Alteration of the N-Terminus to analyze Enhancer of Rudimentary Homolog, ERH function

Senior Honors Thesis

Kennedy Light

Biology Major – SUNY Brockport

May 17, 2021

Thesis Director: Dr. Stuart Tsubota, Professor, Biology

Submitted in Partial Fulfillment of the Requirements for Graduation from the Honors College

Table of Contents

Figures and Tables

Abstract

The *enhancer of rudimentary* gene, *e(r)*, is the gene that encodes the Enhancer of Rudimentary Homolog protein (ERH). The amino acid structure of this protein, especially the first five amino acids which make up the N-terminus and the beginning of the first β-strand in the β-sheet of the protein's secondary structure, have been found to be highly conserved among a wide variety of organisms, including vertebrates, insects, and protists. This suggests that these amino acids may play a key role in the activity of ERH, and that ERH may have a conserved function across organisms. Mutations in amino acids 2 through 5 were constructed and the mutant *e(r)* alleles were inserted into the third chromosome of *Drosophila melanogaster*. The activity of the resulting mutant ERH proteins was assessed by their ability to rescue two mutant phenotypes of *e(r)* deletions – low viability of the single mutant and the synthetic lethality as a double mutant with a low-activity *Notch* mutation. Phenotypic observations of eye color show whether any of the mutations give rise to wild-type ratios of females versus males. The results indicate that the N-terminus of ERH is indeed necessary for its proper function, and that this function may be evolutionarily conserved.

Introduction

The *e(r)* gene was first identified as a mutation which enhanced the wing truncation of certain *rudimentary* mutations (Tsubota and Schedl 1986). The gene was subsequently cloned and sequence, and shown to encode a small, highly conserved protein, now referred to as Enhancer of Rudimentary Homolog or ERH (Wojcik et al. 1994). Additional DNA cloning and sequencing of *e(r)* homologs from *Homo sapiens* (humans), *Mus musculus* (mouse), *Brachydanio rerio* (zebrafish), *Caenorhabditis elegans* (roundworm), and *Arabidopsis thaliana* (flowering plant) revealed a high conservation of e(r) among vertebrates, invertebrates, and higher plants (Gelsthorpe et al. 1997). This study also postulated that the e(r) gene has a crucial role in cell division, given its high expression in cells undergoing DNA synthesis and cell division. Researchers are still looking into all the functions of ERH but arguably the most significant function discovered so far is its role as a promoter of human cancer cell growth through its regulation of pyrimidine synthesis, an essential process in DNA replication, and cellular division (Wojcik et al. 1994; Zafrakas, M. et al. 2008). ERH regulates these processes by playing an essential role in mRNA splicing of the centromere-associated protein E (CENP-E), a mitotic motor protein that stabilizes the kinetochore-microtubule attachment, which is key in chromosome alignment and segregation during mitosis (Fujimura et al. 2012). Without it, aneuploidy can occur, which is a common characteristic of human tumors, suggesting a connection between ERH function and human cancer (Kops et al. 2005). ERH can also play a role in human cancer as a repressor of hepatocyte nuclear factor-1 (HNF-1), and binding the dimerization cofactor of HNF-1 (DCoH), which is a co-activator of HNF-1 (Wan et al. 2005). While ERH is normally expressed in human testes, ovaries, prostate, heart, and liver, there seems to be an overexpression of ERH in cancerous breast, testicular, and ovarian tissue, suggesting ERH overexpression may initiate and promote certain cancers (Zafrakas et al. 2008). The conservation of this protein across species allows us to use research of *e(r)* gene expression and activity in the replicating cells of a species such as *Drosophila*, and apply that information to human cancer cells.

In *Drosophila* cells, it has been discovered that expression of the *e(r)* gene may go through DNA synthesis and replication, which is similar to the way the *e(r)* gene expresses itself in cancer cells of humans. In *Drosophila*, the ERH may also potentially play a role in *Notch*

signaling, which plays a role in the development of tissue through distribution of different cell types within. It is believed ERH may have a role in this due to the knowledge that the pathway requires *e(r)* gene for accurate expression. Previous studies have shown that the *e(r)* gene is a positive regulator of the *Drosophila*'s *Notch* signaling pathway (Tsubota *et al.* 2011).

The striking conservation of ERH's amino acid sequence across a wide variety of species strongly suggests that particular regions of the protein are essential for its proper function. Alterations to the structure of ERH may prevent it from properly binding to other proteins or molecules, disrupting its function and role in certain biological processes. Fig. 1 shows the structure of human ERH, including its β-sheet of four antiparallel β-strands, three amphipathic α-helices, three loop regions connecting the α-helices and β-strands, an N-terminus, and a C-terminus. ERH has an overall monomeric structure, but the interactions between the β-sheet regions give it a dimeric structure as well (Arai et al. 2005). This study focuses on the effect of altering amino acids 2-5 at the end of the N-terminus, and the beginning of the β1 strand, which together makeup of the first 10 amino acids of the total 104 that code for ERH, and are identical in humans and *Drosophila*, as can be seen in Fig. 2. This particular region was chosen as the focus of this study for a few different reasons. Firstly, due to the fact that it is very highly conserved among both vertebrates and invertebrates, suggesting an importance in ERH function. Secondly, it does not have a secondary structure and is solvent accessible, making it more easily accessible to a potential ERH binding partner. Lastly, it is located very close to the β-sheet which is a potential binding site for ERH partners. Proteins can bind to the β-sheet and access the N-terminus, which may be important to ERH function.

The level of ERH amino-acid-sequence similarity is especially shocking in humans and *Drosophila*, considering the apparent physical differences of the two species and the estimated date of divergence of the two species of approximately 783 million years ago (Hedges et al. 2006). As shown in Fig. 2, *Drosophila* and human ERH amino acid sequences are 76% identical, and when including different but structurally similar amino acids in the same position, they are reported as being 84% similar. Previous research (Tsubota *et al.* 2016) showed that the *Drosophila* ERH is functionally equivalent to the human ERH, allowing us to insert a human *e(r)* gene into *Drosophila* and study its effects. A *Drosophila* with a lethal or disadvantageous *e(r)* mutation can be rescued by the insertion of a wild-type human *e(r)* gene. Therefore, the effects of mutations in the human *e(r)* gene on its activity can be assessed by their ability or inability to

rescue the *Drosophila e(r)* mutations. It is preferable to study the human ERH directly, rather than a similar ERH of another species such as *Drosophila*, so that we can relate our findings about the *e(r)* gene and ERH function directly to the human body.

To test our hypothesis that altering a particular site of the ERH will alter its functionality, *Drosophila e(r)* gene mutants were created. Following the insertion of these *e(r)* gene mutants into *Drosophila*, mutant crosses were also performed in order to assess the phenotypes that resulted, and determine which, if any, mutations can be rescued successfully. One of the mutant phenotypes we looked at was low viability caused by a $e(r)^{27\text{-}1}$ single mutation. This mutation is missing 43% of its coding region, along with the promoter sequence, which initiates transcription. Thus, this mutation is a null mutation (no activity). Each combination of mutation has different qualifications for "rescue." All alleles studied are X-linked, as previous research (Wojcik *et al.* 1994) has found the *e(r)* gene to be located in region 8B of the X chromosome. Therefore, a female to male ratio of *Drosophila* possessing a copy of a mutant *e(r)* gene we created will expose each's effect.

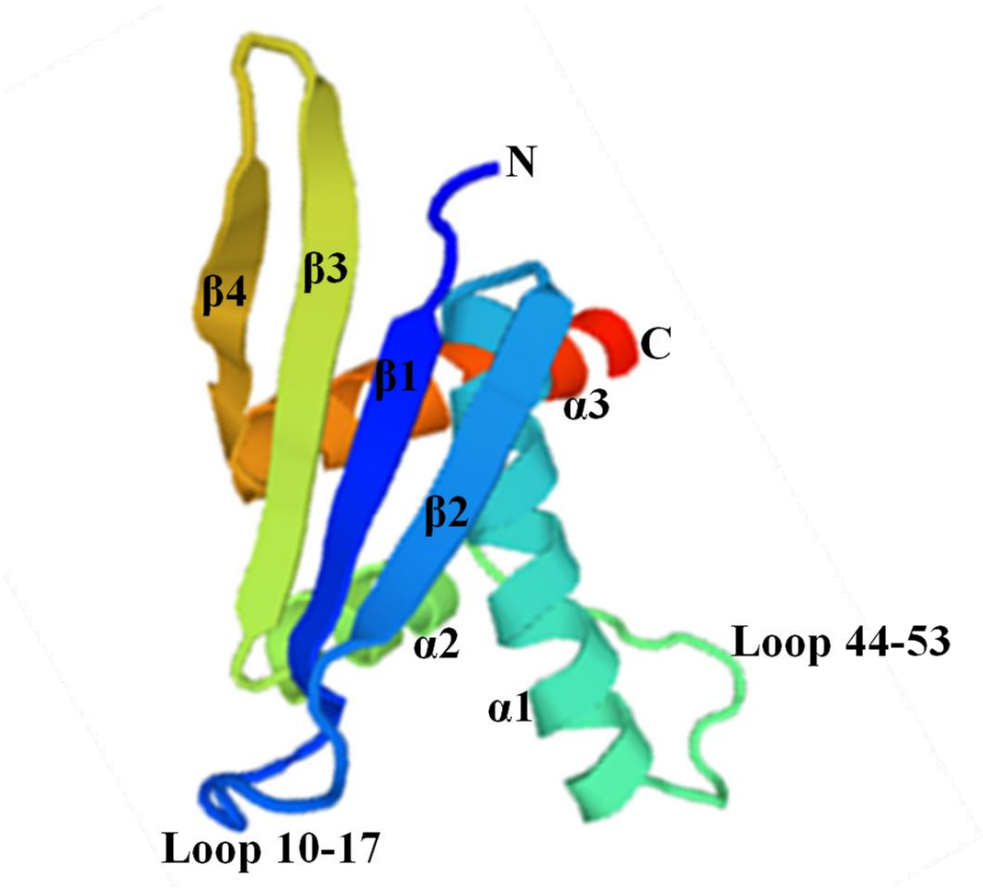

Figure 1. Pictured above is a 3-D model of the crystallized tertiary structure of ERH. The N-terminus, α-helices, β-strands, loops of interest, and C-terminus are labeled as such.

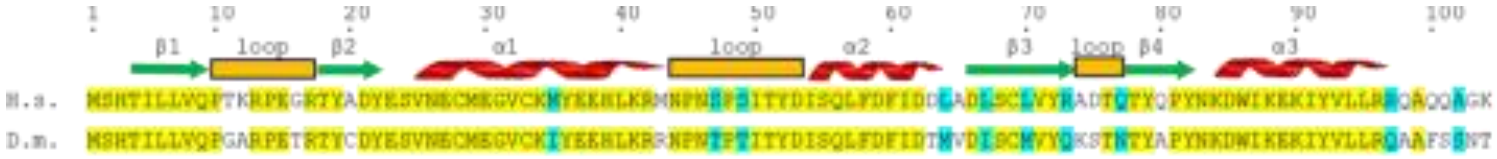

Figure 2. Pictured is the amino acid sequence of ERH in humans and *Drosophila*. In human and *Drosophila* ERH, amino acids highlighted yellow are identical, those highlighted in blue (as well as those in yellow) are similar, and non-highlighted amino acids are dissimilar.

Methods

The following process was used to produce the S2A and T4A *Drosophila e(r)* mutants, while lab mate Brianna Dudley followed a similar procedure to create and clone the H3A and I5D *Drosophila e(r)* mutants, all of which were analyzed in this study.

Part I: Creation of mutant *Drosophila e(r)* coding regions

As the wild-type amino acids of ERH are highly conserved, we are interested in replacing them with amino acids similar in size and shape, but differ in polarity, charge, and affinity for water. Table 1 and Fig. 4 summarize the amino acid changes made and their position in ERH, which are: N-terminus S2A, H3A, T4A, and β1 I5D. Mutants S2A and T4A were designed to prevent phosphorylation at these amino acids, as shown in Fig. 3.

Table 1. Altered Amino Acids Studied as a Result of Mutant *e(r)* Alleles

Structure	Residue Number	Wild-Type Amino Acid	Substituted Amino Acid
N-Terminus	2	S	A
N-Terminus	3	H	A
N-Terminus	4	T	A
β1 Strand	5	I	D

Table 1. Shows which residues within the N-terminus and β1 strand were altered, and which amino acid was substituted. Substituted amino acids were similar in size and shape to the wild-type but differed in another key physical property. Therefore, results can be attributed to this physical property, and the amino acid itself, rather than structural changes to ERH due to dramatic differences in wild-type versus substituted amino acid size and shape. For example, wild-type amino acids serine and threonine at residues two and four, respectively, are hydrophilic. These were replaced with alanine, a hydrophobic amino acid. Histidine, which has a positively charged side chain and is polar, was also replaced with hydrophobic alanine. Isoleucine, which has a neutral hydrophobic side chain, was replaced with aspartic acid, which has a negatively charged side chain.

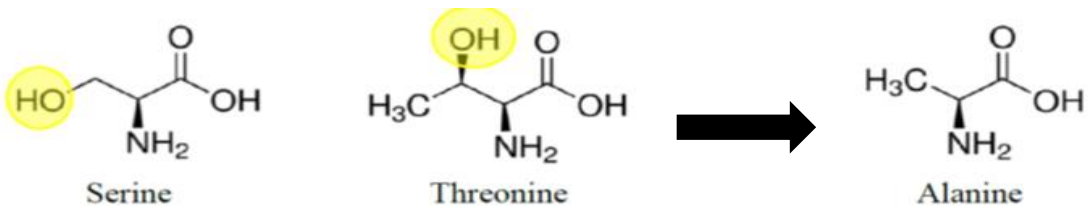

Figure 3. Shows the chemical structure of the wild-type amino acids at positions two and four of the *Drosophila* ERH, serine and threonine, respectively, and the amino acid they were replaced with in the S2A and T4A mutants (alanine). This replacement prevents phosphorylation at the highlighted hydroxyl groups.

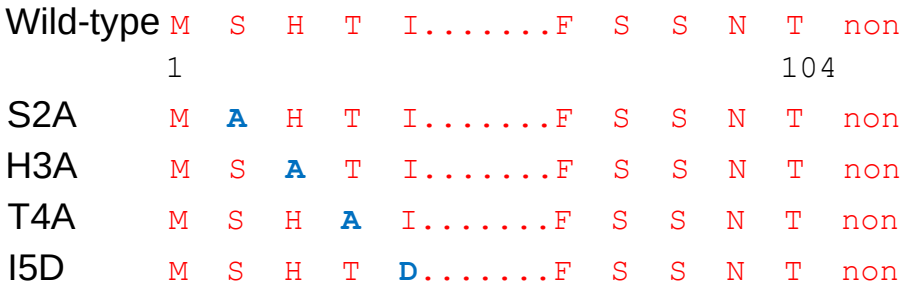

Figure 4. Summarizes the amino acid changes that were made and studied in this project. Each line contains the first five and last six amino acids in the protein, including the STOP codon, indicated by "non" at the end of each sequence. The top line represents the wild type amino acid sequence of *Drosophila* ERH. The last 4 lines represent the amino acid sequence of the mutants, with the blue letter in each line indicating which amino acid was altered in that mutant. Mutants S2A and T4A were constructed by myself, while mutants H3A and I5D were constructed by Brianna Dudley.

The synthesized double-stranded sequence of DNA for cloning the mutant sequences also needed to be created with the possibility of producing sticky ends (terminal DNA single stranded fragments of unpaired nucleotides that can base-pair with one another) at both sides. This was achieved in the mutant sequence by having a string of bases, followed by a PciI restriction enzyme site, followed by start codon ATG at one end, and the completion of the coding region followed by an NcoI site, followed by a few more bases at the other end**.** These restriction sites and the rest of the each sequence can be seen in Fig. 5. A file of this desired mutated DNA sequences was sent to Integrated DNA Technologies, Inc. (https://www.idtdna.com) for synthesis, and were sent back to our lab as a double-stranded mutant DNA sequence.

Figure 5. A more detailed look at the sequence of bases that code for the first five and last six amino acids that make up the wild-type ERH, along with the mutants developed: S2A and T4A. Blue text represents the mutant codons and mutant amino acids. Underlined sections are the digest sites of restriction enzymes NcoI and PciI. The PciI site, (coded by ACATGT) and the NcoI site (coded by CCATGG) produce the same sticky ends, allowing T4A to be ligated into the NcoI site of the empty *e(r)* p-SMART vector. The additional bases on the ends of the mutant *e(r)* sequences are eliminated in the upcoming cloning process

Part II: Ligation of mutant *Drosophila e(r)* gene into empty p-SMART vector

Once the *Drosophila* DNAs with the desired mutations were received, either enzymes PciI and NcoI (T4A mutant) or just NcoI (S2A mutant) were used to digest the sequence and create sticky ends for potential ligation into an empty *e(r)* p-SMART vector (see Fig. 6 for the location of the restriction sites). This vector was created by adding the 5' and 3' untranslated regions of *e(r)*, without the coding region, to a vector called p-SMART that is extremely useful for research (Tsubota et al. 2016). In place of a coding region, the empty *e(r)* p-SMART vector has a NcoI restriction site (CCATGG), allowing for the insertion of the *Drosophila* ERH coding region.

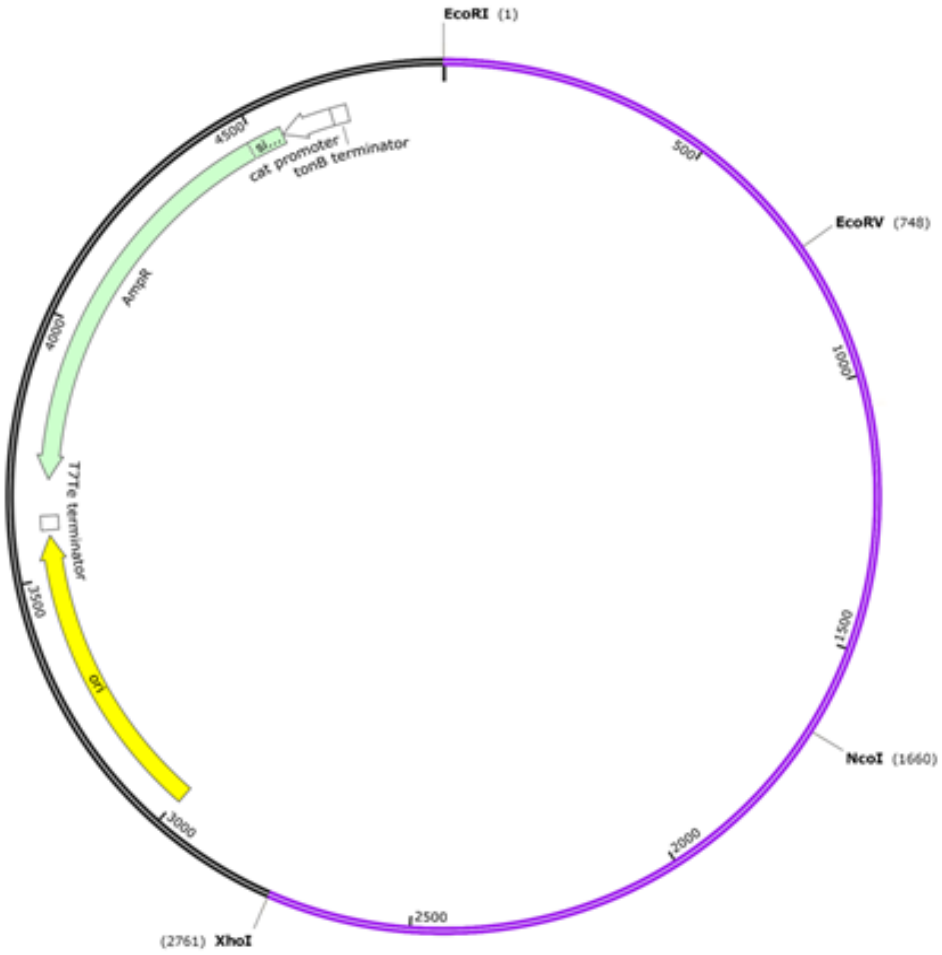

Figure 6. Shows an empty *e(r)* p-SMART vector and its EcoRI, EcoRV, and XhoI restriction sites. The empty *e(r)* p-SMART vector has no *e(r)* coding region, but an NcoI restriction site (CCATGG) instead. This restriction site is used to insert the mutant *e(r)* coding regions into the plasmid. The vector contains a 2.7-kb fragment with the 5' end upstream (EcoRI to NcoI) and 3' end downstream (NcoI to XhoI) of the *e(r)* coding region, which is necessary for normal gene expression.

A column purification procedure was used to purify the DNA that had been digested with NcoI and PciI (T4A mutant) or just NcoI (S2A), and to remove the restriction enzymes and small DNA fragments from each end of the fragment. This was then ligated into the non-coding region of the empty *e(r)* p-SMART vector, producing a 3.0-kb fragment containing the mutant *e(r)* gene. *E.Coli* cells were transformed with the recombinant plasmid and grown on ampicillin agar plates. Cells which were transformed with the plasmid were resistant to ampicillin and, therefore, would grow and produce colonies. Individual colonies were picked from the plates and re-spread onto new plates. Cell cultures of 1.5 ml were prepared for each colony and placed on a shaker-incubator at 250 rpm and 37°C overnight. The recombinant plasmids were isolated from the cells by performing a lysozyme mini-boiling prep on the cultures. The recombinant mutants containing S2A were digested with XbaI and BglII and the recombinant mutants containing T4A were digested with EcoRV and NcoI. Gel electrophoresis was then performed on them to determine the success of the ligation of the mutant DNA fragment into the empty *e(r)* p-SMART vector, and that the ligation was in the correct orientation. When the correct orientation was verified by the gel, the recombinant vectors were digested with EcoRI and XhoI to isolate the 3.0-kb mutant *Drosophila e(r)* gene containing the 5' and 3' untranslated region and coding region. After the double digest, another gel was run, and the 3.0-kb band was extracted and purified from the gel to isolate the *e(r)* gene, and another gel was run to confirm it had been properly isolated.

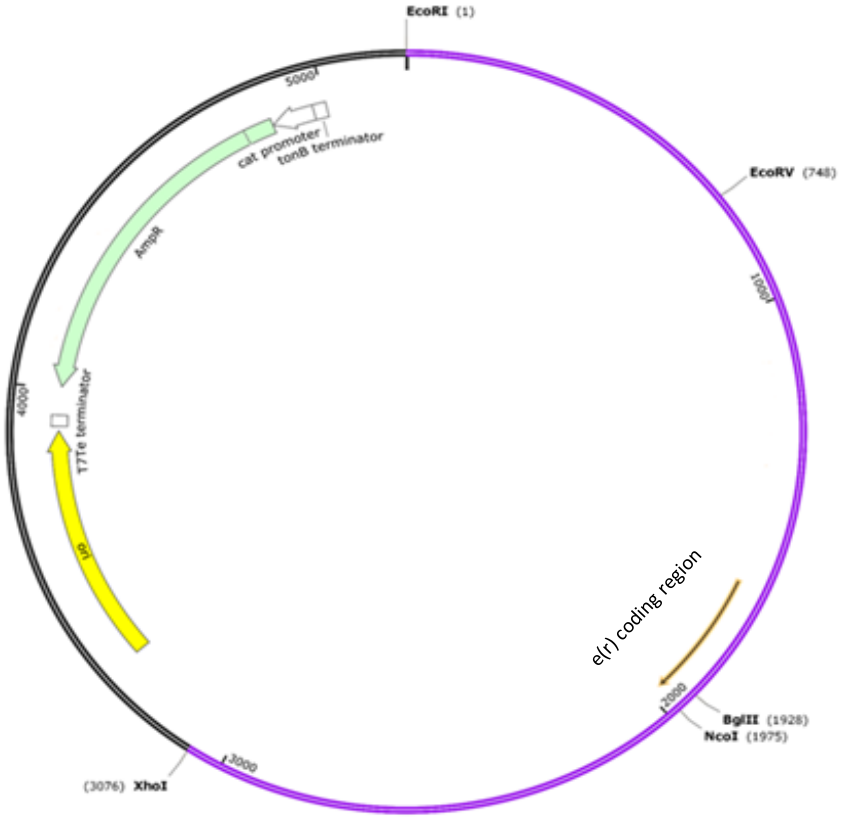

Figure 7. Shows the empty *e(r)* p-SMART vector + the *e(r)* coding region. This particular image depicts what the empty *e(r)* p-SMART vector + T4A mutant would look like. The mutant coding region must be ligated into the NcoI site at the end of the coding region, after the stop codon, in the correct orientation. The empty *e(r)* p-SMART vector being inserted with the S2A mutant coding region had NcoI sites at both the beginning and end of the coding region. The compatible sticky ends created by enzymes PciI and/or NcoI allowed the mutant *e(r)* coding regions to be ligated into the NcoI site the empty *e(r)* pSMART.

Part III: Insertion of mutated *Drosophila e(r)* gene into the final vector (pattB)

The isolated 3.0-kb *Drosophila e(r)* mutant was inserted into a pattB plasmid that had been cut with EcoRI and XhoI. *E. coli* cells were transformed with the recombinant plasmid and grown on ampicillin agar plates. Individual colonies were picked from the plates and used to create 1.5-ml cell cultures that were placed on a shaker-incubator at 250 rpm and 37°C overnight. The recombinant plasmids were isolated from the cells by performing a lysozyme mini-boiling prep on the cultures. An EcoRI-XhoI digest was done on the recombinant plasmids, and gel electrophoresis was run to confirm that the desired 3.0-kb mutant DNA fragment had been successfully ligated into the pattB vector. The colony that was determined to contain the successful recombinant plasmid was re-streaked onto a new plate and used to prepare a larger culture to amplify the amount of recombinant plasmid.

The recombinant plasmids were purified from the cells and two digests, an EcoRI-XhoI double digest and an EcoRV digest, were prepared for the recombinant plasmids. Gel electrophoresis was performed on the digested plasmids to confirm the presence of the desired recombinant plasmid in the large culture. Based on the position of the restriction sites for EcoRI, XhoI, and EcoRV on the pattB vector, as can be seen in Fig. 6, we would expect successfully recombined pattB plasmids to produce two fragments, sized 7.5-kb and 3.0-kb, when cut with EcoRI-XhoI, and three fragments, sized 2.0-kb, 2.9-kb, and 5.6-kb, when cut with EcoRV. The results of this gel electrophoresis is shown in Fig. 7.

Figure 8. Shows a pattB vector that has the 3.0-kb *e(r)* gene (purple section) with the mutation in the coding region (highlighted arrow) inserted into it. The pattB vector alone is about 7.5-kb, making the recombinant plasmid about 10.5-kb total. Attachment site attB (necessary for recombination into Drosophila genome) and the mini-white gene (codes for expression of red eyes in Drosophila) are labeled. These regions help identify the organisms that contain the recombinant plasmid. Restriction enzyme sites of interest are also labeled.

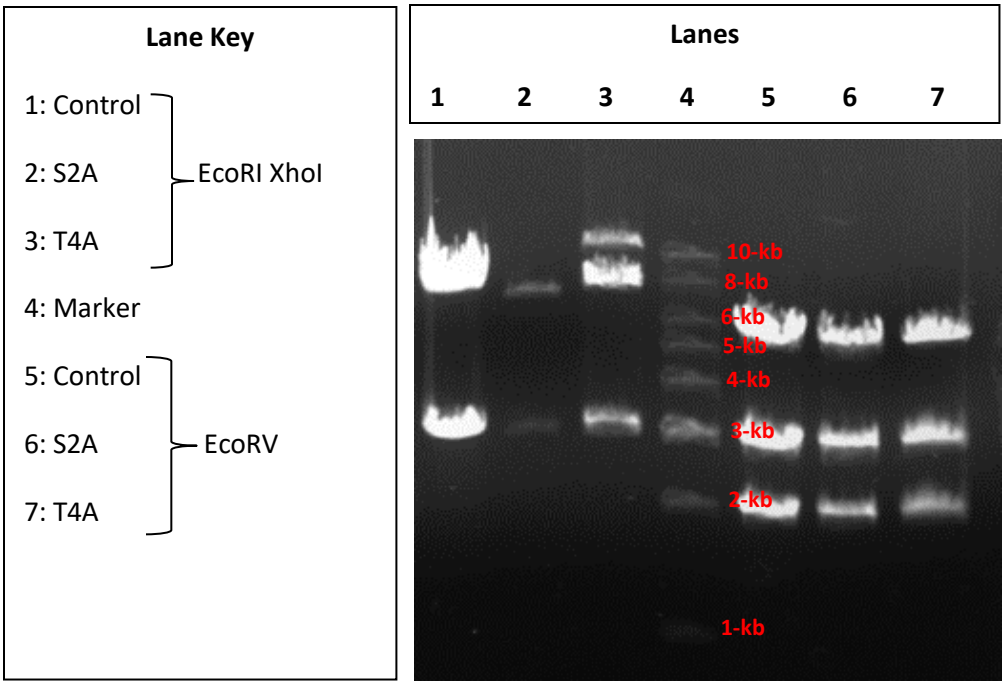

Figure 9. Shows the results of a gel electrophoresis performed on sample to confirm the successful ligation of the 3.0-kb fragment into the pattB vector. Lane markers are above the gel image, and the lane key is to the left. The control samples were a clone of a different mutant used in the past that, when cut with these same enzymes, is known to produce identical fragments as S2A and T4A. Row 4 is the marker lane, used to compare the bands, with the known band sizes indicated in red.

The gel indicated the mutants were cloned successfully, as they produced the same bands as the controls when cut with both EcoRV and with the EcoRI-XhoI double digest. In the EcoRI-XhoI double digest, the T4A mutant (lane 3) does show a minor additional band at 10.5-kb, which is the only deviation from the control. Since EcoRI-XhoI produces a 3.0-kb fragment, and the pattB plasmid is approximately 7.5-kb, the recombinant plasmid turns out to be 10.5-kb total. This indicates that there was an accidental single digest in which only one of the two enzymes cut some of the plasmids. Other than that, the EcoRI-XhoI double digest produced our expected fragment sizes (3.0-kb and 7.5-kb) for all three plasmids samples. The EcoRV digest also produced the expected fragment sizes (2.0-kb, 2.9-kb, and 5.6-kb) for all three plasmids.

The isolated and purified plasmid DNA was then sent to The Best Gene Company to be inserted into chromosome 3 of the *Drosophila*. The company selected for the transformants (flies containing the integrated transgene), and these stocks were sent back to us.

Part IV: Phenotypic analysis of mutant *Drosophila Melanogaster*

When the transformed flies were received back from The Best Gene Company, genetic crosses of *Drosophila* containing the mutant *e(r)* alleles and null *e(r)* mutations were established and performed. The null *e(r)* mutations were $e(r)^{27\text{-}1}$ and $Notch^{nd\text{-}p}$ $e(r)^{37\text{-}6}$. The $e(r)^{27\text{-}1}$ null *e(r)* mutation is associated with low viability (25-40%) in *Drosophila*, and the $Notch^{nd\text{-}p}$ $e(r)^{37\text{-}6}$ null *e(r)* mutation is associated with lethality. All the *Drosophila* with these genetic crosses that were analyzed were kept at 25°C and fed a yeast-glucose medium. Resulting from the crosses, only female offspring with a dark body and round, red eyes and male offspring with a light yellow body and round, red eyes were counted. Examples of *Drosophila* with the desired phenotypes are depicted in Fig. 8. The mini-white gene included in the pattB + e(r) insert vector causes the *Drosophila* containing this vector to have red eyes. This is significant, because red eyes are used to identify which flies have the mutant transgene. In the crosses, flies without the mutant transgene will have white eyes. *Drosophila* with any other phenotypes, including kidney-shaped eyes, white eyes, males with dark bodies, or females with light bodies are excluded because it indicates that they did not acquire the transgene through the cross.

Figure 10. Depicts examples of *Drosophila* with the desired phenotypes. For males (left): light yellow body and round red eyes. For females (right): dark body and round red eyes. These phenotypes indicate their genotype contains the mutant transgene.

One male with red eyes, indicating his DNA contains the transgene, is placed with females from a heterozygous stock of one of the phenotypic mutations we were testing. This heterozygous stock is maintained because an *e(r)* deletion in both sex chromosomes is lethal for the *Drosophila*, but if there is a deletion in only one X chromosome, the wild-type gene on the other X chromosome takes dominance, allowing these female *Drosophila* to survive. The

surviving males in the stock will have the X chromosome with the wild-type *e(r)* gene. This cross is depicted by the Punnett square in table 2. The males' DNA contains normal sex chromosomes. However, it also contains the *e(r)* transgene on both copies of one of their autosomes (chromosome 3), which is how male offspring with an *e(r)* deletion in their single X chromosome allows ERH to still be functional, and consequently demonstrating that the mutated copy of *e(r)* on an autosome is still functional. *Drosophila* males rescued by the transgene will have red eyes. The female offspring always survive since at least one of their two X chromosomes will contain the wild-type, functional *e(r)* gene, from the male. Under normal circumstances in nature, there is an equal chance for offspring to be male or female, resulting in a theoretical 1:1 male to female ratio. However, in this study, all females that inherit the deletion will survive, while the same cannot be said for the males, changing the theoretical male to female ratio. The number of red-eyed females compared to viable, red-eyed males can expose whether a mutation rescued an *e(r)* deletion.

		Heterozygous Female Stock X*e(r)*$^{-}$ / X^{+} *B*; +/+	
		X*e(r)*$^{-}$; +	X^{+} *B*; +
Male with *e(r)* gene mutation X^{+}/Y; *e(r)*transgene/*e(r)* transgene	X^{+}; *e(r)* transgene	Potential Female A: **X*e(r)*$^{-}$ /X^{+}; +/*e(r)* transgene**	Potential Female B: **X^{+}*B*/ X^{+}; +/*e(r)* transgene**
	Y; *e(r)* transgene	Potential Male A: **X*e(r)*$^{-}$ /Y; +/*e(r)* transgene**	Potential Male B: **X^{+}*B* /Y; +/*e(r)* transgene**

Table 2. Punnett square depicting the cross of a heterozygous female with the *e(r)* deletion and a male with the *e(r)* transgene of interest. This is a generalized example and can be applied to each mutant phenotype tested ($e(r)^{27\text{-}1}$ single mutant and $Notch^{nd\text{-}p}$ $e(r)^{37\text{-}6}$ double mutant). Shows the possible combinations of gametes to create different progeny. The following format is used to indicate the contents of the gametes and progeny: sex chromosome(s); autosome allele(s). Progeny of interest are female A and male A to determine if each mutation can rescue. The progeny of interest have red eyes, because of the mini-white gene in the pattB vector. The "B" progeny also have red eyes due to the presence of the mini-white gene, by they have small, kidney-bean-shaped eyes caused by the dominant *B* mutation. Thus the A flies can be distinguished from their B counterparts. The A flies were counted in each cross.

Results

The cross summarized in table 2 was carried out with each *e(r)* mutant and both of the *e(r)* mutant test stocks. The results of the $e(r)^{27\text{-}1}$ mutation cross with each of the four amino acid changes tested are shown in Table 3. The results of the $e(r)^{37\text{-}6}$ $Notch^{nd\text{-}p}$ double mutation cross with each of the four amino acid changes tested are shown in Table 4. In both tests, a control was measured to know what to anticipate for the ratio of males to females. The controls were the same cross but with an unmutated, wild-type transgene for the *Drosophila e(r)* gene, and an empty transgene. The empty *e(r)* control acts as a lower baseline for assessing the ERH activity of the mutant transgene, as it produces no ERH activity. The wild-type *e(r)* control, on the other hand, provides an example of full ERH activity, giving us an upper baseline for assessing the ERH activity of the mutant transgene. These baselines are what we compare the mutant ratio results to in order to determine if the *e(r)* mutants induce rescue, or alternatively, a "knocking out" of wild type activity.

Table 3. $e(r)^{27\text{-}1}$ mutation cross results

Transgene	**Males** ($X^{e(r)27\text{-}1}$/Y + transgene)	**Females** ($X^{e(r)27\text{-}1}/X^{+}$ + transgene)	**Ratio** (males:females)
Empty	134	441	0.304
Wild-type	492	546	0.901
S2A	150	209	0.719
T4A	139	170	0.820
H3A	24	97	0.248
I5D	12	71	0.169

Table 3. A cross of the $e(r)^{27\text{-}1}$ deletion with each transgene was performed. The number of males and females with red eyes were counted, and a male to female ratio was calculated. The control tests are shaded in grey.

The $e(r)^{27\text{-}1}$ cross results (table 3) show rescue from S2A and T4A, but not from H3A or I5D. The controls display signs of slightly lower viability in males, with the empty transgene control having a male to female ratio of 0.304. The wild-type control for this cross, however, is the closest seen to a perfect theoretical one to one ratio of males to females, at 0.901. Mutants

S2A and T4A seem to rescue males from the lower viability of this cross, as seen by the ratios 0.719 for S2A and 0.820 for T4A. These ratios aren't quite as close to one as the wild-type control, but are certainly closer than that of the empty transgene control. Mutants H3A and I5D, on the other hand, appear to knock out activity, based on their lower ratios of 0.248 and 0.169, both of which are lower than that of the empty transgene control.

Table 4. *Notch*$^{nd\text{-}p}$ and *e(r)*$^{37\text{-}6}$ mutation cross results

Transgene	**Males** ($X^{e(r)37\text{-}6\ Nnd\text{-}p}$/Y + transgene)	**Females** ($X^{e(r)37\text{-}6\ Nnd\text{-}p}$/$X^{+}$ + transgene)	**Ratio** (males:females)
Empty	0	295	0.000
Wild-type	130	570	0.228
S2A	37	139	0.266
T4A	77	218	0.353
H3A	3	184	0.016
I5D	2	204	0.010

Table 4. A cross of the *Notch*$^{nd\text{-}p}$ *e(r)*$^{37\text{-}6}$ double mutant with each transgene was performed. The number of males and females with red eyes were counted, and a male to female ratio was calculated. The control tests are shaded in grey.

The *Notch*$^{nd\text{-}p}$ *e(r)*$^{37\text{-}6}$ cross results (table 4) also show rescue from S2A and T4A, but not from H3A or I5D. The controls from this cross show signs of lethality in males, with the empty transgene producing not a single surviving male, and with the wild-type producing a still rather low ratio, at 0.228. The wild type transgene rescues *e(r)*$^{37\text{-}6}$, but the males still have the *Notch*$^{n\text{-}dp}$ mutation, which keeps the male viability lower. Mutants S2A and T4A seem to again show rescuing activity of males from the lethality of this cross, as seen by the ratios 0.266 for S2A and 0.353 for T4A. These ratios are higher than the wild-type control, indicating further rescue, and obviously significantly higher than the empty transgene control, which produced zero surviving males. Mutants H3A and I5D also show similar activity in this cross as they did in the *e(r)*$^{27\text{-}1}$ cross. Again, they appear to knock out wild-type activity, based on their lower ratios of 0.016 for H3A and 0.010 for I5D, both of which show a slight rescue from the total lethality of males seen

with the empty transgene control, but are quite lower than that of the wild-type transgene control, and the S2A and T4A mutant transgene results.

Discussion

In general, both crosses resulted in S2A and T4A retaining wild-type activity, and H3A and I5D knocked out wild-type activity. One likely reason for the differences in the activity of S2A and T4A versus H3A and I5D is the differences in the severity of the structural changes brought on by these mutations: the amino acid replacements in S2A and T4A are more similar to the structure of the wild-type amino acids, and therefore don't change the shape or function of the ERH protein too much, while the H3A and I5D mutants on the other hand are more significant alterations to the wild type amino acid and therefore the protein shape and function.

The S2A and T4A mutations were designed to prevent phosphorylation. Since the results indicate that these mutants were still able to rescue, we can deduce that phosphorylation at these particular amino acids are not necessary for the protein to function properly.

Conclusion

Research up to this point has identified three major areas of interest in ERH: Two areas to do with its own functionality, along with one area of interest in the Notch signaling pathway. My results have shown that mutants S2A and T4A retain wild-type activity, while mutants H3A and I5D knock out activity. A previous study by another student also showed that H3A and I5D in human ERH knock out activity, so the findings of this study have further supported that conclusion.

It is hypothesized that the differences in the structural changes to the protein brought about by the different mutations may provide an explanation for the differences in the activity of S2A and T4A versus H3A and I5D. The decreased viability of H3A and I5D, as compared to the wild-type control, in the $Notch^{nd\text{-}p}$ $e(r)^{37\text{-}6}$ cross could potentially be explained by either an interference with the Notch signaling pathway or an overall decrease in protein activity. Overall, my results have further indicated that the N-terminus of ERH, is indeed necessary for its proper function and that this function may be evolutionarily conserved.

Future Works

Future research performed by Dr. Tsubota will further determine amino acid sites necessary for human ERH function, specifically S2A and T4A as unphosphorylated mimics of S2 and T4. It is possible that either S2 or T4 must be phosphorylated for ERH to be active. While S2A or T4A by itself does not knock out activity, it is possible that double mutant S2A T4A would be inactive. Dr. Tsubota is also looking into a new human ERH mutant, T4Q, which is a more dramatic mutation than T4A, and is hypothesized to result in a more dramatic effect on ERH activity. Research to further determine amino acid sites necessary for ERH function may prove useful in developing future cancer treatments and drugs. By terminating ERH function, cancer cells will struggle to survive and multiply. Whether ERH function termination can be caused by a single or multiple amino acid change to one of the other regions such as the α-helices, β-strands, loops, or C-terminus region will continue to be further researched.

Acknowledgments

I would like to acknowledge my thesis mentor and research director, Dr. Stuart Tsubota, for his guidance and previous research. I would also like to acknowledge my fellow lab mate, Brianna Dudley, for her construction of the H3A and I5D human ERH mutations that I analyzed in the crosses with my S2A and T4A Drosophila ERH mutants. Finally, I would like to acknowledge the SUNY Brockport Honors College for their support and encouragement of student research. I would like to especially thank the Honors College director, Dr. Austin Busch, and associate director, Dr. Pamela Beach, for their help and reassurance throughout my project.

References

Arai, R., Kukimoto-Niino, M., Uda-Tochio, H., Morita, S., Uchikubo-Kamo, T., Akasaka, R., Etou, Y., Hayashizaki, Y., Kigawa, T., Terada, T., Shirouzu, M., Yokoyama, S. 2005. Crystal structure of an enhancer of rudimentary homolog (ERH) at 2.1 Å resolution. *Protein Sci Publ Protein Soc.* 14:1888–1893.

Fujimura, A ., Kishimoto, H., Yanagisawa, J., Kimura, K. 2012. Enhancer of rudimentary homolog (ERH) plays an essential role in the progression of mitosis by promoting mitotic chromosome alignment. *Biochemical and Biophysical Research Communications*. 423 (3): 588-592.

Gelsthorpe,, Pulumati, M., McCallum, C., Dang-Vu, K., Tsubota, S.I. 1997. The putative cell cycle gene, enhancer of rudimentary, encodes a highly conserved protein found in plants and animals. *Gene*. 186:189-195.

Hedges S.B., Dudley J., Kumar S. 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics.* 22:2971–2972.

Kops, G., Weaver, B., Cleveland, D. 2005. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer.* 5: 773-785.

Tsubota, S. and P. Schedl. 1986. Hybrid dysgenesis induced revertants of insertions at the 5' end of the *rudimentary* gene in *Drosophila melanogaster*: transposon-induced control mutations. *Genetics* 114: 165-182.

Tsubota, S.I., Vogel, A.C., Phillips, A.C., Ibach, S.M., Weber, N.K., Kostrzebski, M.A., Spencer, S.A. 2011. Interactions between enhancer of rudimentary and Notch and deltex reveal a regulatory function of enhancer of rudimentary in the Notch signaling pathway in Drosophila melanogaster. *Fly* (Austin). 5(4):275–284.

Tsubota, S., Ryan, T., Rizzo, N., Hing, H. 2016. The human and Drosophila ERH are functionally equivalent: evidence from transgenic studies. *Internal Medicine Review.* 2 (8): 1-14.

Wan, C., Temple, W., Liu, Z., Wang, B., Rose, R.B. 2005. Structure of the Conserved Transcriptional Repressor Enhancer of Rudimentary Homolog. *Biochemistry.* 44 (13): 5017-5023.

Wojcik, E., Murphy, A.H., Fares, H., Dang-Vu, K., Tsubota, S.I. 1994. Enhancer of Rudimentary (p1), E(r)(p1), a Highly Conserved Enhancer of the Rudimentary Gene. *Genetics.* 138:1163–1170.

Zafrakas, M., Losen, I., Knuchel, R., Dahl, E. 2008. Enhancer of the rudimentary gene homologue (ERH) expression pattern in sporadic human breast cancer and normal breast tissue, Enhancer of the rudimentary gene homologue (ERH) expression pattern in sporadic human breast cancer and normal breast tissue. *BMC Cancer*. 8:145–145.