA protein from a bdelloid rotifer (ArLEA1B) interacts with dried liposomes. Another group 3 mitochondrial LEA protein (LEAM) from pea is able to interact with membranes to afford protection in the dry state. The interaction between LEAM and phospholipids and the protective effect of LEAM was demonstrated by differential scanning colorimetry using a liposome desiccation assay (Tolleter et al. 2007).

**Biological glass formation**

Sugar glass is “an amorphous metastable state that resembles a solid, brittle material, but with retention of the disorder and physical properties of a liquid. In the glassy state, the rates of molecular diffusion and chemical reactions are greatly reduced” (Hoekstra et al. 2001). In a desiccating cell, when the water content falls below 10% on a dry weight basis, the cytoplasm vitrifies and enters into the “glassy state” (Buitink and Leprince 2004). In plants, the formation of intracellular glasses is indispensable for survival in the dry state (seeds and pollens). LEA proteins accumulate to high levels in seeds (2-4% of the water soluble proteome) (Roberts et al. 1993), and they increase the density of the sugar glasses by strengthening the hydrogen-bonding of the sucrose/LEA mixture (Buitink and Leprince 2004). The LEA proteins of *Artemia*
have the potential to protect proteins from drying-induced aggregation by forming glasses with trehalose, an abundant cyst sugar (Sharon et al. 2009, Warner et al. 2010, Hand et al. 2011, Toxopeus et al. 2014). Thus a potential role of LEA proteins is their contribution to the formation of biological glasses.

**Hydration buffer**

Another suggested function for LEA proteins is to serve as a “hydration buffer”, whereby unstructured hydrophilic proteins bind greater numbers of water molecules than does a typical globular protein (Mouillon et al. 2006, Hand et al. 2011). LEA proteins might act as hydration buffers, slowing down the rate of water loss during dehydration, osmotic or freezing stress (Garay-Arroyo et al. 2000). Using a knockout mutant of *Arabidopsis*, whose seed exhibited premature dehydration, a role for group 1 LEA protein Atm6 as hydration buffer was proposed (Manfre et al. 2006).

**Applications of LEA proteins in biotechnology**

Several studies showed the positive relationship between transgenic LEA proteins and stress tolerance in plants and animals. In general, the phenotypes of transgenic organisms expressing LEA proteins was enhanced stress endurance, often related to desiccation or salt stress. Most studies reported enhanced growth rates and reduced wilting of the aerial parts of plants under stress in laboratory conditions and in some field trials, demonstrating a real potential of LEA proteins in engineering crops more tolerant to water stress (Leprince and Buitink 2010). In addition to agronomical purposes, LEA proteins could be useful for other biotechnological applications in relation to their capacity to prevent aggregation of proteins. The anti-aggregation properties of group 3 LEA protein, AavLEA1 have been applied to reduce the formation of *in vivo* aggregation. The mammalian cells co-expressing AavLEA1 with
aggregation-prone proteins demonstrated substantially reduced protein aggregation linked to neurodegenerative diseases (Chakrabortee et al. 2010). Finally, LEA proteins plus trehalose might proof useful in the development of techniques to preserve cells and tissues in dried condition at room temperature.

**Objectives**

Although several studies have described the effects of expressing LEA proteins from *Artemia* in the *Drosophila* and mammalian cells, all reports to date worked on single LEA proteins. On the other hand, the presence of multiple LEA proteins in anhydrobiotic cyst of *A. franciscana* suggests that some LEA proteins, if not all, might act together during water stress. Therefore, to better understand the function of the multiple proteins in arthropod cells, I aimed to express combinations of two AfrLEA proteins in cell lines derived from the fruit fly, *Drosophila melanogaster* (Kc167 and S2R+). In order to transcribe two LEA proteins concurrently on the same mRNA transcript, I used a novel pAc5-STABLE2-Neo vector developed by Gonzalez et al. (2011). In this multicistrionic vector, the two transgenes were separated from each other by a T2A peptide sequence (EGRGSLTCLGDVEENPGP) derived from the insect virus *Thosea asigna* (Gonzalez et al. 2011). The self-cleaving 2A peptides were first discovered in picornaviruses, are short and produce equimolar levels of multiple proteins from the same mRNA (Kim et al. 2011). The cleavage occurs between the glycine and proline residues found on the C-terminus meaning the upstream protein will have a few additional residues added to the end, while the downstream protein will start with proline (de Felipe 2004). The principal goal of this experiment is to express pairs of LEA proteins in the *Drosophila* cells, however, the whole project can be divided into three sub-sections:
Expression of fluorescent tagged LEA proteins

Fluorescent proteins are easily imaged reporters extensively used in molecular and cell biology. When a protein is tagged to a fluorescent reporter, interactions between fluorescent proteins can undesirably disturb targeting or function (Shaner et al. 2004). Fluorescent proteins have been used as tools in numerous applications including as markers to track and quantify individual protein (Lippincott-Schwartz and Patterson 2003). The fluorescent protein mCherry at the first position of the pAc5-STABLE2-Neo vector, is a red light emitting monomer which matures extremely rapidly, making it possible to see results very soon after activating transcription. It is highly photostable and resistant to photobleaching (Shaner et al. 2004). The fluorescent marker GFP, at second position of the vector, is generally non-toxic and can be expressed to high levels in different organisms with minor effects on their physiology (Lippincott-Schwartz and Patterson 2003). Thus, my first objective was to express combinations of two LEA proteins one tagged with mCherry and another fused to GFP.

Effect of ribosome recognition sequences on expression level of proteins

The consensus sequence for ribosome recognition in vertebrates is commonly known as a Kozak sequence according to its discoverer, Marilyn Kozak. The Kozak sequence was originally defined as ACC AUG G following an analysis of the effects of single mutations surrounding the initiation codon (AUG) on translation of the prepro-insulin gene (Kozak 1986). Subsequent mutagenesis studies and a survey of 699 vertebrate mRNAs extended the consensus sequence for translation initiation to GCC GCC ACC AUG G (Kozak 1987). However, in Drosophila the consensus sequence (GCC AAC AUG) flanking translational start site differs slightly from the sequence in vertebrates (Cavener 1987). Here, I tested whether these two different ribosome
recognition sequences have any role on the expression level of mCherry tagged AfrLEA1.3 proteins in the *Drosophila* cells.

**Expression of non-fluorescent tagged LEA proteins**

The final objective was to express two LEA proteins with small, non-fluorescent DDK tag and 6X His tag. The 6X His tag is the most commonly used tag for purification and identification of target proteins through immunoblotting. It comprises 6-14 histidine and is typically fused to the N- or C-terminal end of a target protein (Terpe 2003). Due to their hydrophilic and flexible nature, His tag can often increase the solubility of target proteins and only rarely interfere with protein’s function (Hochuli et al. 1988). On the other hand, the sequence GAT TAC AAG GAT GAC GAC GAT AAG is known as a DDK tag since it encodes amino acids DYKDDDDK (D-aspartic acid, K-lysine). DDK is the same as FLAG® which is a registered trademark of Sigma Aldrich. The FLAG-tag system utilizes a short, hydrophilic 8-amino acid peptide that is fused to the protein of interest (Hopp et al. 1988). Finally, I aimed to express DDK tagged LEA proteins at first position and 6X His tagged LEA proteins at second position of the pAc5-STABLE2-Neo vector.

In the current study, I will demonstrate that by using a multicistronic vector, simultaneous expression of group 3 and 6 LEA proteins from the *Artemia franciscana* is possible in the *Drosophila melanogaster* Ka167 cells. I will also show that in comparison to the vertebrate Kozak sequence, *Drosophila* consensus ribosome recognition sequence improves expression level of fluorescent tagged LEA proteins in the *Drosophila* cells. The difficulty in plasmids construction due to vector recombination, and problem in identification of His tagged LEA protein will also be a matter of discussion.
MATERIALS AND METHODS

Vector selection

In order to express two LEA proteins in *Drosophila* cells by single transfection, the multicistronic vector pAc5-STABLE2-Neo (Addgene, Cambridge, MA) was used. The expression of multiple proteins in *Drosophila* cells is driven by the Actin5C promoter. In the vector, the neomycin resistance gene is separated from green fluorescent protein (GFP) by a T2A sequence, peptide found in viral polyproteins and serve as signals to guide apparent self-cleavage of the polyprotein into individual proteins (Ryan et al. 1991). It also contains a sequence encoding a FLAG epitope tagged version of mCherry fluorescent protein. The mCherry and GFP are separated by an additional dT2A sequence, encoded by a degenerate nucleotide sequence to prevent vector recombination (Fig. 1).

![Schematic representation of the pAc5-STABLE2-Neo vector](image)

Figure 1. Schematic representation of the pAc5-STABLE2-Neo vector (modified form Gonzalez et al. 2011).

Primer designing

Upon designing, all the primers were obtained from Integrated DNA Technologies, Coralville, IA. The online program OligoAnalyzer Tool (www.idtdna.com/calc/ analyzer) was used to design all primers. Another online program the Biology
Workbench (http://workbench.sdsc.edu/) was used to identify the longest open reading frame (ORF), a span of genetic material that can be read by the ribosome to produce a single protein. Moreover, the online programs NEB cutter (http://nc2.neb.com/NEBcutter2/) and Webcutter 2.0 (http://rna.lundberg.gu.se/cutter2/) developed by Max Heiman at Yale University were used to examine whether the target genes have any cutting site for the used restriction enzymes. The general principles to design the primers were as follow:

- The length of the primers was between 17-28 bases with some exceptions.
- Guanine and Cytosine (GC) content were at least 50% of total bases.
- Melting temperature of the primers were between 60-80°C.
- Annealing temperature of the primers were around 72°C.
- 3-5 extra bases were added at the 5’ ends outside of the restriction enzymes cutting sites to facilitate enzymes activities.

**Primers for fluorescent tagged LEA proteins**

The primers to clone genes at position 1 in the vector were designed with the two restriction enzymes cutting sites KpnI (GGT ACC) and EcoRI (GAA TTC) to construct chimeric protein composed of the protein of interest and the fluorescent protein mCherry. Primers for insertion of a second gene at position 2 were designed with the two restriction enzymes cutting sites EcoRV (GAT ATC) and XbaI (TCT AGA) to incorporate the gene upstream of GFP.
Table 2: The primers used for cloning of fluorescent marked LEA proteins in the STABLE-2 vector.

<table>
<thead>
<tr>
<th>STABLE-2 vector</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; position mCherry tagged</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; position GFP tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position 1+ Position 2</td>
<td>Forward primer</td>
<td>Reverse primer</td>
</tr>
<tr>
<td>LEA3m + LEA 6</td>
<td>5'- GC <strong>GGT ACC</strong> ATG TTG TCC AAG CGT TTA ATT AAA AGC TTA AGC TGT G -3'</td>
<td>5'- CGGCC <strong>GAA TTC</strong> TCT TCT ATG AGC TCC AGA TGC CGA C -3'</td>
</tr>
<tr>
<td>LEA 1.3 + LEA 6</td>
<td>5'- GT <strong>GGT ACC</strong> GCC GCC ACC ATG GAG CTC TCT TCC -3'</td>
<td>5'- TTAGT <strong>GAA TTC</strong> TCC TCC GCC CTT CTG CCG GCC -3'</td>
</tr>
<tr>
<td>LEA 1 + LEA 6</td>
<td>5'- GC <strong>GGT ACC</strong> ATG GCT GAG CCA GAG GAA CCT CC -3'</td>
<td>5'- CAGCC <strong>GAA TTC</strong> TGC GCC CTT TAT TCG ATC TGC AG -3'</td>
</tr>
<tr>
<td>LEA 2 + LEA 6</td>
<td>5'- GC <strong>GGT ACC</strong> ATG CCA AAA GCA GCA GCT AAA GGT ATT GGG -3'</td>
<td>5'- CAGCC <strong>GAA TTC</strong> TTC TCC AGG GTT TTC TTT TGG AAA CCG TTC TTG ATG AAG TAT TAT CCT ATC TCC -3'</td>
</tr>
<tr>
<td>LEA 3m + LEA 1.1</td>
<td>5'- GC <strong>GGT ACC</strong> ATG TTG TCC AAG CGT TTA ATT AAA AGC TTA AGC TGT G -3'</td>
<td>5'- CAGCC <strong>GAA TTC</strong> TCT TCT ATG AGC TCC AGA TGC CGA C -3'</td>
</tr>
<tr>
<td>LEA 2 + LEA 1.1</td>
<td>5'- GC <strong>GGT ACC</strong> ATG CCA AAA GCA GCA GCT AAA GGT ATT GGG -3'</td>
<td>5'- CAGCC <strong>GAA TTC</strong> TTC TCC AGG GTT TTC TTT TGG AAA CCG TTC TTG ATG AAG TAT TAT CCT ATC TCC -3'</td>
</tr>
</tbody>
</table>

**Bold letters:** sequences for the restriction enzymes
Primers for LEA1.3 with ribosome recognition sequences

To improve the expression level of mCherry tagged LEA1.3 proteins in Kc167 cells, the primers for LEA1.3 were designed with ribosome recognition sequences for vertebrate (Kozak 1986) and Drosophila (Cavener 1987). The ribosome recognition sequences were added on the forward primers of LEA1.3.

Table 3: The primers used for cloning of mCherry marked LEA1.3 proteins in the STABLE-2 vector with and without ribosome recognition sequences.

<table>
<thead>
<tr>
<th>Forward primers</th>
<th>Reverse primers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LEA1.3 control (no ribosome recognition sequence)</strong></td>
<td><strong>LEA1.3 with Drosophila ribosome recognition sequence</strong></td>
</tr>
<tr>
<td>5’- GC <strong>GGT</strong> ACC ATG GAG CTC TCT TCC AGT AAA CTG AAC CGG TC -3'</td>
<td>5’- GC <strong>GGT</strong> ACC GCC AAC ATG GAG CTC TCT TCC AGT AAA CTG AAC CGG TC -3'</td>
</tr>
<tr>
<td>5’- CAGCC <strong>GAA</strong> TTC ATG ATG ATG GTG GTG TCC TCC GCC C -3'</td>
<td>5’- CAGCC <strong>GAA</strong> TTC ATG ATG ATG GTG GTG TCC TCC GCC C -3'</td>
</tr>
<tr>
<td><strong>LEA1.3 with vertebrate Kozak sequence</strong></td>
<td><strong>LEA1.3 with vertebrate Kozak sequence</strong></td>
</tr>
<tr>
<td>5’- GC <strong>GGT</strong> ACC GCC GCC ACC ATG GAG CTC TCT TCC AGT AAA CTG AAC CGG -3'</td>
<td>5’- GC <strong>GGT</strong> ACC GCC GCC ACC ATG GAG CTC TCT TCC AGT AAA CTG AAC CGG -3'</td>
</tr>
<tr>
<td>5’- CAGCC <strong>GAA</strong> TTC ATG ATG ATG GTG GTG TCC TCC GCC C -3'</td>
<td>5’- CAGCC <strong>GAA</strong> TTC ATG ATG ATG GTG GTG TCC TCC GCC C -3'</td>
</tr>
</tbody>
</table>

**Bold letters:** sequences for the restriction enzymes

Primers for non-fluorescent tagged LEA proteins

To express LEA proteins without fluorescent reporter proteins, the primers for genes to be inserted at position 1 in the vector were designed with a DDK tag sequence (CTT ATC GTC GTC GTC ATC CTT GTA ATC) and restriction enzymes KpnI (GGT ACC) and NotI (GC GGC CGC). For the non-fluorescent second protein, the primers were designed with 6X His tag sequence and restriction enzymes EcoRV (GAT ATC) and HindIII (AAG CTT). The sequence for DDK and 6X His tags were added on the reverse primers to express these tags on the C-terminus of the LEA proteins.
Table 4: The primers used for PCR amplification and cloning of LEA proteins in the STABLE-2 vector with 6X His and DDK tags.

<table>
<thead>
<tr>
<th>STABLE-2 vector</th>
<th>1st position DDK tagged</th>
<th>2nd position 6X His tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position 1+</td>
<td>Forward primer</td>
<td>Reverse primer</td>
</tr>
<tr>
<td>Position 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEA1+ LEA 6</td>
<td>5'- GC GGT ACC ATG GCT GCA GGT AAG CCA GAA CCT CC-3'</td>
<td>5'- TAT GC GGC CGC CTT ATC GTC ATC CGC ATC TGG TGC GTA ATC TGG GCC CCG CTG TAT TCG ATC -3'</td>
</tr>
<tr>
<td>LEA2+ LEA 6</td>
<td>5'- GC GGT ACC ATG CCA AAA GCA GCA GCT AAA GGT ATT GGG -3'</td>
<td>5'- CGC GC GGC CGC CTT ATC GTC GTG TGC ATC GTG GTA ATC TTG AGG GTT TTC TTT TGG -3'</td>
</tr>
<tr>
<td>LEA3m+ LEA 6</td>
<td>5'- GC GGT ACC ATG TTG TCC AAG CGT TTA ATT AAA AGC TTA AGC TGT G-3'</td>
<td>5'- TAT GC GGC CGC CTT ATC GTC GTG TGC ATC ATC GTA ATC TCT TTC ATG AGC TCC AGA -3'</td>
</tr>
<tr>
<td>LEA1.3+ LEA 6</td>
<td>5'- CGCGC GGT ACC ATG GAG CTC TCT TCC AGT AAA CTG AAC CGG TC -3'</td>
<td>5'- TAT GC GGC CGC CTT ATC GTC GTG TGC ATC ATC GTA ATC TCC TCC GCC CTT CT -3'</td>
</tr>
</tbody>
</table>

**Bold letters**: sequences for the restriction enzymes

**Polymerase chain reaction (PCR)**

To amplify DNA through PCR the enzyme Q5 High-Fidelity DNA polymerase and the corresponding protocol (New England BioLabs, Ipswich, MA) was used. The
reactions set up for PCR are shown in Table 5. PCR products were run on 1% agarose gel for 1 hour at 120 volts. The TAE buffer (40 mM Tris base, 20 mM acetic acid, and 1 mM EDTA) was used to make and run the agarose gel. A 2-log DNA ladder with mass ranges from 0.1-10 kb was used as a standard to identify the correct PCR products. NucleoSpin Gel and PCR cleanup kit (Macherey-Nagel, Bethlehem, PA) was used to purify DNA from the agarose gels. Purified DNA was quantified through Epoch microplate spectrophotometer (BioTeK Instruments, Winooski, VT) and samples were preserved at -20°C.

Table 5: The reagents used to amplify DNA through PCR.

<table>
<thead>
<tr>
<th>Ingredients</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction buffer (5X)</td>
<td>10 µL</td>
</tr>
<tr>
<td>Deoxy-nucleotide triphosphate (dNTP)</td>
<td>1 µL</td>
</tr>
<tr>
<td>Primers mixture</td>
<td>1 µL</td>
</tr>
<tr>
<td>Template DNA</td>
<td>1 µL</td>
</tr>
<tr>
<td>DNA polymerase</td>
<td>0.5 µL</td>
</tr>
<tr>
<td>dH₂O</td>
<td>36.5 µL</td>
</tr>
<tr>
<td>Total volume</td>
<td>50 µL</td>
</tr>
</tbody>
</table>

**Restriction enzyme digestion**

After purification, both target DNA and the STABLE-2 vector were digested with the same restriction enzymes at 37°C for 1 hour. For fluorescent protein tagged constructs, DNA was digested with restriction enzymes KpnI and EcoRI for genes to be inserted at position 1, and with EcoRV and XbaI for genes to be incorporated at position 2 of the vector. Upon developing mCherry and GFP tagged constructs, the vector STABLE-2 was digested to remove the fluorescent proteins and to insert LEA genes with DDK and 6X His tags. To clone DDK tagged LEA protein at first position, both the genes and the vector were digested with restriction enzymes KpnI and NotI;
and to insert 6X His tagged LEA gene at second position, the inserts and vector were digested with restriction enzymes EcoRV and HindIII. All the restriction enzymes used in this experiment were purchased from New England BioLabs, Ipswick, MA.

<table>
<thead>
<tr>
<th>Ingredients</th>
<th>Vector digestion</th>
<th>Insert digestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutsmart buffer (10X)</td>
<td>10 µL</td>
<td>5 µL</td>
</tr>
<tr>
<td>DNA</td>
<td>2.5 µg</td>
<td>1 µg</td>
</tr>
<tr>
<td>Restriction enzyme 1</td>
<td>2 µL</td>
<td>1 µL</td>
</tr>
<tr>
<td>Restriction enzyme 2</td>
<td>2 µL</td>
<td>1 µL</td>
</tr>
<tr>
<td>De-ionized H2O</td>
<td>varied</td>
<td>varied</td>
</tr>
<tr>
<td>Total volume</td>
<td>100 µL</td>
<td>50 µL</td>
</tr>
</tbody>
</table>

**Vector dephosphorylation**

Digested DNA usually possesses a 5' phosphate group that is required for ligation. In order to prevent self-ligation, the 5' phosphate can be removed prior to ligation. I accomplished dephosphorylation by adding 0.5 µL calf intestinal alkaline phosphatase (CIP) to the digestion reaction for the vector. The reactions for vector dephosphorylation were run for 1 hour at 37°C. Following dephosphorylation, digested vectors were run on 1% agarose gel and purified using NucleoSpin PCR and Gel clean-up kit (Macherey-Nagel, Bethlehem, PA). On the other hand, digested inserts were not run on agarose gel rather purified using the same clean-up kit.

**Ligating the vector and LEA genes**

The enzyme T4 DNA Ligase (New England BioLabs, Ipswick, MA) was used to ligate the digested vector and the inserts. Both room temperature (one hour for sticky ends and two hours for blunt ends) and 16°C (16 hours) were used for ligation. While reactions were set up at 16°C, PCR machine was used to maintain the correct
temperature. Ligated plasmids were either frozen at -20°C or immediately used to transform chemically competent *E. coli* cells.

Table 7: The reagents used to set up ligation reactions.

<table>
<thead>
<tr>
<th>Ingredients</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vector</td>
<td>3-4 µL (&gt;120 ng)</td>
</tr>
<tr>
<td>Insert</td>
<td>1-2 µL (&gt;50 ng)</td>
</tr>
<tr>
<td>T4 DNA Ligase buffer</td>
<td>1 µL</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>1 µL</td>
</tr>
<tr>
<td>dH₂O</td>
<td>varied</td>
</tr>
<tr>
<td>Total volume</td>
<td>10 µL</td>
</tr>
</tbody>
</table>

**Transformation in *E. coli* cells**

One shot TOP 10 chemically competent *E. coli* cells and the manufacturer’s guideline (Thermo-Fisher Scientific, Carlsbad, CA) was used for transformation. Briefly, one vial of the competent cells was removed from -80°C and placed on ice to thaw frozen bacteria. Then 5 µL of ligated products were added to the bacteria and mixed by gentle tipping. The mixture of bacteria and plasmids were incubated on ice for 30 minutes. After incubation on ice, heat shock was given at 42°C for exactly 30 seconds in a pre-heated water bath. Following heat shock, the bacteria were kept on ice for 2 minutes and 250 µL of S.O.C media (Thermo-Fisher Scientific, Carlsbad, CA) was added. The bacteria were cultured for 1 hour at 37°C on a shaker rotating at 225 rpm. Finally, bacteria were spread on ampicillin (100 µg/mL) containing LB (Luria Bertani) plates and grown in incubator overnight at 37°C. I also used α-select chemically competent bacteria cells (Bioline, Taunton, MA) for cloning of non-fluorescent tagged LEA proteins because these cells show reduced recombination of cloned DNA.
Clone selection and plasmids DNA purification

Following the day of transformation, growth of bacteria on LB plates was observed and 3-6 colonies were picked and grown in 5 mL of liquid LB medium containing 100 µg/mL ampicillin. The bacteria were cultured overnight (12-16 hours) at 37°C on a shaker rotating at 225 rpm. The next day, 500 µL of overnight grown bacteria were mixed with 250 µL 3X glycerol solution (65% glycerol, 0.1 M MgSO₄, 0.2 M Tris·Cl, pH 8) in cryopreservation vials and preserved at -80°C for long term storage. The remaining 4.5 mL of bacteria were used for plasmid isolation using the NucleoSpin plasmid isolation kit and the corresponding protocol (Macherey-Nagel, Bethlehem, PA).

Verification of genes in the plasmids

Purified plasmids were quantified with the Epoch microplate spectrophotometer (BioTek Instruments, Winooski, VT) by following the manufacturer’s guideline. The insertion of LEA genes in the plasmids were verified by two methods. One way was by running PCR products on 1% agarose gel where isolated plasmids from cloned bacteria were used as template DNA. Another way to verify success of cloning was digestion of the isolated plasmids by restriction enzymes and comparing plasmids size with the empty vector on 0.8% agarose gels.

Drosophila cell culture

The embryonic cell lines Kc167 and S2R+ of the fruit fly, Drosophila melanogaster were obtained from the Drosophila Genomics Resources Center, Bloomington, IN. Cells were grown in Schneider’s Drosophila medium (Caisson Laboratories, Smithfield, UT) supplemented with 10% fetal bovine serum (FBS) and 1% PSA (MP Biomedicals, Santa Ana, CA) to prevent bacterial and fungal
contaminations. PSA is a triple antibiotic solution of penicillin (10000 IU/mL), streptomycin (10 mg/mL), and amphotericin B (25 µg/mL). Cells were usually cultured on 10 mL petri dish with initial concentration of 2 million/mL and subcultured every 2-3 days. Cells were cultured in a refrigerator cabinet at 25°C.

Transflecting plasmid DNA into *Drosophila* cells

1. The day before transfection, 20 million cells were aliquoted from the regular medium and re-suspended in FBS and PSA free medium, and plated.
2. On the transfection day, two tubes were labeled as A and B. In tube A, 120 µL Grace insect medium and 18 µL Cellfectin were mixed; and in tube B, 120 µL Grace insect medium, 2 µg plasmid DNA, and 2 µL PLUS Reagent were mixed. Both solutions were incubated for 30 minutes at room temperature. All the reagents for transfection were brought from the Thermo-Fisher Scientific, Carlsbad, CA.
3. The solutions of tube A and B were mixed together and incubated for another 30 minutes at room temperature.
4. During the incubation time, previously plated cells were counted and 5 million cells were re-suspended in 2.5 mL FBS and antibiotics free medium and plated on a 6-wells plate.
5. Following incubation, the transfection solution was mixed with the cells by shaking the plate and incubated for 24 hours in cell culture chamber to uptake the plasmids.
6. Transfection medium were removed around 24 hours of transfection, and cells were re-suspended in FBS containing Schneider medium.
7. After 48-72 hours of transfection, G418 antibiotic (Thermo-Fisher Scientific, Carlsbad, CA) was added to the medium at a concentration of 2 mg/mL to select cells that have the neomycin resistance gene.

8. During the selection process, transfected cells were monitored, counted, and medium was changed every 3-4 days.

9. After 3-4 weeks of transfection, Kc167 cell lines started to thrive in G418 medium. However, S2R+ cell lines were sluggish during the selection process and it took around 7-8 weeks to obtain stable S2R+ cell lines.

10. Stable cell lines were maintained in lower concentrated G418 medium (1 mg/mL).

**Visualizing fluorescent tagged LEA proteins**

The day after transfection, expression of mCherry and GFP tagged LEA proteins in *Drosophila* cells were observed under the fluorescence microscope. The cells were monitored under the microscope once in a week during the selection process, even after Western blot confirmation of fluorescent marked LEA proteins. The images of stable cell lines that expressed fluorescent tagged LEA proteins were taken with a fluorescence microscope (Olympus, Model BX50F4, Japan) at 400X magnification.

**Preservation of stable Drosophila cell lines**

When transfected cells started to thrive, a portion of cells were preserved in liquid N2 tank for future use and the remaining cells were continuously cultured in G418 supplemented medium. Cell preservation was done according to the guidelines of the Drosophila Genomic Resource Center, Bloomington, IN. Briefly, the freezing medium was prepared by adding 10% FBS and 10% dimethyl sulfoxide (DMSO) to Schneider’s medium and filtered. Cells were pelleted by centrifugation at 4000 rpm for 10 minutes.
and then re-suspended in the freezing medium at a concentration of 20-30 million/mL. In each cryopreservation vial, 1 mL of suspended cells was transferred and immediately kept at -80°C. After 48-72 hours in the -80°C freezer, the vials were transferred into a liquid N₂ tank for permanent storage. Although I generated both stable Kc167 and S2R+ cell lines, only Kc167 were maintained and used for further experiments because of their faster growth rate compared to S2R+.

**Protein isolation**

For the isolation of total proteins from Kc167 cells, 40 million cells were harvested in conical tubes and pelleted by centrifuging at 4000 rpm for 10 minutes. The supernatants were discarded and cells were re-suspended in 1 mL of phosphate buffer saline (PBS). Cells were centrifuged again for 10 minutes at 4000 rpm to remove PBS. Cells were re-suspended in 200 µL of 1X RIPA sample buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 5 mM EDTA, 1% NP-40 or 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS), and 1 µL of 200 mM PMSF (phenyl-methane-sulfonyl fluoride) that is 200X stock solution. Cells were sonicated on ice for 60 seconds to disrupt membranes and liberate proteins. Then the lysates were centrifuged for 1 hour at 14000 rpm at 4°C to separate dissolve proteins from other cellular components. After centrifugation the supernatant was transferred into a fresh micro tube. Then 120 µL of supernatant was mixed with 120 µL of 2X Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue, 5% β-mercaptoethanol which added fresh). The proteins in the Laemmli buffer were denatured by heating at 95°C for 10 minutes and preserved at -20°C for immunoblotting. Another 50 µL of supernatant were aliquoted to measure total protein concentration by Bradford assay.
**Protein quantification**

The Bradford assay were used to determine total protein concentration in the samples. The standard curve was generated by taking absorbance readings of Bradford reagent (Bio-Rad Laboratories, Hercules, CA) containing 0, 125, 250, 500, 750, 1000, 1500, 2000 µg/mL bovine serum albumin (BSA). The optical density (OD) of the samples were taken at 595 nm using a spectrophotometer (Evolution 300 UV-Vis, Thermo-Fisher Scientific, Carlsbad, CA).

Table 8: Optical density of standard BSA samples in Bradford reagent at 595 nm.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Protein concentration, µg/mL</th>
<th>OD at 595 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2000</td>
<td>1.088</td>
</tr>
<tr>
<td>B</td>
<td>1500</td>
<td>0.998</td>
</tr>
<tr>
<td>C</td>
<td>1000</td>
<td>0.709</td>
</tr>
<tr>
<td>D</td>
<td>750</td>
<td>0.557</td>
</tr>
<tr>
<td>E</td>
<td>500</td>
<td>0.412</td>
</tr>
<tr>
<td>F</td>
<td>250</td>
<td>0.194</td>
</tr>
<tr>
<td>G</td>
<td>125</td>
<td>0.102</td>
</tr>
</tbody>
</table>

**Immunoblotting**

**Preparing SDS-PAGE**

In order to run protein samples, nine welled 0.75 mm thick sodium dodecyl sulfate-polyacrylamide gels were used. To prepare two 10 % resolving gels, 4.1 mL dH2O, 3.3 mL acrylamide/bis (37.5:1) solution, 2.5 mL gel buffer (1.5 M Tris-HCl, pH 8.8) and 0.1 mL of 10% SDS were mixed together and degassed for 15 minutes. Then, 50 µL of 10% fresh ammonium persulfate and 5 µL TEMED solutions were mixed properly with degassed solution and casted. After 45 minutes, 5% stacking gel was prepared by mixing 5.7 mL dH2O, 1.7 mL acrylamide/bis (37.5:1) solution, 2.5 mL gel
buffer (0.5 M Tris-HCl, pH 6.8) and 0.1 mL of 10% SDS. The mixture of the stacking
gel also degassed for 15 minutes. Similar to resolving gel, 50 µL of 10% fresh
ammonium persulfate and 5 µL TEMED solutions were mixed properly with the
degassed solution and casted on top of the resolving gel. After polymerization, gels
were used either immediately or kept in refrigerator at 4°C for up to 5 days.

**Running SDS-PAGE**

Two polymerized gels were placed together in the gel running box (Bio-Rad
Laboratories, Hercules, CA) and half of the box was filled with 1X running buffer (25
mM Tris, 190 mM glycine, 0.1% SDS, pH 8.3). Before loading, all samples and
standards were heated at 95°C for 2 minutes. Then, 30 µL protein samples, 10 µL
Kaleidoscope pre-stained standards (Bio-Rad Laboratories, Hercules, CA) and 30 µL
biotinylated protein ladder (Cell Signaling Technology, Danvers, MA) were loaded in
different wells. The distribution of the samples were monitored by looking on the
migration of pre-stained protein standards. The gels were run at 120 volts until lowest
band (10 kDa) of the Kaleidoscope separated from other bands which took on average
1 hour.

**Transferring proteins on membrane**

After electrophoresis, gels were washed in transfer buffer (25 mM Tris, 190 mM
glycine, 0.1% SDS, 20% methanol, pH 8.3) for 15 minutes on a shaker. The small
amount of SDS in the transfer buffer may give the proteins enough charge to move
unidirectionally towards the anode and in most cases should not denature the protein. I
used nitrocellulose membrane to transfer proteins from the gel for Western blotting. A
sandwich was made by combining a fiber pad with soaking paper, SDS gel, membrane,
soaking paper, and a final fiber pad (bottom to top), to transfer proteins. The sandwich
was placed into the transfer cassette and the tank was filled with transfer buffer and run
for 1 hour at 60 volts. While transferring proteins, the tank was kept on ice to avoid high temperature due to electric current.

**Visualizing transferred proteins on membrane**

To see whether transfer was successful, membranes were stained with Ponceau Red (0.2% w/v Ponceau S, 5% glacial acetic acid) for 5 minutes. Then membranes were washed with water for three times and band of proteins in the samples became visible.

**Blocking membrane**

Blocking buffer was made by dissolving 5% nonfat dry milk powder in TBS-T (20 mM Tris pH 7.5, 150 mM NaCl, 0.1% Tween 20) solution. The membranes were incubated in the blocking buffer for 1 hour at room temperature on a shaker. Blocking the membrane in milk solution prevents unspecific binding of primary and secondary antibodies.

**Incubation with primary antibody**

Rabbit anti-mCherry and anti-GFP primary antibodies (both from Rockland Immunochemicals, Limerick, PA.) were used for detecting fluorescent tagged LEA proteins. Mouse anti-DDK (OriGene, Rockville, MD) and rabbit anti-His (Cell Signaling Technologies, Danvers, MA) antibodies were used for detecting non-fluorescent tagged LEA proteins. The primary antibodies were diluted in the blocking buffer at 1:1000 and membranes were incubated overnight on a shaker at 4°C. The following morning, the membranes were washed with TBS-T for three times (each 5 minutes) prior to incubation with secondary antibody.

**Incubation with secondary antibody**

Anti-rabbit and anti-mouse antibodies for the samples and anti-biotin for the biotinylated proteins standard were used as secondary antibodies (Cell Signaling Technologies, Danvers, MA). Secondary antibodies were also diluted in TBS-T
solution at 1:1000. After incubation in secondary antibody for 1 hour at room temperature, membranes were washed with TBS-T for three times (each 5 minutes).

**Developing X-ray film**

The membranes were incubated in Lumiglow (Cell Signaling Technologies, Danvers, MA) for 1-2 minutes at room temperature. Lumiglow which is a substrate for horseradish peroxidase (HRP) was diluted with deionize water at 1:10. Finally, the membranes were exposed to X-ray film in a dark room for 30-120 seconds and films were developed.
RESULTS

1. Expression of fluorescent tagged LEA proteins

Construction of plasmids

The insertion of the target genes in the STABLE-2 vector was verified in two ways: (1) amplification of the desired genes through PCR using plasmids isolated from bacterial colonies as template DNA and (2) comparing the size of the isolated plasmids with the size of the empty vector on agarose gel using restriction enzyme digestion to confirm LEA inserts.

Cloning of LEA 3m at the first position of the STABLE-2 vector

The first gene inserted at the first cloning site of the vector was the mitochondrial targeted protein LEA3m that belongs to group 3 LEA proteins. Following transformation, plasmids were isolated from five distinct bacterial colonies and PCR carried out by using the isolated plasmids as template DNA. The PCR results confirmed the presence of gene LEA3m in all five isolated plasmids as gel image showed a DNA band of about 921 base pairs (bp) corresponding to the molecular size of LEA3m (Fig. 2). Furthermore, three (# 2, 3, and 4) of the five plasmids showed sizes above the size of the empty vector indicating that LEA3m has been incorporated into these three plasmids. However, two plasmids (# 1 & 5) were smaller in size than the empty vector suggesting that these two plasmids did not incorporate the gene of interest (Fig. 3) despite the positive PCR result. Therefore, plasmids number 2-4 were used for further experiments.
Figure 2: PCR amplification of LEA3m DNA (921 bases) using isolated plasmids from bacterial colonies as template DNA. Lane 1: 2-log DNA ladder, lane 2-6: PCR products of LEA3m from five isolated plasmids. All five plasmids successfully yielded LEA3m DNA when used as template.

Figure 3: Restriction enzyme digestion of the plasmids that yielded LEA3m DNA in PCR. Lane 1: empty vector, lane 2-6: isolated plasmids # 1-5 respectively. The size of the plasmids in lanes 3-5 (# 2-4) were larger than empty vector (lane 1) indicates that these plasmids incorporated the target gene LEA3m. Conversely, smaller size of the plasmids in lanes 2 and 6 (# 1 & 5) than parental vector implies that these two were not the desired plasmids.
Inserting LEA6 at the second cloning site of the vector STABLE-2+LEA3m

After verification of LEA3m at the first position in the plasmid construct, the gene LEA6 was inserted at the second cloning site of the vector. Similar to the cloning of LEA3m, plasmids were isolated from five distinct colonies and PCR carried out for LEA6 by using purified plasmids as template. The gel image of the PCR products suggested that all five plasmids have the target gene LEA6 (Fig. 4). Size comparison of the linearized vectors indicated that four (# 2-5) of the five plasmids that showed positive result in PCR, incorporated the desired gene LEA6 as their size were larger than vector with only LEA 3m (Fig. 5, lane 3). Therefore, I got four plasmids that incorporated LEA3m at first and LEA6 at the second cloning sites. However, restriction enzyme failed to cut one plasmid (Fig. 5, lane 4), so this one was discarded and the four correct plasmids were used for further experiments.

Figure 4: PCR products of LEA6 (771 bases) using isolated plasmids as template DNA. Lane 1: 2-log DNA ladder, lane 2-6: amplification of LEA6 from plasmids number 1-5 respectively. All five selected plasmids contain LEA6 DNA as indicated by agarose gel electrophoresis.
Cloning of LEA1 and LEA2 at the first position of the vector STABLE-2+LEA6

After generating first complete construct (STABLE-2+LEA3m+LEA6), this plasmid was digested with the restriction enzymes KpnI and EcoRI to remove the first gene LEA3m. Upon gel purification, the digested vector having LEA6 at the second position was ligated with either LEA1 or LEA2 at the first position. After transformation, bacteria were grew on ampicillin containing LB plate and five colonies for LEA1 and LEA2 were cultured and plasmids were isolated. The PCR products showed that three plasmids (# 1-3) contained LEA1 but two (# 4 & 5) did not, and all five selected colonies for LEA2 had the gene (Fig. 6). Plasmids that showed positive result in PCR were next digested and the size comparison further suggested insertion of LEA1 and LEA2 in the isolated plasmids (Fig. 7). Thus, the cloning process of another two constructs (LEA1+LEA6) and (LEA2+LEA6) were completed.
Figure 6: Amplification of LEA1 (1074 bases) and LEA2 (1092 bases) through PCR using isolated plasmids as DNA template. Lane 1-5: PCR products of LEA1 from plasmids number 1-5 respectively, lane 6: DNA ladder, lane 7-11: PCR products of LEA2 from plasmids number 1-5 respectively. Among five selected plasmids for LEA1, three contained the gene (lane 1-3). Also, all five isolated plasmids incorporated LEA2 gene as suggested by the PCR products (lane 7-11).

Figure 7: Linearized plasmids that showed positive PCR result for LEA1 and LEA2. Lane 1-3: plasmids with LEA1+LEA6, lane 4: vector with only LEA6, lane 5: empty vector, lane 6-10: plasmids with LEA2+LEA6. All the plasmids that successfully amplified LEA1 and LEA2 in the PCR were larger in size than empty vector (lane 5).
**Insertion of LEA1.3 at the first cloning site of the vector STABLE-2+LEA6**

Similar to developing plasmids with LEA1+LEA6, and LEA2+LEA6, another mitochondrial protein LEA1.3 was inserted at the first position of the vector to generate the LEA1.3+LEA6 construct. Five plasmids were isolated and digested prior to running PCR, and found that the sizes of three (# 3-5) out of five plasmids were larger than empty vector (Fig. 8). Then PCR was carried out by using two isolated plasmids (# 2 and 5) as template DNA and found that the gene LEA1.3 was present in both plasmids (Fig. 9). Thus, construction process of another combination (LEA1.3+LEA6) was completed.

![Figure 8: Single digestion of the plasmids isolated from the bacteria cloned with LEA1.3+LEA6. Lane 1: empty vector, lane 2-6: plasmids purified from colonies # 1-5 respectively. Plasmids in lane 4-6 (# 3-5) were larger in sizes than the size of empty vector (lane 1), but plasmids in lanes 2-3 (# 1-2) were about the same size of the parental plasmid (lane 1).](image-url)
Figure 9: PCR products of LEA 1.3 (600 bases) using isolated plasmids as template DNA. Lane 1: DNA ladder, lane 2: PCR product using plasmid # 2 as template, lane 3: PCR product using plasmid # 5 as template. Although the size of the plasmid number 2 was about same to the empty vector (Fig. 8), this plasmid also yielded LEA1.3 DNA. However, for further experiments the plasmid isolated from colony number 5 was used since it showed both a positive PCR result and larger size than the empty vector.

**Expression of fluorescent tagged LEA proteins in Drosophila cells**

Following generation of the above described four plasmid constructs (LEA1+LEA6, LEA2+LEA6, LEA3m+LEA6, and LEA1.3+LEA6), the *Drosophila* cell lines Kc167 and S2R+ were transfected with these plasmids. Furthermore, the parental vector STABLE-2 encoding for GFP and mCherry alone was transfected in both cell types. After 3-4 weeks of selection in the G418 medium, Kc167 cell lines started to grow robustly. However, growth rates of S2R+ cell lines were slower and it took around 7-8 weeks to obtain stably transfected S2R+ cell lines. Images of Kc167 and S2R+ cell lines demonstrated that most of the cells expressed the fluorescent tagged LEA proteins after selection in the G418 medium. However, the expression level of
LEA1.3+LEA6 (panel E). Images were taken by fluorescence microscope after 21 days of selection in the G418 medium.

Figure 11: Expression of mCherry (red, column 1) and GFP (green, column 2) fused LEA proteins in S2R+ cell lines. The cells were transfected with empty vector (panel A), LEA1+LEA6 (panel B), LEA2+LEA6 (panel C), LEA3m+LEA6 (panel D) and
fluorescent markers in Kc167 cell lines (Fig. 10) were lower than that in S2R+ cell lines (Fig. 11). Nonetheless, the Kc167 cell lines were used for Western blotting and other experiments due to their faster growth rate in the G418 medium. S2R+ cell lines were preserved in the liquid N2 tank for future studies.

Figure 10: Expression of mCherry (red, column 1) and GFP (green, column 2) tagged LEA proteins in Kc167 cell lines. The cells were transfected with empty vector (panel A), LEA1+LEA6 (panel B), LEA2+LEA6 (panel C), LEA3m+LEA6 (panel D) and
Confirming expression of fluorescent tagged LEA proteins in Kc167 cells

Immunoblot revealed that Kc167 cells transfected with the described four plasmids expressed mCherry tagged LEA1, LEA2, LEA3m, and LEA1.3 proteins. However, expression of GFP tagged LEA6 proteins was only successful in the LEA3m+LEA6 cell line, but not in other cell lines. Cells transfected with empty vector expressed only mCherry (Fig.12C) and GFP (Fig. 12G). In the LEA1+LEA6 cell line, immunoblot detected the right band for mCherry-LEA1 protein (Fig. 12D) but not GFP-LEA6 rather only the fluorescent marker (Fig. 12H). However, Western blotting confirmed successfully double expression of LEA3m and LEA6 as it detected both mCherry tagged LEA3m (Fig. 12E) and GFP fused LEA6 proteins (Fig.12I). Another immunoblot demonstrated expression of mCherry fused LEA1.3 (Fig. 13D) in cells transfected with LEA1.3+LEA6 construct, but anti-GFP antibody bound with a protein similar to the size of the only GFP (Fig. 14E) rather than the chimeric protein LEA6-GFP. Western blot also confirmed that cells transfected with LEA2+LEA6 construct, expressed mCherry tagged LEA2 (Fig. 13E) but identification of GFP tagged LEA6 remained elusive as anti-GFP antibody bound with only GFP (Fig. 14E). Thus, immunoblotting confirmed expression of mCherry tagged LEA proteins in all four cell lines, however, GFP tagged LEA expression was proved in only one cell line.
Figure 12: Immunoblotting for identification of fluorescent tagged LEA proteins in the *Drosophila Kc167* cell lines. Lane A: biotinylated protein ladder, lane B: non-transfected cells, lane C: empty vector control, lane D: LEA1+LEA6, lane E: LEA3m+LEA6, lane F: non-transfected cell, lane G: empty vector control, lane H: LEA1+LEA6, lane I: LEA3m+LEA6. Numbers indicate molecular weight of the probed proteins (kDa).
Figure 13: Immunoblotting to detect mCherry tagged LEA1.3 and LEA2 proteins in the *Drosophila* Kc167 cell lines. Lane A: biotinylated proteins standard, lane B: non-transfected Kc167 cells, lane C: empty vector control, lane D: LEA1.3+LEA6, lane E: LEA2+LEA6. The blot demonstrates that Kc167 cells expressed mCherry-LEA1.3 (lane D) and mCherry-LEA2 (lane E) proteins. Numbers indicate molecular weight of the probed proteins (kDa).
Figure 14: Immunoblotting for identification of GFP tagged LEA protein in the *Drosophila Kc167* cell lines. Lane A: biotinylated proteins standard, lane B: empty lane, lane C: non-transfected cells, lane D: empty vector, lane E: LEA1.3+LEA6, lane F: LEA2+LEA6, lane G: empty lane, and lane H: LEA3m+LEA6 (positive control). The blot shows that the chimeric protein GFP-LEA6 was only expressed in the LEA3m+LEA6 cell line. Numbers indicate molecular weight of probed proteins (kDa).

**Inserting LEA1.1 at the second position of the STABLE-2 vector**

Next LEA6 was replaced by LEA1.1 which is also a cytoplasmic protein belongs to group 1 LEA proteins. The gene LEA6 was removed from the plasmids LEA3m+LEA6 and LEA2+LEA6 by restriction enzymes digestion, and LEA1.1 was ligated at the second cloning site of the vectors. Following transformation, plasmids were isolated from five distinct colonies and PCR carried out by utilizing isolated plasmids as template. PCR showed that LEA1.1 was present in all isolated plasmids for both combinations (Figs. 15 and 17). Single digestion revealed that the size of the five
plasmids cloned with LEA3m+LEA1.1 were larger than empty vector (Fig. 16). However, restriction enzyme failed to cut the plasmids having LEA2+LEA1.1, probably due to bacterial recombination during the cloning process (Fig. 18). Therefore, the effort to construct plasmids with LEA3m+LEA1.1 was successful, but that of LEA2+LEA1.1 was fruitless.

Figure 15: PCR products of LEA1.1 (546 bases) using purified plasmids as template DNA. Lane 1: 2-log DNA ladder, lane 2-6: PCR products from the five isolated plasmids. PCR products in lanes 2-5 (from plasmids # 1-4) indicated the gene LEA1.1 was present in these plasmids, but DNA size in lane 6 (from plasmid # 5) was too smaller than the gene of interest.

Figure 16: Single digestion of the plasmids that showed positive PCR result for LEA1.1. Lane 1: empty vector, lane 2: vector with only LEA3m, lane 3-7: plasmids with
LEA3m+LEA 1.1. The size of the all isolated plasmids (lane 3-7) were larger than empty vector (lane 1), and vector with only LEA3m (lane 2).

Figure 17: Amplification of LEA1.1 DNA (546 bases) through PCR using isolated plasmids as template DNA. Lane 1: DNA ladder, lane 2-6: PCR products of LEA1.1 from plasmids # 1-5 respectively. The image indicates that all five selected plasmids have the insert LEA1.1 at the second position of the vectors.

Figure 18: Single digestion of the plasmids that showed positive result in PCR for gene LEA1.1. Lane 1: parental vector, lane 2: STABLE-2+LEA3m, lane 3-7: plasmids with LEA2+LEA1.1. Although restriction enzyme EcoRI digested empty vector and plasmid with only LEA3m, it did not work for the isolated plasmids.
2. Cloning of LEA 1.3 with ribosome recognition sequences

The Kozak consensus sequence (GCC GCC ACC AUG G) was named according to the Marilyn Kozak who discovered the pattern on vertebrate mRNA (Kozak 1986). This sequence occurs on vertebrate mRNA molecule is recognized by the ribosome as the translational start site, from which a protein is produced according to the coding template of a gene carried on that mRNA molecule (Valasek 2013). However, the consensus translation initiation site described by Kozak (1986) was derived primarily from vertebrate mRNA sequences. *Drosophila* nuclear genes exhibit a significantly different consensus sequence (GCC AAC AUG) for translation start than that of vertebrate (Cavener 1987). The idea of Kozak sequence in this study came from the fact that *Drosophila* cells transfected with the LEA1.3+LEA6 vector construct expressed very low level of fluorescent tagged LEA proteins than other cell lines (Figs.10 and 11). Retrospectively, I figured out that the gene sequence of LEA1.3 (obtained from the IDT DNA Technologies) had GCC GCC ACC (vertebrate Kozak) sequence at the upstream of translation start codon ATG. I hypothesized that the additional vertebrate Kozak sequence might be the reason for lower expression of the fluorescent markers in *Drosophila* cells. Hence, the gene LEA1.3 was cloned with *Drosophila* ribosome recognition sequence (GCC AAC ATG) and with vertebrate Kozak sequence (GCC GCC ACC ATG G) at the first position of STABLE-2+LEA6 vector. Another plasmid for LEA1.3 control (no ribosome recognition sequence) was generated to compare expression levels of the protein. The verification processes of inserting LEA1.3 with and without ribosome recognition sequences were similar to that of other constructs. Three plasmids for LEA1.3 control, LEA1.3 with *Drosophila* ribosome recognition sequence, and LEA1.3 with vertebrate Kozak sequence were isolated and PCR carried out. PCR results demonstrated that one plasmid from each of
the three groups incorporated the gene LEA1.3 (Fig. 19). Size comparison of the plasmids that showed positive result in PCR confirmed insertion of LEA1.3 DNA into these plasmids as their size were larger than the size of the empty vector (Fig. 20).

Figure 19: PCR products of LEA1.3 (600 bases) using isolated plasmids as template DNA. Lane 1: DNA ladder, lane 2-4: LEA1.3 control, lane 5-7: LEA1.3 with Drosophila ribosome recognition sequence, lane 8-10: LEA1.3 with vertebrate Kozak sequence. Among three selected colonies, plasmid in lane 2 (# 1) for control, plasmid in lane 6 (# 2) for Drosophila, and plasmid in lane 9 (# 2) for vertebrate ribosome recognition sequences showed positive PCR results.

Figure 20: Digestion of the plasmids that showed PCR products of LEA1.3. lane 1: empty vector, lane 2: STABLE-2+LEA1.3 control, lane 3: STABLE-2+LEA1.3 with Drosophila translation initiation sequence, lane 4: STABLE-2+LEA1.3 with vertebrate Kozak sequence. The larger sizes of the isolated plasmids than empty vector indicated that these plasmids have the target gene LEA1.3.
Expression of the protein LEA1.3 with ribosome recognition sequences

In accordance with the assumption, expression of fluorescent tagged proteins was lowest in Kc167 cells transfected by construct LEA1.3 with vertebrate Kozak sequence (Fig. 21C). Intuitively, *Drosophila* ribosome recognition sequence helped in the initiation of translation as this cell line expressed highest amount of fluorescent tagged LEA proteins (Fig 21B). However, LEA1.3 control cells (without ribosome recognition sequence) showed intermediate level of expression in comparison to other two cell lines. Furthermore, the three cell lines expressed proportional amount of GFP tagged LEA6 proteins. Therefore, ribosome recognition sequences played important role in expression of proteins in the cell lines derived from the fruit fly.

Figure 21: Expression of mCherry tagged LEA1.3 with and without ribosome recognition sequences (red, column 1) and GFP fused LEA6 proteins (green, column 2). Cells transfected with LEA1.3 control+LEA6 (panel A), LEA1.3 with *Drosophila* ribosome recognition sequence+LEA6 (panel B), and LEA1.3 with Kozak sequence + LEA6 (panel C). The images were taken by fluorescence microscope after 28 days of selection in the G418 medium.
3. Expression of non-fluorescent tagged LEA proteins

Cloning of His tagged LEA6 at the second position of the STABLE-2 vector

In order to express LEA6 without the fluorescent protein, the vector STABLE-2 was digested with the restriction enzymes EcoRV and HindIII to remove GFP from the vector. Following gel purification of the digested vector, LEA6-His was ligated at the second cloning site of the vector. Similar to previous constructs, five plasmids were isolated and PCR carried out using purified plasmid DNA as the template. PCR results suggested that the target gene LEA6-His was inserted in the five isolated plasmids (Fig. 22). However, vector digestion revealed that the sizes of only four isolated plasmids (#1, 2, 4, and 5) were larger than empty vector implying that they incorporated the desired gene. On the other hand, one plasmid was identical in size with the empty vector (Fig. 23), so this plasmid was not used in further experiments.

Figure 22: PCR amplification of LEA6-His (771 bases) using isolated plasmids as DNA template. Lane 1: 2-log DNA ladder, lane 2-6: LEA6 from plasmids number 1-5 respectively. All five purified plasmids yielded PCR products for LEA6 indicated that the gene was present in the isolated plasmids.
Figure 23: Linearized empty vector and plasmids that yielded DNA of LEA6-His. Lane 1: DNA ladder, lane 2: empty vector (double digested), lane 3-7: purified plasmids from bacterial colony # 1-5 respectively. The sizes of the plasmids in lane 3, 4, 6, and 7 were larger than the parental vector (lane 2), but plasmid in lane 5 was similar in size to empty vector.

**Inserting DDK tagged LEA1 at the position of STABLE-2+LEA6-His vector**

After verification of LEA6-His at the second position, this vector was digested with the restriction enzyme KpnI and NotI to remove the fluorescent protein mCherry from position 1 of the vector. Then, the gene LEA1-DDK was ligated at the first cloning site of the vector. After transformation, four plasmids were purified from distinct colonies and the DNA of LEA1 was amplified through PCR by using isolated plasmid as template. PCR results suggested that the gene LEA1 was present in four isolated plasmids (Fig. 24). Vector digestion revealed that the size of three plasmids (# 1-3) were larger than plasmid with only LEA6 but one plasmid (# 4) was similar in size to the STABLE-2+LEA6 (Fig. 25). Hence, insertion of both LEA1-DDK and LEA6-His in the STABLE-2 vector was verified, thus this construct became ready to transfect in *Drosophila* cells for the expression of two LEA proteins without fluorescent markers.
Figure 24: PCR amplification of LEA1-DDK (1074 bases) using isolated plasmids as template DNA. Lane 1: DNA ladder, lane 2-5: LEA1-DDK from plasmids number 1-4 respectively. The four selected plasmids incorporated the target gene LEA1-DDK as indicated by the PCR products.

Figure 25. Digestion of the plasmids that showed positive result in PCR for LEA1-DDK. Lane 1: DNA ladder, lane 2-3: plasmids with only LEA6-His (double digested), lane 4-7: plasmids isolated from colony # 1-4 respectively. The sizes of plasmids in lanes 4-6 were larger than vector with only LEA6-His (lane 2-3). However, one plasmid (lane 7) was identical in size to the STABLE-2+LEA6-His, so it was excluded from further experiments.
Cloning of DDK tagged LEA2 and LEA3m at the first position of the vector

Next, LEA2-DDK and LEA3m-DDK were cloned at the first site of the vector STABLE-2+LEA6-His. In the similar fashion, these genes were ligated at the first position of the vector and PCR carried out to verify the desired genes in the isolated plasmids. Among five selected plasmids, LEA2 was present in three (# 1, 4, & 5) but was absent in two (# 2, & 3) plasmids (Fig. 26). On the other side, the gene LEA3m was present in only one (# 5) among five selected plasmids (Fig. 27). Plasmids that showed positive result in PCR were double digested to confirm the target genes. As expected, double digested vectors revealed DNA of LEA3m and LEA2 on the agarose gel (Fig. 28). Thus, two additional plasmids were generated for expression of two LEA proteins without the fluorescent reporter.

Figure 26: PCR amplification of LEA2-DDK (1092 bases) using isolated plasmids as template DNA. Lane 1: DNA ladder, lane 2-6: PCR products of LEA2 from the isolated plasmids. The plasmids in lane 2, 5, and 6 yielded DNA of LEA2-DDK, implied that these plasmids incorporated the gene of interest.
Figure 27: PCR product of LEA3m-DDK from the isolated plasmids. Lane 1: DNA ladder, lane 2-6: amplification of LEA3m from plasmids number 1-5 respectively. Only plasmid in lane 6 (# 5) had incorporated the gene LEA3m-DDK as suggested by the PCR product.

Figure 28: Single and double digestion of the plasmids (# 5) that showed positive result in PCR for LEA2 and LEA3m. Lane 1: DNA ladder, lane 2: double digestion of plasmid with LEA3m DDK+LEA6 His, lane 3: LEA3m control DNA, lane 4: double digestion of plasmid with LEA2 DDK+LEA6 His, lane 5: LEA2 control DNA, lane 6: single
digestion of plasmid with LEA3m DDK+LEA6 His, lane 7: single digestion of vector with LEA2 DDK+LEA6 His, lane 8: empty vector.

**Cloning of DDK tagged LEA1.3 at first position of STABLE-2+LEA6 His vector**

The gene LEA1.3-DDK was inserted at the first position of the vector STABLE-2+LEA6-His to generate the final plasmid construct for this study. Similarly, PCR was used to amplify LEA1.3 DNA by using six isolated plasmids as DNA template. All six plasmids showed PCR products that run with the correct size on agarose gels for LEA1.3 (Fig. 29). Double digestion of three plasmids (# 1-3) also revealed that LEA1.3 was present in two (# 1 & 3) plasmids (Fig. 30).

![Figure 29: PCR products of LEA1.3-DDK (600 bases) from the isolated plasmids. Lane1: 2-log DNA ladder, lane 2-7: amplification of LEA1.3 using isolated plasmids as template. All purified plasmids had the target gene LEA1.3-DDK as indicated by the gel picture.](image-url)
Figure 30: Double digestion of the three isolated plasmids (#1-3) that showed positive PCR result for LEA1.3-DDK. Lane 1: DNA ladder, lane 2-4: plasmids number 1-3 respectively. Double digestion revealed that LEA1.3-DDK (546 bp) was present in lanes 2 and 4 (#1 & 3 respectively). However, same restriction enzyme cut a large DNA fragment (1100 bp) from the plasmid in lane 3 (#2), so it was discarded.

**Confirmation of DDK tagged LEA proteins in Drosophila Kc167 cells**

In order to express LEA proteins without the fluorescent probes mCherry and GFP, *Drosophila Kc167* cells were transfected by the plasmid constructs LEA2 DDK+LEA6 His, and LEA3m DDK+LEA6 His. As before, transfected cell lines were selected in the G418 medium (2 mg/mL) to obtain stable cell lines. When transfected cells started to thrive, proteins were isolated and ran on a Western blot. Immunoblotting detected a clear band for DDK tagged LEA2 (Fig. 31E) and a faint band for LEA3m-DDK tag (Fig. 31D). However, identification of protein LEA6-His transcribed at second position of the constructs remained elusive as anti-His tag antibody did not bind with any protein extracted from these two cell lines except protein standard (data not shown).
Figure 31: Immunoblotting to detect DDK tagged LEA3m and LEA2 proteins in *Drosophila* Kc167 cell lines. Lane A: biotinylated protein ladder, lane B: non-transfected cells, lane C: empty vector, lane D: LEA3m DDK+LEA6 His, lane E: LEA2 DDK+LEA6 His. The blot showed a clear band for DDK tagged LEA2 protein (lane E) and a faint band for protein LEA3m-DDK tag (lane D).
DISCUSSION

Expression of fluorescent tagged LEA proteins

Anhydrobiosis represents a unique example of organisms’ adaptation to water loss, where an organism can exist in an ametabolic state until water returns (Gusev et al. 2014). The cyst of the brine shrimp *Artemia franciscana* exhibit anhydrobiosis by entering diapause, a state of developmental arrest and greatly enhanced stress tolerance (Clegg 1967, Hand et al. 2007, McRae 2016). Probably the main feature distinguishing anhydrobiotic organisms including *Artemia* is that they produce many types of highly hydrophilic proteins in preparation for severe dehydration (Tunnacliffe et al. 2010). LEA proteins are hydrophilic and non-globular proteins, and recent findings show that they play various roles in dehydrating cells, including homeostasis of proteins and nucleic acids, stabilizing cell membranes, redox balance, and the formation and stability of glassy state (Tunnacliffe and Wise 2007). Despite several studies on *Artemia* LEA proteins, their functions, roles, and localizations in the anhydrobiotic cyst remain unknown (Kim et al. 2015). Here, I aimed to express several combinations of two AfrLEA proteins in the *Drosophila* cells to understand the functions of multiple LEA proteins.

In this study, I was able to demonstrate that simultaneous expression of two different LEA proteins in *Drosophila melanogaster* cells is possible by using a multicistronic vector. I used the vector pAc5-STABLE2-Neo because in this construct, mCherry and GFP tagged LEA proteins are separated from each other by a ‘self-cleaving’ T2A peptide sequence. This vector transcribes tricistronic mRNA transcripts that are efficiently processed in S2R+ and Kc167 cells (Gonzalez et al. 2011). The self-cleaving nature of the T2A sequence allowed for correct processing of the LEA3m-
mCherry and LEA6-GFP proteins in Kc167 cells as demonstrated by the correct molecular weight (Fig. 12).

Aside from the LEA3m+LEA6 combination, in three other Kc167 cell lines, only the sequence inserted at position 1 was correctly processed (Figs. 12 & 13) and the anti-GFP antibody failed to detect the LEA6-GFP fusion protein. In these cell lines only the fluorescent marker protein was detected (Fig. 14). Why other cell lines did not express the correct chimeric construct remain elusive, although the verification process to confirm insertion of the genes into all the plasmids was similar. The negative result in the Western blot can arise due to several factors related to antibody, antigen, or buffer used (Liu et al. 2014). Since the antibody reacted with GFP and the LEA6-GFP construct in the LEA3m+LEA6 cell line (Fig. 14), the problem was not related to the antibody and buffer system used. The probability of undetectably low expression levels of the fusion protein can also be excluded because immunoblot showed a prominent band for GFP (Fig. 14). The failure to express one of the two transgenes driven by two separate promoters may be due to interference and/or silencing of promoters, vector rearrangements, and deletions (Curtin 2008). Since pAc5-STABLE2-Neo is a multicistronic vector, where expression of multiple proteins is driven by the single promoter Actin5C, the problem was not germane to the promoter.

However, there are three probable reasons for the negative Western blot result: (1) Drosophila cells might have deleted LEA6 gene through recombination while incorporating the construct into their genome, (2) the constructs that were transfected might have a mutated LEA6 gene, and (3) protease activity might have cleaved off GFP from the fusion protein LEA6 during post-translational modification. The assumption that protease might have cleaved off the marker GFP from the LEA6 is less likely since it did not happen for the cell line LEA3m+LEA6. Plasmid recombination is a common
problem in sub-cloning project specifically when either or both the vector and the insert fragments are large, or contain repeated sequences that destabilize DNA (Bzymek and Lovett 2001). LEA genes which are usually enriched in repeated motifs (Goyal et al. 2003), probably susceptible to DNA recombination. Furthermore, the size of the pAc5-STABLE2-Neo vector is 7605 base pairs which is fairly large and it showed recombination during generation of the LEA2+LEA1.1 construct, as restriction enzymes failed to cut the vector (Fig. 18), although PCR confirmed the insertion of LEA1.1 (Fig. 17). Recombination is facilitated by DNA breaks or nicks, something that can result from UV damage during gel viewing or by harsh chemical reagents in the DNA purification kits (Bernstein et al. 2011). Thus, recombination of the constructs and/or mutation of the gene LEA6 may be the reason for the negative result in immunoblotting.

In the pAc5-STABLE2-Neo vector the antibiotic resistance gene neomycin is detached from the fusion proteins by another T2A sequence which avoids the need to use co-transfection with a separate vector expressing an antibiotic resistance gene or the need for a dual promoter (Gonzalez et al. 2011). Similar to the finding of Gonzalez et al. (2011), I observed that G418-based selection is efficacious and took 3–4 weeks to establish a stable population of Kc167 cells. Images of the fluorescence microscopy showed that most of the G418 selected Kc167 and S2R+ cells (Figs. 10 & 11 respectively) expressed fluorescent proteins mCherry and GFP. However, the expression levels of the fluorescent proteins in S2R+ were higher than in Kc167 cells. Although both Kc167 and S2R+ cell lines were derived from embryos of Drosophila, Kc167 cells are small and round (10 µm) whereas S2R+ cells are large and flat (averaging 50 µm) and strongly adherent to surface (Kiger et al. 2003). Gene expression varies between cells in a multicellular organism even though they shared identical
genomes (Alberts et al. 2002). Thus, difference in expression level of fluorescent tagged LEA proteins in these two cell lines was not surprising. On the other side, higher amount of the transgenes expression might be the reason of slow proliferation rate of S2R+ cells in the G418 selection medium. Wang et al. (2013) reported that the level of GFP as the second transgene in the T2A vector is always lower than the level of murine MHC class II chaperone, invariant chain (Ii) in the context of lentiviral transduction, whereas GFP and Ii expression were comparable in 293T cells directly transfected with T2A vector. Although images of the cells (Figs 10 & 11) showed lower expression of the second transgene GFP, Western blot did not support that observation as the thickness of the protein bands for both transgenes were about same (Fig. 12).

**Effect of ribosome recognition sequences on expression level of proteins**

The Kozak consensus sequence plays a major role in the initiation of the translation process (Kozak 1986). This sequence found on an mRNA molecule is recognized by the ribosome as the translational start site (De Angioletti 2004). However, the consensus sequence for eukaryotic translation initiation sites by Kozak (1986) was derived primarily from vertebrate mRNA sequences. The nuclear genes of *Drosophila* show a substantially different translation start consensus sequence (Cavener 1987). The results showed that consensus ribosome recognition sequence of *Drosophila* has helped to improve the expression level of mCherry tagged LEA1.3 proteins in Kc167 cells above control levels. Surprisingly, the vertebrate ribosome recognition sequence (Kozak sequence) inhibited expression level of the mCherry labelled LEA1.3 proteins in *Drosophila* cells. The images of fluorescence microscopy demonstrated that cells with *Drosophila* ribosome recognition sequence not only increased expression level of the mCherry tagged protein but also GFP tagged second LEA protein. Similar to the first protein mCherry-LEA1.3, expression level of GFP
fused LEA6 was highest in cells transfected with *Drosophila* ribosome recognition sequence and lowest was in cells with Kozak sequence (Fig. 21). Hence, I conclude that the ribosome recognition sequence for *Drosophila* described by Cavener (1987) is effective for enhancing expression level of protein in *Drosophila* cells. Conversely, vertebrate Kozak sequence is not useful for protein expression in *Drosophila* cells, rather this sequence is inhibitory for non-vertebrate cells.

**Expression of non-fluorescent tagged LEA proteins**

Although fluorescent probes are widely used in molecular biology to monitor and locate subcellular localization of expressed transgene, their large size (mCherry 29 kDa, and GFP 27 kDa) is a matter of concern. Moreover, the size of LEA proteins from *Artemia franciscana* ranges between 16 to 41 kDa, which are about equal in size to the fluorescent proteins. The vector pAc5-STABLE2-Neo is versatile and open reading frames (ORFs) with 6X His, glutathione S-transferases (GST), or others tag that facilitate protein identification and purification can be easily used instead of GFP and mCherry (Gonzalez et al. 2011). Thus, to avoid potential interferences with LEA functions by the fluorescent proteins, I cloned LEA proteins with two alternative tags (DDK and 6X His). Although T2A mediated cleavage seems efficient in both S2R+ and Kc167 cells, it is highly recommended to confirm processing of polyproteins by Western blot (Gonzalez et al. 2011). I performed immunoblotting to detect proteins from the cells transfected with (LEA2 DDK+LEA6 His) and (LEA3m DDK+LEA6 His) constructs. The immunoblot revealed that anti-DDK antibody detected both LEA2 and LEA3m proteins (Fig. 31). However, the band for LEA3m was faint probably due to low level of expression of this mitochondrial protein.
Another reason might be that binding of anti-DDK antibody with LEA3m was interfered due to post-translational modification (PTM) of this protein. Schmidt et al. (2012) reported that despite the heavy use of FLAG/DDK in numerous laboratories worldwide, it was surprising that in insect cells a PTM interferes with the FLAG-anti-FLAG interaction rendering this tag system ineffective for secreted proteins. However, LEA3m is not a secretory protein, but PTM of this intrinsically disordered protein might be possible. Warner et al. (2010) hypothesized that PTM in group 1 LEA proteins from brine shrimp *Artemia franciscana* is possible since they contain multiple sites with a high probability of phosphorylation. A complex combination of different PTMs, including phosphorylation, acetylation, methylation and deamination in the native form of group 1 LEA protein Emb564 have been reported by Amara et al. (2012). Phosphorylation of Rab17, a group 2 LEA protein has been reported much earlier (Plana 1991). In LEAM, a pea mitochondrial LEA protein belonging to group 3, the occurrence of deaminations and oxidations has been proposed to contribute to the functional conformation of the protein (Tolleter et al. 2007). Therefore, some form of PTM in AfrLEA3m, a member of group 3 LEA protein, may be happened that interfered interaction with anti-DDK antibody.

Nonetheless, immunoblot probed with anti-His antibody did not show any band for LEA6 protein in these two cell lines. Several factors might be responsible for the negative results in Western blotting of 6X His tagged recombinant proteins. These include the availability of the His tag to the antibody, the location of the His tag on the individual protein, protein purity, antibody dissociation constant, and the length of the His tag (Debeljak et al. 2006). Since the immunoblot showed bands for the protein standard, the anti-His tag antibody was functional. Debeljak et al. (2006) also experienced similar problem in detecting C-terminal 6X His tagged recombinant
protein through immunoblot despite successful incorporation of the 6X His tag into the
cDNA constructs. The 6X His tag on the C-terminus of the recombinant antigen
underwent some extra folding such that the tag was inaccessible to the antibody during
the Western blot (Kadasia 2012). Hence, I assumed that the 6X His tag has been buried
within the hydrophilic LEA6 protein and did not bind to the antibody.

Another reason might be that exocytosis leads to diffusion of the protein LEA6
through the plasma membrane and rendered undetectable. de Felipe et al. (2010)
showed that due to cleavage inefficiency of 2A peptide sequence, a large proportion of
the translation products are uncleaved, leading to translocation of the fusion protein into
the endoplasmic reticulum (ER) without presence of any such sequence and finally
exocytic pathway. They described this event as a form of ‘slipstream’ translocation
where downstream proteins, without signals, were translocated through a translocon
pore. The slipstream translocation is a result of inhibition of the 2A reaction (cleavage
between cysteine and proline residue) by the C-terminus of upstream proteins when
translocated into the ER (de Felipe et al. 2010). Although I did not observe uncleaved
products on immunoblot, exocytosis of fusion protein could an explanation for the
negative Western blotting result.
CONCLUSION

I have demonstrated that by using a multicistronic vector, concurrent expression of the two AfrLEA proteins is possible in *Drosophila* Kc167 cells. I also showed that *Drosophila* ribosome recognition sequence improved, but vertebrate Kozak sequence inhibited expression of fluorescent tagged LEA proteins in *Drosophila* cells. For non-fluorescent tags, expression of DDK tagged LEA proteins cloned at first position of the vector were confirmed, but His tagged LEA6 protein transcribed at second position of the vector was not identified probably due to unavailability of His tag to the antibody and/or exocytosis of the protein. However, this problem might be specific to only LEA6 protein which is recently discovered in *Artemia* and is not well characterized yet. The results suggest that pAc5-STABLE2-Neo is a unique vector for simultaneous expression of two proteins in *Drosophila* cells by single transfection in lieu of a traditional two different vectors system. Nonetheless, future efforts are needed to clone other LEA protein at the second position of the vector. Another possible option might be use of other tags instead of 6X His tag for the LEA6 cloned at the second position of the vector. Use of the N-terminal tag instead of the C-terminus tag that I used might be another option for non-fluorescent tagged expression.
REFERENCES


Clegg JS (1965). The origin of trehalose and its significance during the formation of encysted dormant embryos of *Artemia salina*. Comparative Biochemistry and Physiology, 14: 135-143


Kozak M (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell, 44: 283-292


proteins suggest protein-nucleic acid binding functionality. Computational and Mathematical Methods in Medicine, 2013: 470390

Leprince O and Buitink J (2010). Desiccation tolerance: from genomics to the field. Plant Science, 179: 554-564


