THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

THE OVEREXPRESSION OF COMPLEXES AND SUBCOMPLEXES OF THE DROSOPHILA MELANOGASTER NELF PROTEINS

NATHAN CHRISTOPHER HAZI

Summer 2010

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biochemistry and Molecular Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Song Tan, Ph.D.
Associate Professor of Biochemistry and
Molecular Biology
Thesis Supervisor

David Gilmour, Ph.D.
Professor of Biochemistry
Honors Advisor

Scott Selleck, Ph.D., Department Head
Department of Biochemistry and
Molecular Biology

*Signatures are on file in the Schreyer Honors College.

Abstract

The eukaryotic transcription cycle is controlled by a dynamic system of essential transcription machinery. After the preinitiation complex, including RNA Polymerase II, begins to transcribe nascent pre-mRNA, transcription elongation of many genes may be controlled by the negative elongation factor (NELF). NELF has been found to inhibit elongation in some cases, or to stall elongation, allowing a rapid induction response. NELF is a complex composed of 4 subunits: NELFA, NELFB, NELFD and NELFE. I attempted to clone all four *Drosophila* NELF subunits into the pST44 polycistronic expression plasmid in hopes of expressing the complex, subcomplexes and subunits in *E. coli.* I was able to create a tricistronic expression plasmid including dNELFD, dNELFB and HISdNELFE, but I could not subsequently incorporate the dNELFA gene into this plasmid. Each of the individual subunits and some subcomplexes were expressed at varying temperatures to determine optimal expression and solubility levels. The addition of a 6x Histidine affinity tag onto dNELFE allowed for effective affinity purification of HISdNELFE. Large-scale expression and purification of HISdNELFE was attempted, although final purification was not achieved due to truncations of the full-length protein that proved difficult to remove.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Transcriptional Regulation

The regulation of the transcription cycle through RNA Polymerase II (Pol II) is critically important for almost every facet of cellular function. The mRNA transcription cycle has three main phases: initiation, elongation and termination. Each phase can be further divided into substeps (Figure 1.1) to understand the mechanisms of transcriptional regulation. Before any transcription can occur, the essential transcription machinery must assemble around the promoter to form the preinitiation complex (PIC) (Orphanides *et al.*, 1996).

The PIC is composed of Pol II and general transcription factors that are bound to DNA near the transcriptional start site. As the PIC is assembled, the packaging of nearby chromatin is altered to be more permissive of upcoming transcription. Transcription activators that recognize specific promoter or enhancer sequences, together with protein complexes that remodel and modify chromatin, allow the formation of the PIC. Once preinitiation steps are completed, initiation may occur and the first phosphodiester bonds in the newly synthesized RNA transcript are formed. However, initiation is often a complex process, as transcription up to the +9 position may have multiple rounds of abortive initiation (Saunders *et al*., 2006). If transcription progresses through the +9 position and abortive initiation, it has reached the promoter escape transition. At this point, transcription becomes prone to other problems such as arrest, backtracking or

slippage. Between +20 and +50, promoter-proximal pausing may occur for some genes, and transcription is paused until induction resumes elongation. Once elongation has resumed, the nascent mRNA will grow, be capped, and upon completion, transcription is terminated.

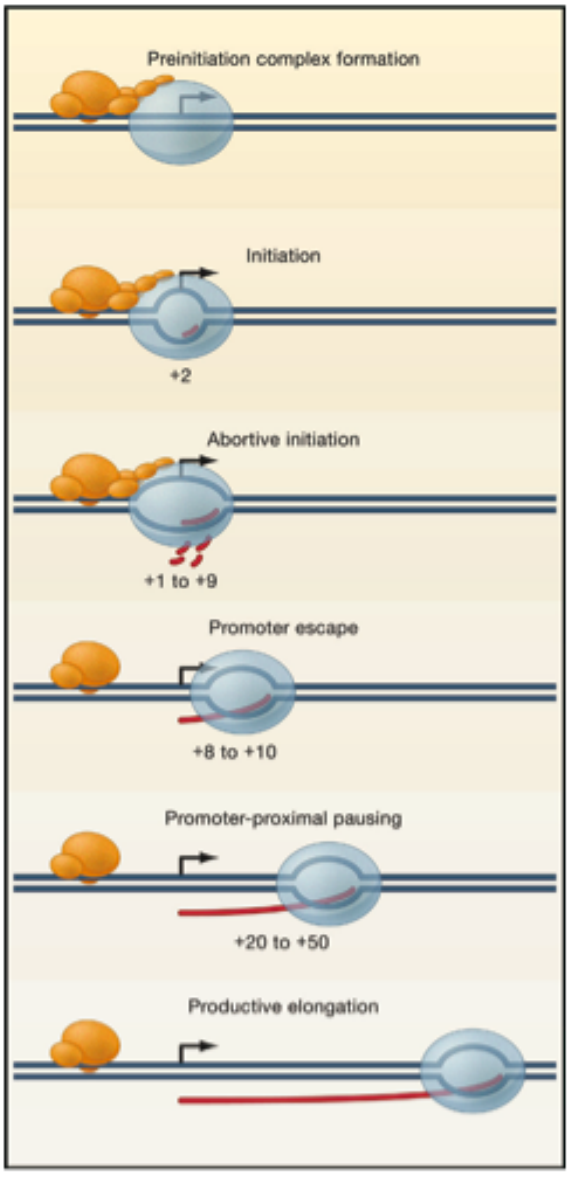

Figure 1.1: The stages of transcription initiation. Pol II is displayed in blue and general transcription factors in orange. The arrow marks the start site of initiation and numbers indicate the region on the template DNA that has been reached by the Pol II active site (Saunders et al., 2006).

Recent genome-wide studies are changing the view that Pol II recruitment is the rate-limiting and most regulated step for the transcription of most eukaryotic genes (Ptashne and Gann, 1997). Studies have elucidated regulation during the elongation phase and identified this phase as an important and common stage for regulation of significant sets of eukaryotic genes.

In a genome-wide survey of stationary phase *Saccharomyces cervisiae*, Pol II occupied hundreds of early growth response genes without active transcription (Radonjoc *et al.*, 2005). Similarly, a genome-wide survey of human fibroblasts indicated the presence of assembled PIC's on numerous inactive genes (Kim *et al*., 2005). Elongation regulation in embryonic and differentiated mammalian cells was demonstrated by comparing mRNA transcript levels to initiation hallmarks such as Pol II occupancy and nucleosomes with H3K4me3 and H3K9,14Ac modifications (Guenther *et al.*, 2007). These hallmarks of transcription initiation were observed at promoters for the majority of protein encoding and transcriptionally inactive genes, regardless of transcription activity. Together, these findings suggest that preinitiation regulation is not the rate-limiting step for many genes.

1.2 A System of Transcriptional Control

Pol II is regulated during elongation by a system of proteins that is not fully understood. Two factors, DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) have been shown to inhibit elongation by Pol II (Yamaguchi *et al.,* 1999). Inhibition by these factors is alleviated by the positive elongation factor b (P-TEFb)

kinase, which phosphorylates the C-terminal domain of Pol II (Wada *et al.,* 1998; Yamaguchi *et al.,* 2002). The kinase activity of P-TEFb can be inhibited by the nucleoside analog DRB (5,6-dichloro-1-ß-D-ribofuranosylbenzimidazole). Thus DSIF and NELF regain the ability to inhibit Pol II elongation in the presence of DRB. This system of counteracting effects is likely an important regulation system that governs elongation. A well-understood example of promoter-proximal pausing has been demonstrated in *hsp70*. NELF and DSIF were shown to cause promoter-proximal pausing in the *hsp70* gene in *Drosophila* under normal conditions (Wu *et al.,* 2003). Induction by heat shock causes P-TEFb association with *hsp70* (Boehm *et al*., 2003), the dissociation of NELF but not DSIF, and continued elongation of the transcript.

NELF may have more significance than just as an elongation inhibitor. In estrogen-stimulated genes, NELF was found to counteract the induction by estrogen (Aiyar *et al.,* 2004). While NELF-mediated elongation inhibition is well documented, emerging evidence also suggests that it may actually enhance the expression of a large number of genes. A genome-wide analysis in *Drosophila* demonstrated that upon NELF depletion, roughly one-third of NELF-targeted genes were transcriptionally up-regulated, while roughly two-thirds of NELF-targeted genes were down-regulated (Gilchrist *et al*., 2008). By utilizing the stalled Pol II initiation complex, NELF maintains a permissive chromatin structure that allows transcription to proceed easily upon induction. The removal of NELF allowed the chromatin structure to increase nucleosome density and caused a decrease in the H3K4me3 active chromatin mark.

Through basic Pol II binding and promoter-proximal pausing, NELF inhibits some transcription, while allowing rapid transcription upon induction for other genes. Elongation regulation through NELF has proven to be a dynamic, efficient and frequently used source of gene regulation.

1.3 Molecular Characterization of the NELF Complex

NELF is a complex of four proteins that is reasonably well conserved throughout organisms in which it is present. NELF is found in vertebrates and insects, but no orthologs have been discovered in other organisms (Narita *et al*., 2003). It is composed of four subunits, A, B, C/D, and E. NELFB and NELFD are well conserved, while NELFA and NELFE show nonconserved regions between N- and C-terminal regions.

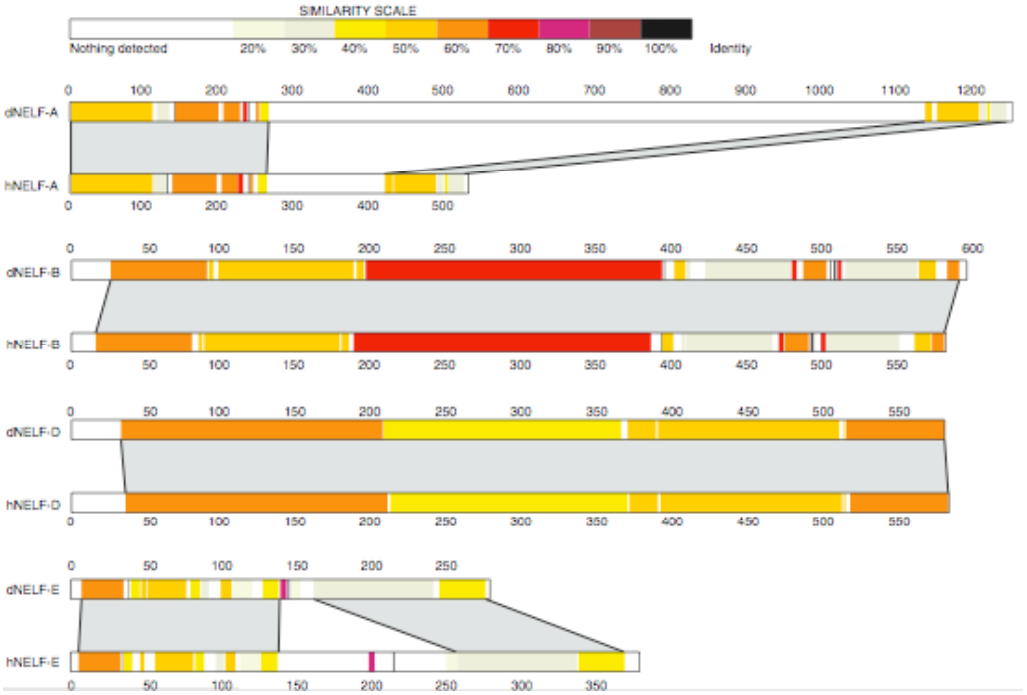

Figure 1.2: Sequence identity between human and *Drosophila* NELF genes (Wu *et al*., 2005).

NELFB is identical to the protein named COBRA1 (Cofactor of BRCA1), named because it was found to associate with the protein product of *BRAC1*, which produces a breast cancer susceptibility protein (Ye *et al., 2001*). NELFB forms a coiled-coil structure that interacts with NELFE. NELFC/D are translational variants of the same mRNA. NELFC has nine additional N-terminal amino acids compared to NELFD, the result of using an alternative start codon. In NELF complexes, only four subunits are present: A, B, C or D, and E (Narita *et al*., 2003). NELF complexes containing either NELFC or NELFD have not demonstrated different activity, and it is likely that the N-termini differences of these subunits are not functional (Narita *et al*., 2003).

NELFA is encoded by *WHSC*2, Wolf-Hirschhorn syndrome candidate 2, which has been linked to developmental defects (Wright *et al.,* 1999). The N-terminal segment of NELFA shares sequence similarity with hepatitis delta antigen, a viral polypeptide that binds Pol II (Yamaguchi *et al.,* 2001). NELFE has an N-terminal leucine zipper motif, while the C-terminal domain has a RNA recognition motif (RRM). This RRM is necessary for NELF function, as the RRM binds to some RNA sequences (Yamaguchi *et al.,* 2002).

1.4 Orthologous NELF Complexes

Human and *Drosophila* NELF appear to be functionally similar because depletion of either desensitizes transcription to DRB (Wada *et al* 1998; Wu *et al.,* 2005). Sequence similarity is high between orthologs of NELFB and NELFD (Figure 1.2), with several

regions of >50% sequence identity throughout the entire polypeptides. *Drosophila* NELFA contains a nonconserved region between amino acids 300-1100 that contains a series of poly-glutamine, poly-asparagine, poly-threonine and poly-alanine that is not contained in human NELFA.

Human and *Drosophila* NELFE share sequence homology, especially at the N- and C-terminal region. Human NELFE has an RNA binding motif that is essential for NELF-mediated elongation inhibition (Narita *et al*., 2003). This region (167-232 amino acids in dNELFE) has 30% sequence identity between the orthologs. There are two significant differences between human and Drosophila NELFE. The region spanning amino acids 196-242 in hNELFE, comprised mainly of alternating arginine and aspartic acid residues, and of unknown function, is absent in dNELFE. Also absent from dNELFE are serine residues 181, 185, 187 and 191 of hNELFE, which are known to be phosphorylated *in vitro* by P-TEFb to reduce the RNA-binding ability of hNELFE (Fujinaga *et al.*, 2004).

The NELF complex has been reconstituted from recombinant proteins to form a functionally active complex *in vitro* (Narita *et al*., 2003). Narita *et. al*. were able to purify the human NELF complex from insect cell lysates using one-step affinity chromatography. The NELF subunits were multiply infected into insect Sf9 cells by recombinant baculoviruses. Immunoprecipitation interaction studies revealed that NELFE directly interacts only with NELFB. NELFB interacts with only NELFE and NELFD, and NELFA only directly interacts with NELFD. This information, and information of the function of each subunit, led Narita *et. al*. to theorize that NELFD and NELFB act to

bring NELFA and NELFE together (2003).

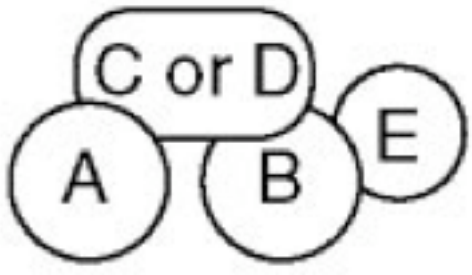

Figure 1.3: Schematic representation of the interactions of the NELF subunits (Narita et al., 2003).

1.5 Polycistronic Expression Plasmid

The expression of multiple subunits of the NELF complex is enabled by the pST44 polycistronic expression system developed in the Tan lab (Tan *et al.,* 2005). This system, based on the pST39 system, allows up to four genes to be inducibly expressed from a single plasmid under the control of a single T7 promoter. Each of the four genes can be cloned into the pST44 vector using unique transfer cassettes with unique restriction sites. The use of transfer cassettes also allows easy addition of affinity tags to the N- or C-terminal end of a protein within the polycistronic plasmid. Once inserted into the pST44 plasmid, each gene has a translational enhancer, Shine-Dalgarno sequence and termination site. All expression plasmids of this system contain the pMB1 ColE1 origin of replication and ampicillin resistance conferred by the ß-lactamase gene (Tan *et al,* 2005).

1.6 Aims of This Research

The success of Narita *et al*. proved recombinant NELF could be expressed and purified to form a functionally active complex (2003). I attempted to clone all four dNELF subunits into a polycistronic expression plasmid to express recombinant dNELF in *E. coli* cells. With the completed polycistronic expression plasmid, I hoped to express the NELF complex, subcomplexes, and subunits. Purification of the NELF complex, subcomplexes and subunits would potentially provide reagents for biochemical transcription assays and pure material for molecular structure determination by X-ray crystallography.

2. Materials and Methods

2.1 Bacterial Strains Used

These strains of *E. coli* were used as host strains in cloning and protein expression:

TG1: Δ (lac-pro), supE, thi, hsdD5/F', traD36, $proA^{+}B^{+}$, $lacI^{q}$, lacZΔ, m15

BL21(DE3)pLysS: B F^{-}, dcm, ompT, hsdS($r_B^{-}m_B^{-}$), gal[lambda](DE3)[pLysS Cam^{r}]

HB101: F^{-}Δ(gpt-proA)62 leuB6 glnV44 ara-14 galK2 lacY1 Δ(mcrC-mrr) rpsL20(Str^{r}) xyl-5 mlt-1 recA13

2.2 Polymerase Chain Reaction (PCR) Amplification

PCR amplification of single genes was performed to amplify a gene or to add restriction sites outside of the coding region for further cloning procedures. The PCR mixture contained 10 µl 10xThermo Pol Buffer, 10 µl 2.5 mM dNTP, 0.5 µl of 10 ng/µl template plasmid, 5 µl 10 µM forward primer, 5 µl 10 µM reverse primer, 0.5 µl 2 units/µl Pfu polymerase and water to a final volume of 100 µl. The mixture was placed in a thermocycler with this general cycle: 2 minutes of 95ºC followed by 5 cycles of (30 seconds of 95ºC, 30 seconds of T_m-5°C, T_{ext} for 75ºC), then 20 cycles of (30 seconds of 95ºC, 30 seconds of 60°C, T_{ext} of 75ºC), then 3 minutes of 75ºC, where T_m is the lower

temperature of melting for the primers used and T_{ext} is an extension time of 1 minute per 1000 base pairs (bp) of the expected amplification product. After thermocycling was completed, the PCR product was confirmed by combining 5 µl of the PCR mixture and 1 µl of 6xGLB [0.25% (w/v) Bromophenol Blue, 0.25% (w/v) Xylene Cyanol, 30% (v/v) glycerol, 60 mM EDTA] to a final concentration of 1xGLB and analyzed for 40 minutes on a 1% (w/v) high gelling temperature (HGT) agarose gel at 125 V. Ethidium bromide [0.33 µg/ml in the 0.5xTBE buffer] staining allowed viewing of the gel on a long wavelength UV transilluminator. While 40 minutes of electrophoresis was standard with all agarose gel electrophoreses, the sample was exposed to 125 V for a shorter time period with small DNA fragments (<1000 bp), or for a longer time period with larger DNA fragments (>3000 bp) to ensure DNA band separation.

Upon verifying the size of the PCR product, the remaining PCR mixture was Phenol/CIA (CIA is 24 parts chloroform and 1 part isoamyl alcohol) extracted, CIA extracted and ethanol precipitated to remove unwanted reagents and proteins. The precipitated pellet was resuspended in 30 µl TE(10, 0.1) [10 mM Tris-HCl pH 8.0, 0.1 mM EDTA].

2.3 PCR-Based Site-Directed Mutagenesis

The PCR-based site-directed mutagenesis procedure was adapted from the QuikChange method from Stratagene (La Jolla, CA). It was used to remove or create an internal restriction site by introducing a silent mutation, or to revert a base mutation introduced by PCR amplification. In the case of removing or adding an internal restriction site, the

restriction site was used to screen the final PCR mutagenesis product. Two oligonucleotide primers were designed to introduce each mutation. The PCR mixture contained 17.6 µl water, 2.5 µl 10xThermo Pol Buffer, 2.5 µl of 2.5 mM dNTP, 0.5 µl of 10 ng/µl template plasmid, 0.7 µl of 10 µM mutagenesis forward primer, 0.7 µl of 10 µM mutagenesis reverse primer, and 0.5 µl of 2 units/µl Pfu polymerase. This mixture was placed in the thermocycler with this general cycle: 2 minutes of 95ºC followed by 12 cycles of (30 seconds of 95ºC, 1 minute of 55ºC, T_{ext} of 68ºC), then 15ºC where T_{ext} is an extension time of 1 minute per 1000 bp of expected PCR product.

After thermocycling was completed, selection of the desired PCR mutagenesis product was performed. An undigested sample of 2 µl of PCR mixture was set aside in a sterile Eppendorf tube. The remainder of the PCR mixture was digested by adding 0.5 µl of 10 units/µl DpnI and incubated for 1 hour at 37ºC. Following digestion, 2 µl of each digested and undigested sample were transformed separately into 100 µl competent TG1 cells. After growth at 37ºC, colonies from the DpnI digested plate were PCR screened to determine if the desired mutagenesis including the restriction site change had occurred. Colonies were restreaked, and colonies from the restreaked plate were chosen for plasmid preparation, restriction digestion, and sequencing.

2.4 Gel Purification

For each cloning construct to be created, insert and vector DNA were restriction digested by incubating the plasmid or PCR product with the appropriate restriction enzymes at

37ºC. The digestion mixture contained water, appropriate New England Biolabs restriction buffer, 0.1 mg/ml BSA, 3.3 mM DTT, and restriction enzymes (see Appendix A for reaction mixture recipes) (Tan *et al.*, 2005; Tan, 2001). Following digestion, 6xGLB was added to the digestion mixture to a final concentration of 1xGLB. The sample was analyzed on a1% (w/v) HGT agarose gel in 0.5xTBE buffer [45 mM Tris, 45 mM boric acid, 1 mM EDTA] at 125 V. Ethidium bromide staining allowed viewing of the gel on a long wavelength UV transilluminator. The desired DNA fragment was removed with a razor blade and placed in a 0.5 ml Eppendorf tube with a hole in the bottom punctured by a very hot needle. A small amount of siliconized glass wool was also placed in the bottom of the tube to retain the solid agarose during centrifuging. The 0.5 ml Eppendorf tube was placed into a labeled 1.5 ml Eppendorf tube and this apparatus was centrifuged for 5 minutes at 7,000 rpm (4,656 g) in a microcentrifuge (Sorvall Biofuge). By centrifuging, the desired DNA fragment was eluted in TBE buffer into the 1.5 ml Eppendorf tube.

2.5 Ligation

To ligate the gel purified insert and vector with corresponding sticky ends, a ligation mixture was created using water, T4 DNA ligase buffer, 5 mM DTT, gel purified vector DNA, gel purified insert DNA, and 4 units/µl T4 DNA ligase (see Appendix A for reaction mixture recipes). This mixture was allowed to incubate at room temperature for a minimum of 15 minutes. A concurrent control mixture lacking the gel purified insert DNA was also prepared.

2.6 Plasmid Transformation for 18ºC Competent Cells

Following ligation, both the ligation mixture (vector + insert) and the control mixture (vector only) were transformed into TG1, HB101, or BL21(DE3)pLysS *E. coli* competent cells. Prior to transformation, the competent cells previously stored at –80ºC were allowed to thaw on ice. For each 100 µl aliquot of competent cells, 2 µl of either ligation mixture or control mixture was added and allowed to incubate on ice for 15 minutes (only 0.5 µl was used with BL21(DE3)pLysS cells). Following this incubation, the cells were heat-shocked in a 42ºC water bath for 30 seconds, and then returned to an ice bath for 10 seconds.

To induce bacterial growth, the heat-shocked competent cells were combined with 0.5 ml 2xTY media [1.6% (w/v) bacto tryptone, 1% (w/v) yeast extract, 0.5% (w/v) sodium chloride] and placed in a 37ºC shaking incubator for 15 minutes. After incubation, 0.3 ml of the cell culture was plated onto TYE media plates [1.0% (w/v) bacto tryptone, 0.5% (w/v) yeast extract, 0.8% (w/v) NaCl, 1.5% agar (w/v)] with the appropriate antibiotic(s) [100 µg/ml ampicillin for TG1 cells, 15 µg/ml kanamycin for HB101 cells and 100 µg/ml ampicillin + 25 µg/ml chloramphenicol for BL21(DE3)pLysS cells] and incubated at 37ºC for 10-18 hrs (20-24 hrs for HB101 cells).

2.7 PCR Screening

To identify clones that had been transformed with the desired plasmid, a PCR screening procedure was used. With a sterile inoculating loop, a single colony was picked up and added to 100 µl water in a 1.5 ml Eppendorf tube and then restreaked onto a TYE media plate with the appropriate antibiotic(s). When possible, 8 (vector + insert) colonies were selected in this way for PCR screening. The re-streaked TYE media plates were incubated at 37ºC until distinct colonies were visible.

The cell suspensions in 1.5 ml Eppendorf tubes were then prepared for PCR screening by vortexing for 10 seconds. A PCR screening mixture was created using 10xThermo Pol Buffer, 0.25 mM dNTP, 0.5 µM forward primer, 0.5 µM reverse primer, 2 units/µl Pfu polymerase and water. The volume of each of these was dependent upon the number of colonies being screened (see Appendix A for full recipe). Forward and reverse oligonucleotide primers were selected to ensure that the desired insert had been ligated into the desired vector in the correct orientation. When possible, one vector-specific and one insert-specific primer were chosen to produce a PCR product under 2000 base pairs long. If primers could not be chosen specifically for the vector and the insert, two primers were chosen for the insert. When PCR screening for a site-directed mutagenesis that introduced or removed a restriction site, the PCR product was designed to contain the introduced or removed restriction site so that restriction digestion can verify the presence or absence of the site.

From each vortexed cell suspension, 1 µl was added to 19 µl of PCR screening mix and was placed in a thermocycler for 2 minutes of 95ºC followed by 20 cycles of (30 seconds of 95ºC, 30 seconds of T_m -10ºC, T_{ext} of 75ºC) then 15ºC. T_m is the melting temperature of the primers used and T_{ext} is an extension time of 1 minute per 1000 bp of the expected PCR product. Following thermocycling, 6xGLB was added to the digestion mixture to a final concentration of 1xGLB. The sample was analyzed on 1% (w/v) HGT agarose gel in 0.5xTBE buffer at 125 V for 40 minutes. Ethidium Bromide staining allowed viewing of the gel on a long wavelength UV transilluminator. Two clones displaying the desired PCR products were selected for plasmid preparation

2.8 100 ml Alkaline Lysis Plasmid Preparation

Each clone selected from the PCR screening was used to inoculate a 500 ml Erlenmeyer flask with 100 ml 2xTY media [1.6% (w/v) bacto tryptone, 1% (w/v) yeast extract, 0.5% (w/v) sodium chloride] containing the appropriate antibiotic(s) to retard the growth of untransformed bacteria. Flasks were incubated in a 37ºC shaking incubator for 16-20 hrs. The contents of the flasks were transferred to 250 ml centrifuge bottles and were centrifuged for 5 minutes at 4,000 rpm (4,204 g) (Heraeus rotor #7570 G) to pellet the cells. The supernatant was discarded, and the pellet was resuspended in 5 ml lysis buffer [50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA]. This cell suspension was transferred to a 50 ml Falcon tube, and cells were lysed by adding 10 ml of NaOH/SDS solution [0.2 M NaOH, 1% (w/v) SDS] and by shaking vigorously until clumps were no longer visible. After 5 minutes of incubation on ice, 10 ml cold KAc/HAc solution [5 M

KAc, 2.5 M HAc] was added, and the tube was once again vigorously shaken and incubated on ice for 5 minutes to precipitate chromosomal DNA. The tube was centrifuged at 4,000 rpm (4,024 g) for 3 minutes at 20ºC (Heraeus rotor #7570 G). Following centrifugation, the supernatant was poured through a scintered glass funnel to remove precipitates, and then transferred to a 50 ml round bottom polypropylene centrifugation tube. To precipitate plasmid DNA and RNA, 12.5 ml isopropanol was added and incubated at room temperature for 5 minutes. Following incubation, this mixture was again centrifuged at 13,000 rpm (16,060 g) for 5 minutes at 20ºC (Sorvall SS-34). The supernatant was discarded, and the pellet was resuspended in 0.5 ml 70% ethanol before being transferred to a 1.5 ml Eppendorf tube. The ethanol was removed by centrifuging at 13,000 rpm (16,060 g) for 1 minute at room temperature in a microcentrifuge (Sorvall Biofuge). The resulting nucleic acid pellet was resuspended in 0.15 ml TE (10, 50) [10 mM Tris, 50 mM EDTA], and 2.5 µl 10 mg/ml RNase A was added before 37ºC incubation for 15 minutes.

This digested mixture was phenol/CIA extracted with 0.3 ml 1:1 phenol/CIA, vortexed for 15 seconds and centrifuged at 13,000 rpm (16,060 g) for 1 minute at room temperature in a microcentrifuge (Sorvall Biofuge). The aqueous phase was removed, placed into a new 1.5 ml Eppendorf tube and was CIA/phenol extracted and centrifuged again in the same manner. The aqueous phase was transferred into another Eppendorf tube and sample was CIA extracted and centrifuged using 0.5 ml CIA.

To remove small, unwanted pieces of nucleic acids and to exchange the plasmid-containing buffer, a Sephacryl S400 spun column was prepared by stuffing some siliconized glass wool into a Gilson blue pipette tip that was then fitted into a 5 ml polypropylene tube with the use of the top half of a 1.5 ml Eppendorf tube. The Gilson blue pipette tip was filled to the top with Sephacryl S400 resin equilibrated in TE (10, 0.1) and the whole construction was centrifuged at 2,000 rpm (1,050 g) for 10 minutes at 20ºC (Heraeus rotor #7570 G). Any liquid that eluted into the bottom polypropylene tube was discarded.

After the final CIA extraction of the plasmid-containing liquid, the aqueous phase was loaded into the newly created Sephacryl S400 spun column and centrifuged at 2,000 rpm (1,050 g) for 5 minutes at 20ºC (Heraeus rotor #7570 G). The final eluent contained the purified plasmid in TE (10, 0.1). To determine the final plasmid concentration, UV spectroscopy measured the absorbance at 260 nm (A_{260}) and at 320 nm (A_{320}). Plasmid samples were diluted 1:100 in TE (10, 0.1). Knowing that the adjusted A_{260} of a 50 µg/ml DNA sample is 1.0, the concentration of the stock solution plasmid sample was determined by multiplying the observed (A_{260}-A_{320}) by a factor of 100 and then by 50 µg/ml to give a concentration in units of µg/ml.

2.9 Restriction Mapping and DNA Sequencing

Following plasmid preparation, plasmids were restriction mapped to ensure that the ligation produced the desired plasmid. Restriction mapping utilized one or more

restriction enzymes that would cut the isolated plasmid in multiple regions, creating distinct DNA fragments that would be used to confirm the identity of the plasmid. Ideally, restriction enzymes would cut at least once in the vector region and once in the insert region and would create DNA fragments that were distinct from the parent plasmid.

Analytical restriction enzyme digesting was performed by using 1 µl10x New England Biolabs recommended buffer, 1 µl 1 mg/ml BSA, 0.5 µl 100 mM DTT, 0.5 µl of the each appropriate restriction enzyme, and water to reach a final volume of 9 µl. This digestion mixture was used to digest 1 µl of plasmid DNA (0.2 to 1 µg of DNA) and was incubated at 37ºC for 1 hour. Following digestion, 2 µl of 6xGLB was added and the sample was analyzed on a 1% (w/v) high gelling temperature (HGT) agarose gel at 125 V for 40 minutes. Ethidium bromide staining allowed viewing of the gel on a long wavelength UV transilluminator.

Following successful restriction mapping, plasmids that were synthesized by using PCR were submitted for DNA sequencing at the Nucleic Acid Facility at the Huck Institutes of the Life Sciences at the Pennsylvania State University to verify the DNA sequence.

2.10 Small-Scale Protein Expression

Small-scale expressions were used to test the protein expression products and levels from expression plasmids. The desired expression plasmid was transformed into BL21(DE3)pLysS competent cells and incubated at 37ºC. BL21(DE3)pLysS cells confer

chloramphenicol resistance and all of the expression plasmids that were expressed conferred ampicillin resistance. The liquid growth media of choice for expression was 100 ml 2xTY media + 50 µg/ml ampicillin + 25 µg/ml chloramphenicol in a 500 ml Erlenmeyer flask. After 12-18 hours of growth, 3 colonies from the transformation plate were used to inoculate the liquid media prior to 37ºC incubation.

The optical density of the cell culture was monitored at 600 nm (OD_{600}) until it reached the absorbance range of 0.5 to 0.9. At this point the culture was induced with 100 µl 0.2 M Isopropyl ß-D-thiogalactopyranoside (IPTG). A 500 µl sample of uninduced culture (0 hour) was removed prior to induction. All time point samples were prepared for SDS-PAGE by centrifuging in a 1.5 ml Eppendorf tube at 13,000 rpm (16,060 g) for 1 minute at room temperature in a microcentrifuge (Sorvall Biofuge), aspirating off the supernatant, and resuspending the pellet in 100 µl PGLB [0.5 M Bis-Tris pH 6.8, 20% (v/v) glycerol, 10% (w/v) SDS, 5 M 2-mercaptoethanol, 0.4 mg/ml bromophenol blue]. Following the 0 hour time point sample, 250 µl samples were taken at 1, 2, 3, and 4 hours. At the 3 hours post-induction time point, 50 ml of the cell culture was centrifuged at 4,000 rpm (4,204 g) for 10 minutes at room temperature (Heraeus rotor #7570 G) in a 50 ml Falcon tube. The supernatant was discarded and the pellet was resuspended in 10 ml P300–EDTA buffer [50 mM sodium phosphate pH 7.0, 300 mM sodium chloride, 1 mM benzamidine, 5 mM 2- mercaptoethanol] for future solubility determination and test purification. This sample was flash frozen by submersing the Falcon Tube in liquid nitrogen for a few minutes and was subsequently stored at -20ºC.

To test protein expression and solubility at lower temperatures, small-scale expression was sometimes performed at 28ºC or 18ºC. The beginning steps of 28ºC or 18ºC expression were the same as 37ºC expression until the OD_{600} reached 0.1 to 0.15. At this point the cell cultures were transferred to incubators at either 28ºC or 18ºC to continue growth for several cell cycles. For expression at 28ºC, the remainder of the procedure was the same as 37ºC expression, with time point samples taken and a 50 ml sample taken after 3 hours. For 18ºC expression, following induction and taking a 0 hour time point sample, no further samples were taken until hour 16-22, at which point one time point sample was taken and the 50 ml sample was prepared and flash frozen.

2.11 Solubility Testing and Small-Scale Talon Purification

Following small-scale expression, solubility tests revealed the solubility of various proteins from the expression. The 50 ml sample that was pelleted and resuspended near the end of the small-scale expression was thawed, transferred to a 50 ml glass beaker and sonicated. Sonication was performed twice at 10 seconds with 40% maximum power and 50% duty cycle (Branson Digital Sonifier #S-450D), while storing the beaker on ice in between sonication steps. Following sonication, the whole cell extract (WCE) sample was prepared for SDS-PAGE by removing 25 µl of the sample and adding 25 µl PGLB. All samples for SDS-PAGE were prepared in this manner. To create the supernatant and pellet samples, the remainder of the WCE was transferred to 4x1.5 ml Eppendorf tubes and was centrifuged in a microcentrifuge (Sorvall Biofuge) at 13,000 rpm (16,060 g) for 5 minutes at room temperature. From the supernatant, an SDS-PAGE sample was

prepared, and the remaining supernatant was set aside for small scale Talon purification if applicable. A pellet from one Eppendorf tube was resuspended in P300-EDTA to a volume equal to the original volume that was centrifuged. From this, a pellet sample was prepared for SDS-PAGE.

For small-scale expressions that contained 6x Histidine-tagged dNELFE, Talon Superflow resin (Clontech) was used for small-scale affinity purification. To prepare the Talon resin, 1 ml (a 50% ethanol suspension) was placed in a 15 ml Falcon tube and was washed by adding 10 ml water and centrifuging at 1,800 rpm (700 g) for 2 minutes at room temperature to pellet the resin (Heraeus rotor #7570 G). The supernatant was discarded, and the washing process was repeated with 10 ml of P300-EDTA to equilibrate the Talon resin.

After centrifuging and decanting the equilibrated supernatant, the cell supernatant from the solubility test (~5 ml) was added to the 15 ml Falcon tube containing the equilibrated Talon resin. To allow the resin to bind the proteins, the Falcon tubes were incubated at 4ºC on a rotating platform for 20 minutes. Following incubation, the suspension was centrifuged at 1,800 rpm (700 g) at room temperature for 5 minutes (Heraeus rotor #7570 G). The resulting supernatant (flow through) was decanted into a new 15 ml Falcon tube, and a sample was prepared for SDS-PAGE. The resin was washed twice with 10 ml P300-EDTA (wash A and B) and resuspended in 3 ml P300-EDTA. This resuspended resin was transferred to a Bio-Rad Bio-Spin column. Elution buffer (P300-EDTA+100mM imidazole) was added and 4x0.5 ml fractions were collected in 1.5 ml

Eppendorf tubes. Samples for SDS-PAGE were prepared from each of the 4 fractions. All samples were run on an 18% SDS-PAGE gel and were stained with Coomassie Blue.

2.12 Large-Scale Protein Expression at 18ºC

After small-scale expression and small-scale Talon purification of HISdNELE, 6-liter large-scale expression was performed in hopes of optimizing large-scale purification of the protein. After transforming BL21(DE3)pLysS competent cells with the HISdNELFE expression plasmid and allowing colonies to grow for 12-16 hours, 5 colonies were used to inoculate a starter culture of 100 ml 2xTY media + 50 µg/ml ampicillin + 25 µg/ml chloramphenicol in a 500 ml Erlenmeyer flask. The starter culture was incubated at 37ºC in a shaking incubator until the OD_{600} was between 0.5 and 0.6, at which point 12 large Erlenmeyer flasks each containing 500 ml of 2xTY media + 50 µg/ml ampicillin + 25 µg/ml chloramphenicol were inoculated using 5 ml starter culture.

These large flasks were incubated at 37ºC until the OD_{600} was between 0.1 and 0.15, at which point the incubation temperature was reduced to 18ºC. When the culture reached OD_{600} between 0.5 and 0.6, a 0 hour time point sample was prepared for SDS-PAGE, and the cultures were induced with 500 µl 0.2 M IPTG per flask. A time point was taken at 20 hours post-induction, and the cells were harvested by centrifuging in 500 ml centrifuge bottles at 7,000 rpm (8,381 g) for 7 minutes at 20ºC (Sorvall SLA-3000 rotor). Supernatant was discarded and each cell pellet was resuspended in 150 ml of P300-EDTA before flash freezing and storage at -20ºC. Time points were analyzed on an 18%

SDS-PAGE gel.

2.13 Preparation of Crude Extract

To begin purification of the frozen large-scale protein expression, the frozen crude extract was allowed to thaw in a 30ºC water bath. As soon as the crude extract was fully thawed, it was divided into 5x50 ml beakers and was sonicated twice at 14 seconds with 50% maximum power and 50% duty cycle (Branson Digital Sonifier #S-450D), while storing on ice in between sonication steps.

After a homogenous mixture was reached by sonication, the whole cell extract was transferred to 50 ml polypropylene tubes and was centrifuged at 18,000 rpm (39,121 g) for 20 minutes at 4ºC (SS-34 rotor). The supernatant was transferred to a 250 ml flask and was kept on ice for later use. Samples of whole cell extract, supernatant and the re-suspended pellet were prepared for SDS-PAGE.

2.14 Western Blotting

Western blotting was used to identify HISdNELFE in HPLC elution fractions. Elution fraction samples were analyzed on an 18% SDS-PAGE gel. After electrophoresis, the gel was removed and equilibrated in 30 ml Western transfer buffer [25 mM Tris, 192 mM glycine, pH 8.3] for 5 minutes. Following equilibration, a trans-blot sandwich was constructed, which was arranged from negative to positive side in the following order: a

3M fiber pad, Whatman filter paper, the SDS-PAGE gel, nitrocellulose blotting membrane, Whatman filter paper, and a 3M fiber pad. This assembly was clamped and transferred to the blotting apparatus. An ice block and stir bar were placed in the apparatus, and it was filled with Western transfer buffer with 20% methanol. The protein was blotted onto the nitrocellulose membrane overnight at 30 V constant voltage at 4ºC.

The nitrocellulose membrane was removed and placed in 25 ml of 1xTBS buffer [0.25 M Tris, 1.5 M NaCl, pH 8.0] for 5 minutes, followed by 1 hour in 25 ml preincubation buffer [2% nonfat dry milk, 1xTBS] to block nonspecific binding sites on the nitrocellulose membrane. The membrane was then washed twice for 5 minutes each with 50 ml 1xTTBS [1xTBS, 0.05% Tween 20]. The nitrocellulose membrane was incubated with the primary antibody [a 1:1000 dilution of antibody in TTBS] for 1 hour, and then washed 3 times for 5 minutes each in 25 ml TTBS. The membrane was incubated with the secondary antibody [a 1:1000 dilution of antibody in TTBS] for 1 hour and then washed 3 times for 5 minutes each in 25 ml TTBS. The membrane was incubated in a 1:1 solution of ECL detection solution 1 [0.1 M Tris-Cl, 4.4 mM luminal, 4.3 mM PIP pH 9.35] and ECL detection solution 2 [0.1 M Tris-Cl, 0.012% H_2O_2, pH 9.35] for 5 minutes. The membrane was wrapped in plastic film and exposed onto film.

2.15 Experimental Purification of HISdNELFE from 18ºC Large-Scale Expression

Following a 6-liter 18ºC expression of HISdNELFE, and preparation of the crude extract, ~160 ml supernatant was loaded onto the equilibrated 20 ml Talon column at 4ºC. The

column was equilibrated with P300-EDTA buffer [50 mM sodium phosphate pH 7.0, 300 mM sodium chloride, 1 mM benzamidine, 5 mM 2-mercaptoethanol]. After the supernatant was loaded onto the column, the column was eluted with P300-EDTA+100 mM imidazole to elute bound proteins from the resin. Key elution fractions were analyzed by 18% SDS-PAGE to identify fractions containing the 33 kDa HISdNELFE. Fractions containing HISdNELFE that eluted from the column were pooled into 12-14 kDa dialysis tubing and dialyzed overnight against H50 buffer [10 mM HEPES pH 7.5, 50 mM NaCl, 10 mM 2-mercaptoethanol] at 4ºC.

The following day, analytical Source S1 cation exchange chromatography was performed. The Source S1 column was equilibrated in H50 buffer [10 mM HEPES pH 7.5, 50 mM NaCl, 10 mM 2-mercaptoethanol]. Before loading, the sample from the Talon pool was centrifuged in a microcentrifuge at 13,000 rpm (16,060 g) for 5 minutes at room temperature (Sorvall Biofuge). The supernatant was loaded onto the column and was eluted with an increasing salt gradient from 0% to 100% H500 [10 mM HEPES pH 7.5, 500 mM NaCl, 10 mM 2-mercaptoethanol] buffer. Elution fractions were analyzed by SDS-PAGE. The Talon pool was dialyzed against H100 [10 mM HEPES pH 7.5, 100 mM NaCl, 10 mM 2-mercaptoethanol] overnight.

The following day, analytical Source S1 cation exchange chromatography was performed as explained above with H100 and H500 buffers, and the column was eluted with an increasing salt gradient from 0% to 73%. The sample was 2 ml of Talon pool dialyzed against H100. The limits of HISdNELFE solubility were analyzed by an analytical

ammonium sulfate precipitation of the Talon pool.

In preparation for analytical Source ISO1 hydrophobic interaction chromatography, the Source ISO1 column was equilibrated in TA1000 buffer [20 mM Tris-Cl pH 8.0, 1000 mM ammonium sulfate, 10 mM 2-mercaptoethanol]. Before loading, ammonium sulfate was added to the sample to bring it to a final concentration of 1.0 M ammonium sulfate. The sample was centrifuged in a microcentrifuge at 13,000 rpm (16,060 g) for 5 minutes, and the supernatant was loaded onto the Source ISO 1 column and eluted with a 100% to 0% decreasing salt gradient. The low salt buffer was TA0 [20 mM Tris-Cl pH 8.0, 10 mM 2-mercaptoethanol]. Elution fractions were analyzed by SDS-PAGE.

Stronger hydrophobic interactions were tested with analytical Source PHE1 hydrophobic interaction chromatography. The buffers, sample and salt gradient were the same as the previous Source ISO trial. Elution fractions were analyzed with SDS-PAGE. The Talon pool was centrifuged at 12,000 rpm (17,387 g) for 10 minutes at 4ºC (Sorvall SS-34) and the pellet was mixed with 20% glycerol. Both the pellet and the supernatant were flash frozen and stored at -20ºC.

Weeks later, the frozen supernatant was thawed and 1.5 ml was dialyzed against pH 7.5 buffer [10 mM HEPES pH 7.5, 100 mM NaCl, 10 mM 2-mercaptoethanol], pH 6.5 buffer [10 mM HEPES pH 6.5, 100mM NaCl, 10 mM 2-mercaptoethanol] and pH 5.5 buffer [10 mM HEPES pH 5.5, 100 mM NaCl 10 mM 2-mercaptoethanol]. An analytical test of solubility at different pH was performed by SDS-PAGE. In preparation for analytical

Source S1 cation exchange chromatography, 10 ml of the supernatant was dialyzed against pH 5.5 buffer [10 mM HEPES pH 5.5, 100 mM NaCl, 10 mM 2-mercaptoethanol] overnight.

The following day, analytical Source S1 cation exchange chromatography was performed. The Source S1 column was equilibrated in H100 buffer [10 mM HEPES pH 5.5, 100 mM NaCl, 10 mM 2-mercaptoethanol]. Before loading, the sample from the Talon pool was centrifuged in a microcentrifuge at 13,000 rpm (16,060 g) for 5 minutes (Sorvall Biofuge). The supernatant was loaded onto the column and eluted with an increasing salt gradient from 0% to 100% H1000 buffer [10 mM HEPES pH 5.5, 1000 mM NaCl, 10 mM 2-mercaptoethanol] buffer. Elution fractions were analyzed by SDS-PAGE and Western blotting.

3. Results

3.1 Molecular Cloning of Drosophila NELF Subunits

The initial cloning strategy was to subclone all four NELF subunit genes (A, B, D and E) into a polycistronic expression plasmid as illustrated in Figure 3.1. The original NELF genes were obtained from Dr. David Gilmour of the Center for Gene Regulation at The Pennsylvania State University. These template plasmids were amplified through polymerase chain reaction (PCR) to obtain usable quantities of DNA. The oligonucleotides that were designed for these PCR amplifications added 5' and 3' restriction sites to facilitate the cloning steps necessary to create a final polycistronic expression plasmid containing all four NELF genes. Each of the subunit gene inserts was first cloned into either pWM528 or pWM529 high-copy cloning plasmids (Mandecki *et al*., 1990). To verify successful cloning, each plasmid was restriction mapped, and was sequenced through the entire coding region. A summary of the initial constructs, including template DNA and initial PCR amplification, can be seen in Table 3.1.

DNA Template	5’ Restriction Sites Added	3’ Restriction Sites Added	Amplified Fragment	Plasmid Product
GH10333 cDNA	EcoRI-NdeI	XbaI-NgoMIV	dNELF-B (1-596)	pWM529- dNELFB
pA5CΔP-Flag-RD	EcoRI-BamHI	XbaI-BsrGI	dNELF-E (2-294)	pWM529- dNELFE
LD42626 cDNA	EcoRI-NdeI	XbaI-BsrGI	dNELF-D (2-578)	pWM529- dNELFD
pOT2-SD09448	XbaI-NdeI	HindIII-BsrGI	dNELF-A (1-1248)	pWM529- dNELFA

Table 3.1: A summary of initial PCR amplifications.

Creating pWM528-dNELFA proved to be difficult and did not proceed as illustrated in Figure 3.1. Initial PCR amplification of the template DNA was unsuccessful, possibly because the region to be amplified was over 3,780 base pairs long. By PCR amplifying dNELFA as 5' and 3' halves using internal oligonucleotide primers, two halves of the dNELFA gene were successfully amplified and cloned into pWM529 vectors. An internal dNELFA restriction site and a different restriction site in the pWM529 vector then allowed cloning of both dNELFA halves into the pWM529 vector, creating pWM529-dNELFA. To allow subsequent subcloning of dNELFA, PCR mutagenesis was used to remove an internal NdeI site, creating pWM529-dNELFAx1.

Sequencing of pWM529-dNELFE revealed two single-base, non-conservative mutations and 11 silent mutations in the codon wobble positions compared to the expected sequence of dNELFE. Since the silent changes did not affect the amino acid sequence, this was deemed an acceptable allelic variation. The two non-conservative mutations were reverted back to the expected sequence.

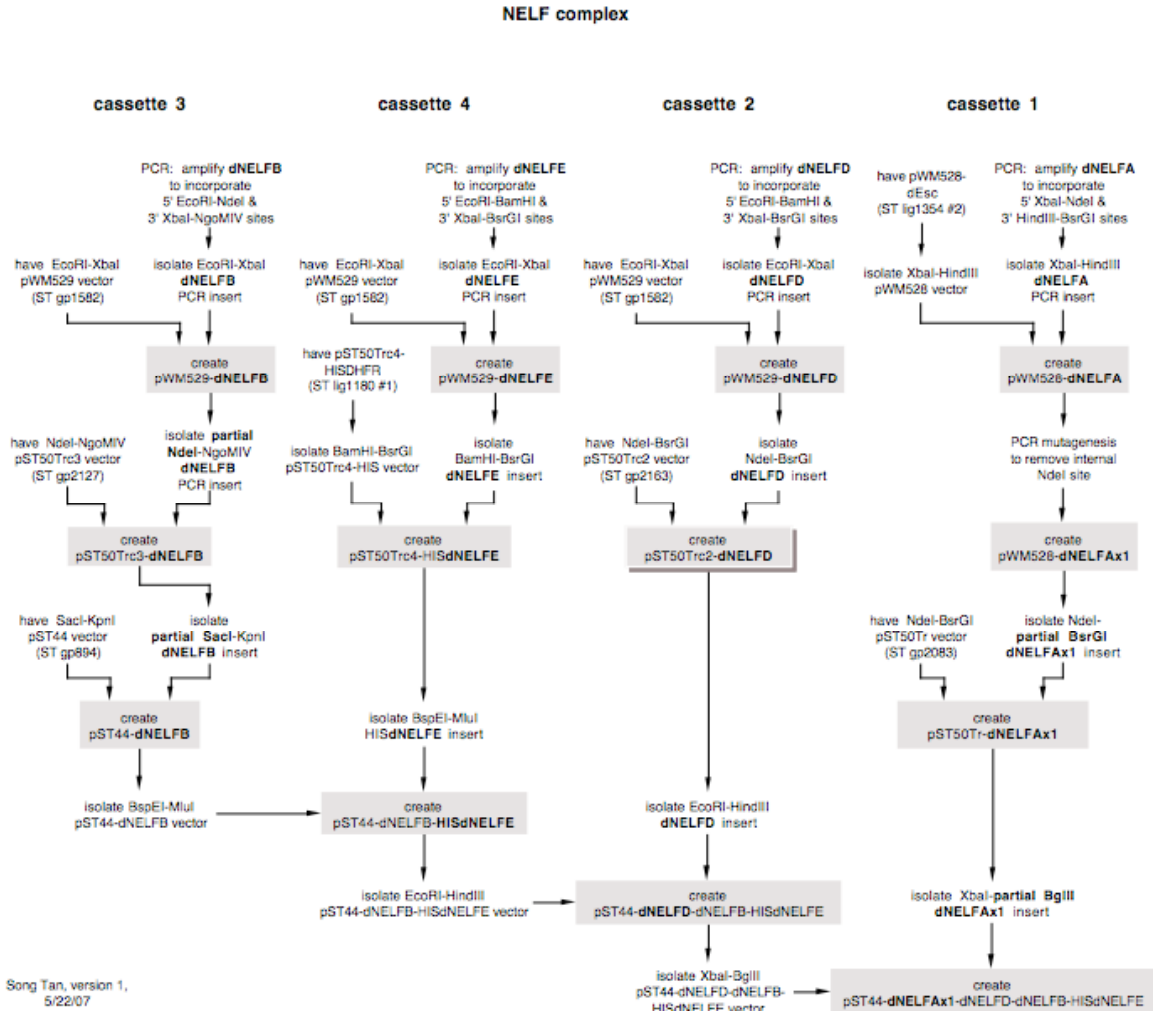

Figure 3.1: Cloning flow chart of the dNELF complex. Figure designed by Dr. Song Tan.

Each NELF subunit was then cloned into a transfer cassette that was capable of inducible protein expression. Recombinant human and *Drosophila* NELF complexes containing all four subunits have previously been purified by using N-terminal affinity tags (Narita *et al*., 2003), so in this cloning step, a 6x Histidine tag was added to the N terminus of dNELFE to create HISdNELFE and to allow future affinity purification with Talon resin. Each pST50Tr expression plasmid construct was verified by restriction mapping.

3.2 Molecular Cloning to Create the Polycistronic Expression Plasmid

To create polycistronic expression plasmids, the transfer cassettes were cloned into the pST44 polycistronic vector as illustrated in Figure 3.1. Each construct was verified with restriction mapping upon completion. All constructs were successfully created with the exception of the final construct, pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE.

Isolation of the dNELFAx1 (XbaI-BglII) insert from pST50Tr-dNELFAx1 required a complete XbaI digestion and a partial BglII digestion (Figure 3.2A). The pST50Tr-dNELFAx1 plasmid contained 6,546 base pairs (bp) with an XbaI restriction site located at 101 bp, and BglII restriction sites located at 468 and 3,902 bp. An XbaI cut at 101 bp and a BglII cut at 3,902 bp would yield the dNELFAx1 fragment of 3,801 bp. Assuming a complete XbaI digestion, BglII digestion at only the 468 bp site would yield fragments of 6,179 + 367 bp, BglII digestion at only the 3,902 bp site would yield fragments of 3,801 + 2,745, and BglII digestion at both sites would yield fragments of 3,434 + 2,745 + 367 bp.

As seen in Figure 3.2B, complete XbaI and partial BglII digestion of 5 and 8 minutes produced only expected DNA fragments. Distinct fragments of 3,801 bp were excised and gel purified to create the dNELFAx1 insert for use in attempts to create pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE.

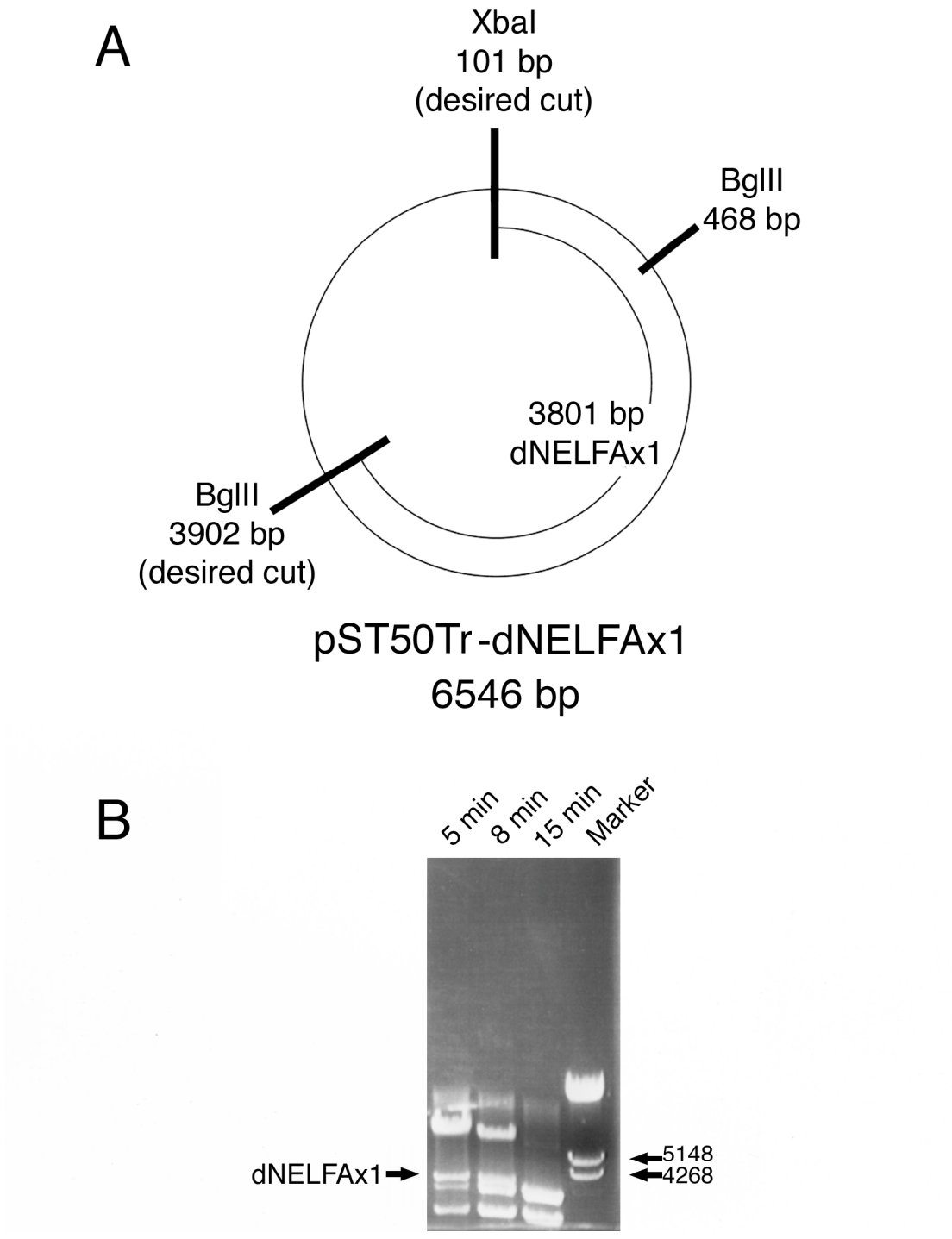

Figure 3.2: Restriction digest to create dNELFAx1. (A) An illustration of pST50Tr-dNELFAx1 showing XbaI and BglII restriction sites for the restriction digest to create the dNELFAx1 (XbaI-BglII) insert. (B) Agarose gel electrophoresis of XbaI-partial BglII digestion of pST50Tr-dNELFAx1 to produce the dNELFAx1 (XbaI-BglII) insert.

Several attempts to clone the dNELFAx1 insert into the pST44-dNELFD-dNELFB-HISdNELFE vector were unsuccessful. These results were characterized by poor (vector + insert)/(vector only) transformation colony ratios, negative PCR screening results when using even dNELFAx1-specific internal oligonucleotide primers that should produce expected PCR products of <1,000 bp, and failure to isolate plasmid that produced the expected restriction mapping fragments. A high transformation colony ratio would have suggested that the desired insert was ligated into the vector, and that the vector was not simply religating without the insert.

It was speculated that the large size of the desired plasmid (11,126 bp) could be detrimental to its existence within the host cell. Following disappointing results using TG1 competent cells as a host strain, HB101 competent cells were used because they were deficient in RecA, an *E. coli* DNA repair and maintenance protein. The RecA protein has a tendency to remove large pieces of exogenous DNA, and its absence could have potentially aided in ensuring the stability of extrachromosomal DNA. Attempts to subclone dNELFAx1 into other vectors such as pST44-dNELFB (XbaI-BglII) and pST53 (XbaI-BglII) were unsuccessful. During transformations in these failed attempts, concurrent positive control transformations performed did not indicate that either the dNELFAx1 insert or the pST44-dNELFD-dNELFB-HISdNELFE vector was directly related to poor transformant colony counts.

3.3 Small-Scale Expressions, Solubility Tests and Small-Scale Purifications of dNELF Subunits and Subcomplexes

After dNELF genes were cloned into expression plasmids, the encoded proteins were expressed at various temperatures to determine the optimal expression temperatures to maximize expression levels and solubility. Expression of recombinant proteins in *E. coli* at lower temperatures can change protein expression levels and increase protein solubility (Tolia *et al.,* 2005; Niiranen *et al.,* 2007). Expression plasmids were tested for induction at 37ºC, 28ºC and 18ºC. Some expression plasmids contained HISdNELFE, which allowed small-scale affinity purification with Talon resin via the hexahistidine affinity tag with the goal of determining optimal purification conditions. Table 3.2 summarizes the expression plasmids that were successfully created and expressed.

Protein or Complex	Expression Plasmid	MW (kDa)	pI
dNELF B	pST44-dNELFB	68.1	6.66
dNELF E	pST50Trc4-**HIS**dNELFE	33.1	10.28
dNELF D	pST50Trc2-dNELFD	65.8	4.83
dNELF A	pST50Tr-dNELFAx1	134.5	10.72
dNELF BE	pST44-dNELFB-**HIS**dNELFE	68.1, 33.1	6.66, 10.28
dNELF BDE	pST44-dNELFB-dNELFD-**HIS**dNELFE	68.1, 65.8, 33.1	6.66, 4.83 10.28

Table 3.2: Created dNELF plasmids. Histidine affinity tags are bolded.

Small-scale expression of dNELFB revealed optimal expression at 28ºC: medium expression levels of a protein with the expected molecular weight (68.1 kDa) (Figure 3.3). Expression at 37ºC demonstrated slightly lower levels of expression and 18ºC expression resulted in low expression levels (data not shown). Solubility tests revealed

that the majority of dNELFB protein was insoluble at all temperatures. The solubility test from 28ºC expression can be seen in Figure 3.4. Many bands were visible in lane 3 (the pellet) of Figure 3.4, possibly indicating that the rapid expression of dNELFB might not have allowed all of the proteins to be folded properly, forming aggregates of misfolded, insoluble proteins that are primarily dNELFB, but that also contain *E. coli* proteins.

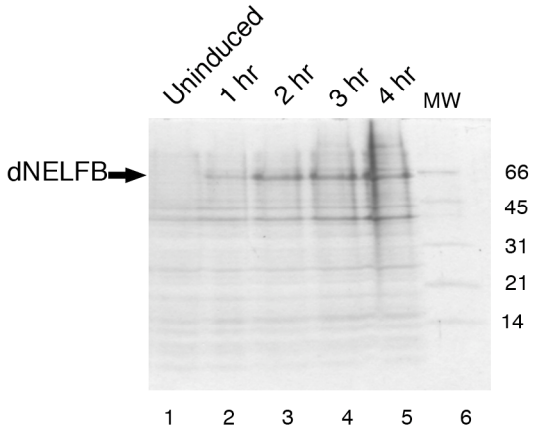

Figure 3.3: Small-scale expression of dNELFB. The figure above shows expression in BL21(DE3)pLysS cells at 28ºC.

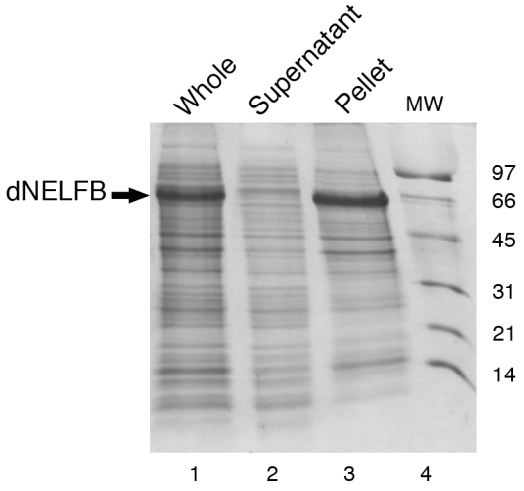

Figure 3.4: Solubility test of dNELFB. The figure above shows the solubility of the small-scale expression of dNELFB in BL21(DE3)pLysS cells at 28ºC.

Small-scale expression of HISdNELFE revealed optimal expression at 18ºC: high expression levels of a protein with the expected molecular weight (33 kDa) (Figure 3.5). Expression at 37ºC and at 28ºC demonstrated medium expression levels (data not shown). Small-scale Talon purification of the 18ºC expression confirmed that roughly two-thirds of HISdNELFE was soluble. Some HISdNELFE was present in the flow through, but a large amount of HISdNELFE was present in all elution fractions (Figure 3.5). Some low-molecular weight contaminants were present in the elution fractions. Small-scale Talon purification of 37ºC expression revealed some purification of HISdNELFE in elution fractions and 50% solubility of HISdNELFE. Small-scale Talon purification of 28ºC expressions also resulted in some purification visible in elution fractions, but >50% solubility of the HISdNELFE (data not shown). Based on the results, I chose to perform large-scale expression of HISdNELFE at 18ºC.

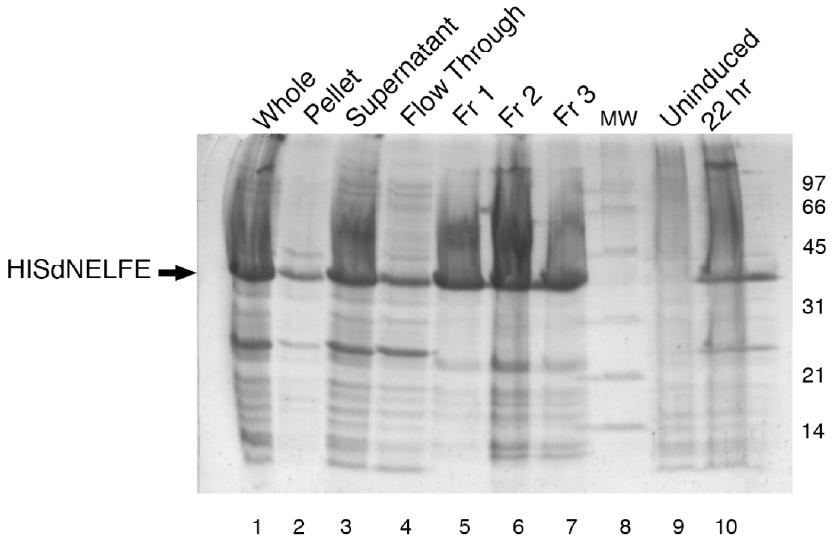

Figure 3.5: Small-scale expression and small-scale purification of HISdNELFE. The figure above shows expression in BL21(DE3)pLysS cells at 18ºC. The initial supernatant volume was ~5 ml, flow through volume was ~ 5 ml, and each eluent fraction volume was ~0.5 ml. A sample of each of these was combined 1:1 with PGLB and 5 µl was loaded into each lane.

Small-scale expression of dNELFD revealed low levels of expression at both 37ºC and 18ºC. A protein of expected molecular weight (65.8 kDa) was visible, but expression of this protein was not significantly stronger than the background proteins as shown in the 18ºC expression (Figure 3.6). A solubility test of this putative dNELFD protein revealed that it was roughly 50% soluble (Figure 3.6). Data for the 37ºC expression and solubility test are not shown. A high proportion of background *E. coli* proteins were seen to be insoluble, possibly caused by aggregation of misfolded proteins under the induced expression.

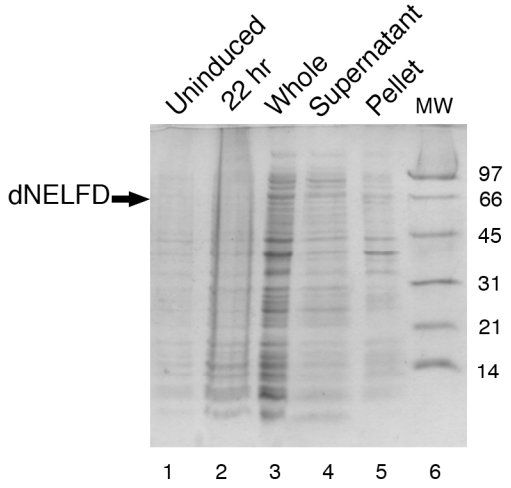

Figure 3.6: Small-scale expression and solubility test of dNELFD. The figure above shows expression in BL21(DE3)pLysS cells at 18ºC.

Small-scale expression of dNELFAx1 did not definitively display a protein with expected molecular weight (134.5 kDa) at 37ºC, 28ºC, or 18ºC. The solubility test of dNELFAx1 (Figure 3.8 B lanes 1 and 2) showed two protein bands of ~135kDa, while the small-scale expression and solubility test of dNELFD (Figure 3.6) only appeared to show a single protein band of ~135 kDa. However, these results were not strong enough to confirm the expression of dNELFAx1. Solubility tests revealed that the protein that migrated at

~135kDa was ~50% soluble (Figure 3.8). Data for 28ºC and 18ºC small-scale expressions and solubility tests are not shown.

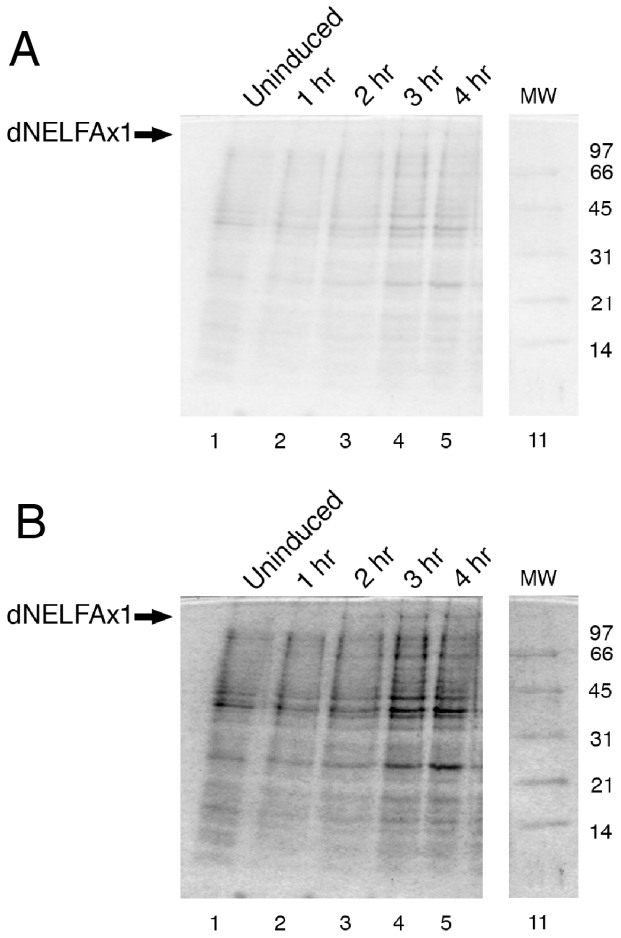

Figure 3.7: Small-scale expression of dNELFAx1. The figure above shows the expression in BL21(DE3)pLysS cells at 37ºC. (A) Is the unaltered image of the SDS-PAGE gel. (B) The contrast has been adjusted in Adobe Photoshop (Adobe Systems Inc.) to reveal faint bands.

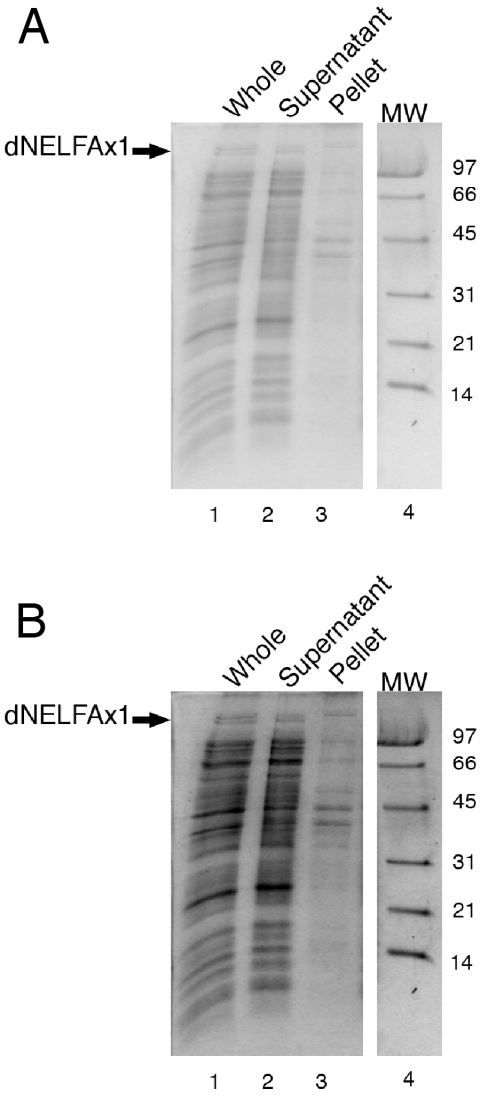

Figure 3.8: Solubility test from the small-scale expression of dNELFAx1. The figure above shows expression in BL21(DE3)pLysS cells at 37ºC. (A) Is the unaltered image of the SDS-PAGE gel. (B) The contrast has been adjusted in Adobe Photoshop (Adobe Systems Inc.) to reveal faint bands.

Small-scale expression of dNELFB-HISdNELFE found optimal expression at 37ºC: high expression levels of a protein with the HISdNELFE expected molecular weight (33 kDa), and medium expression levels of a protein with the dNELFB expected molecular weight (68.1 kDa) (Figure 3.9). Expression at 28ºC and 18ºC revealed high levels of HISdNELFE expression, but decreasing levels of dNELFB expression at lower

temperatures of expression (data not shown).

At all expression temperatures, HISdNELFE could be partially purified by Talon affinity chromatography, but slightly more of the protein was soluble at lower temperatures of expression. dNELFB displayed <50% solubility at 37ºC, but due to a significant drop-off in expression levels, it was difficult to conclude the solubility at lower expression temperatures. There did not appear to be copurification of dNELFB with HISdNELFE. Small-scale purification at 37ºC for the dNELFB-HISdNELFE binary complex is shown in Figure 3.10.

The optimal temperature for small-scale purification of HISdNELFE was determined to be 37ºC because while the recombinant protein remained somewhat constant in solubility and Talon purification at all temperatures of expression, dNELFB revealed highest expression and solubility at 37ºC.

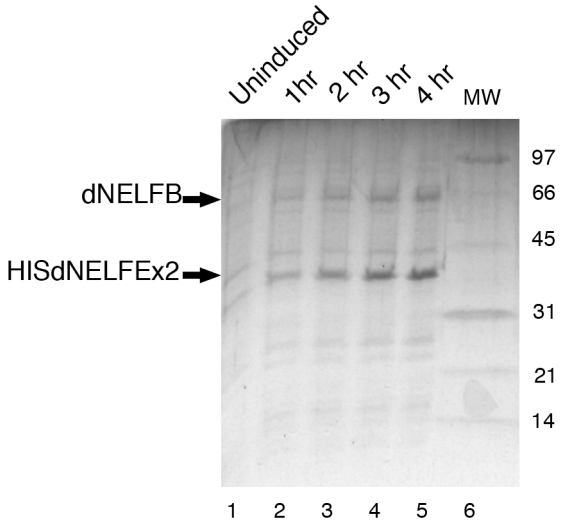

Figure 3.9: Small-scale expression of dNELFB-HISdNELFE. The figure above shows the expression in BL21(DE3)pLysS cells at 37ºC.

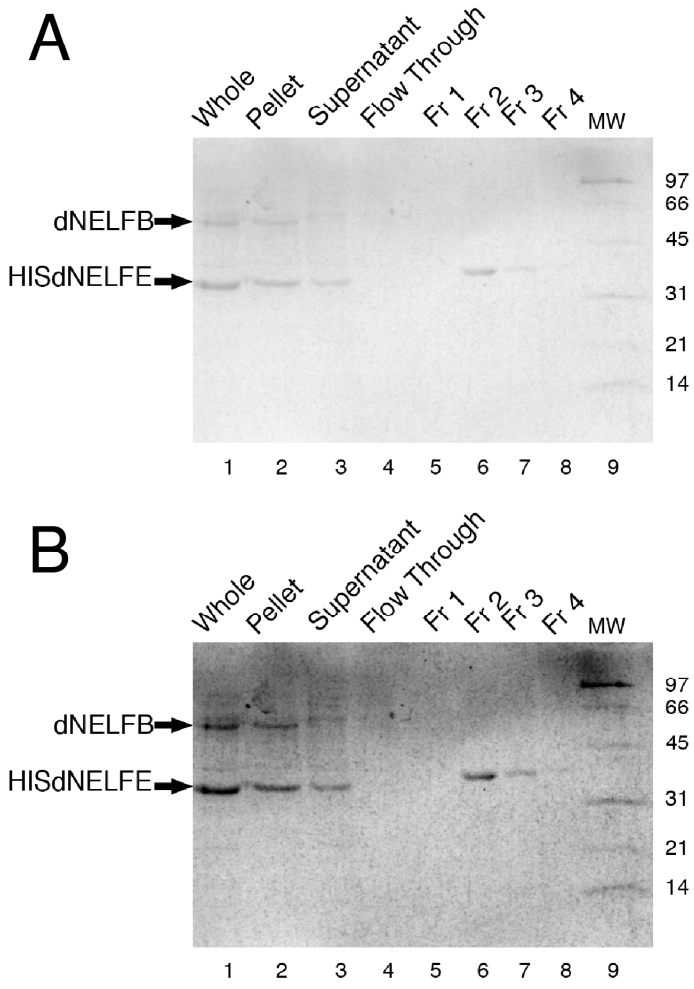

Figure 3.10: Small-scale purification of dNELFB-HISdNELFE. The figure above shows the small-scale purification from a small-scale expression in BL21(DE3)pLysS cells at 37ºC. (A) Is the unaltered image of the SDS-PAGE gel. (B) The contrast has been adjusted in Adobe Photoshop (Adobe Systems Inc.) to reveal faint bands.

Small-scale expressions of dNELFD-dNELFB-HISdNELFE at 37ºC, 28ºC, and 18ºC all showed medium levels of expression for HISdNELFE (33 kDa) and low levels of expression for dNELFB (68.1 kDa) and dNELFD (65.8 kDa). The expression at 18ºC can be seen in Figure 3.11.

The results of small-scale purification of this complex at the three different expression temperatures were similar to each other. HISdNELFE could clearly be seen in the whole, pellet, supernatant and flow through fractions. It revealed roughly 50% solubility, and the majority of the soluble protein bound the Talon resin before elution. dNELFB and dNELFD both appeared in the whole, pellet, supernatant and flow through fractions, and appeared to be partly soluble. The elution fractions showed purification of HISdNELFE, a small amount of what appeared to be dNELFB, and some low molecular weight contaminants. dNELFB, though not nearly at a 1:1 stoichiometric ratio with HISdNELFE, showed better purification than the background host cell proteins. Some low molecular weight contaminants also appeared. Small-scale purification of the 18ºC expression can be seen in Figure 3.11.

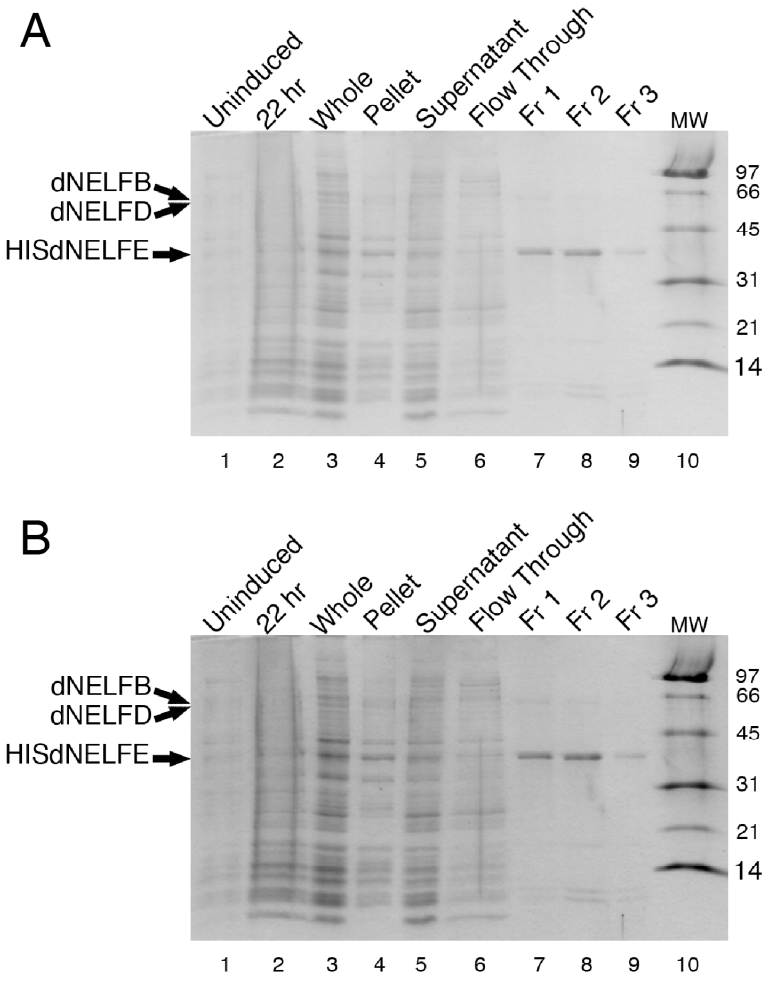

Figure 3.11: Small-scale expression and small-scale purification of dNELFD-dNELFB-dNELFE. The figure above shows expression in BL21(DE3)pLysS cells at 18ºC. (A) Is the unaltered image of the SDS-PAGE gel. (B) The contrast has been adjusted in Adobe Photoshop (Adobe Systems Inc.) to reveal faint bands.

3.4 Large-Scale Purification of HISdNELFE

Following small-scale expression and small-scale purification of HISdNELFE, optimal expression conditions proved to be 18ºC expression for 16+ hours to achieve the highest protein expression, solubility and Talon purification. HISdNELFE was expressed on a 6-liter scale and the cells were harvested after 4 hours of induction. After the cells were

lysed by thawing and sonication, the crude extract was centrifuged and prepared for purification on a Talon affinity column. After elution with 100 mM imidazole, selected fractions were analyzed on SDS-PAGE for the presence of HISdNELFE (Figure 3.12).

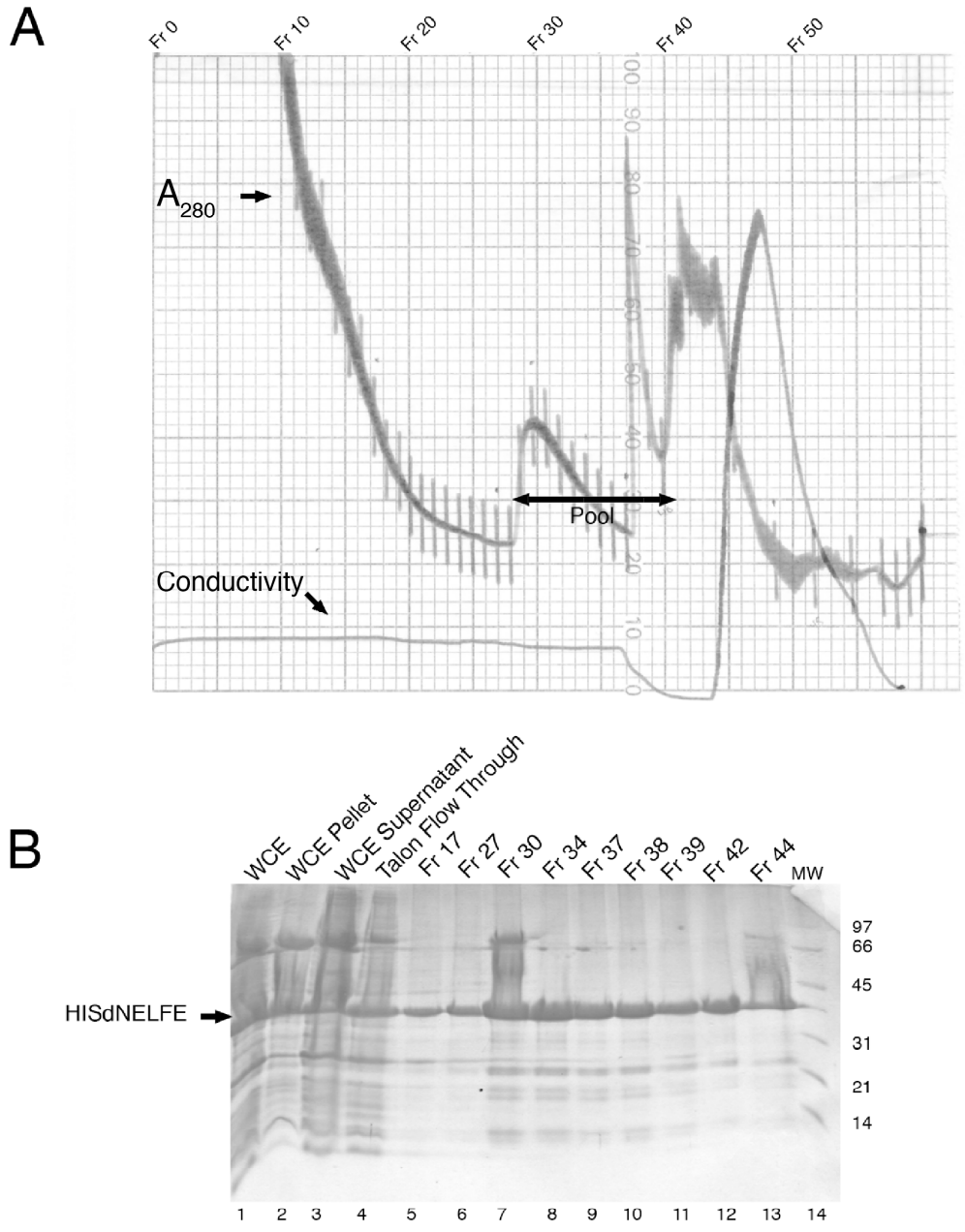

Figure 3.12: **Large-scale Talon column affinity purification of HISdNELFE. (A) Elution profile of HISdNELFE from the Talon column as represented by A_{280}. (B) SDS-PAGE analysis of selected elution fractions. Fractions 28-41 were pooled.**

The Talon chromatogram displayed multiple elution peaks. SDS-PAGE analysis showed the presence of high amounts of HISdNELFE in all fractions, possibly suggesting high levels of expression and Talon column saturation. Fractions 28-41 were then pooled and dialyzed overnight against H50 buffer [10 mM HEPES pH 7.5, 50 mM NaCl, 10 mM 2-mercaptoethanol].

The isoelectric point of HISdNELFE was 10.28, so it was appropriate to attempt purifying a sample of the Talon pool over an analytical Source S1 cation exchange column as shown in Figure 3.13.

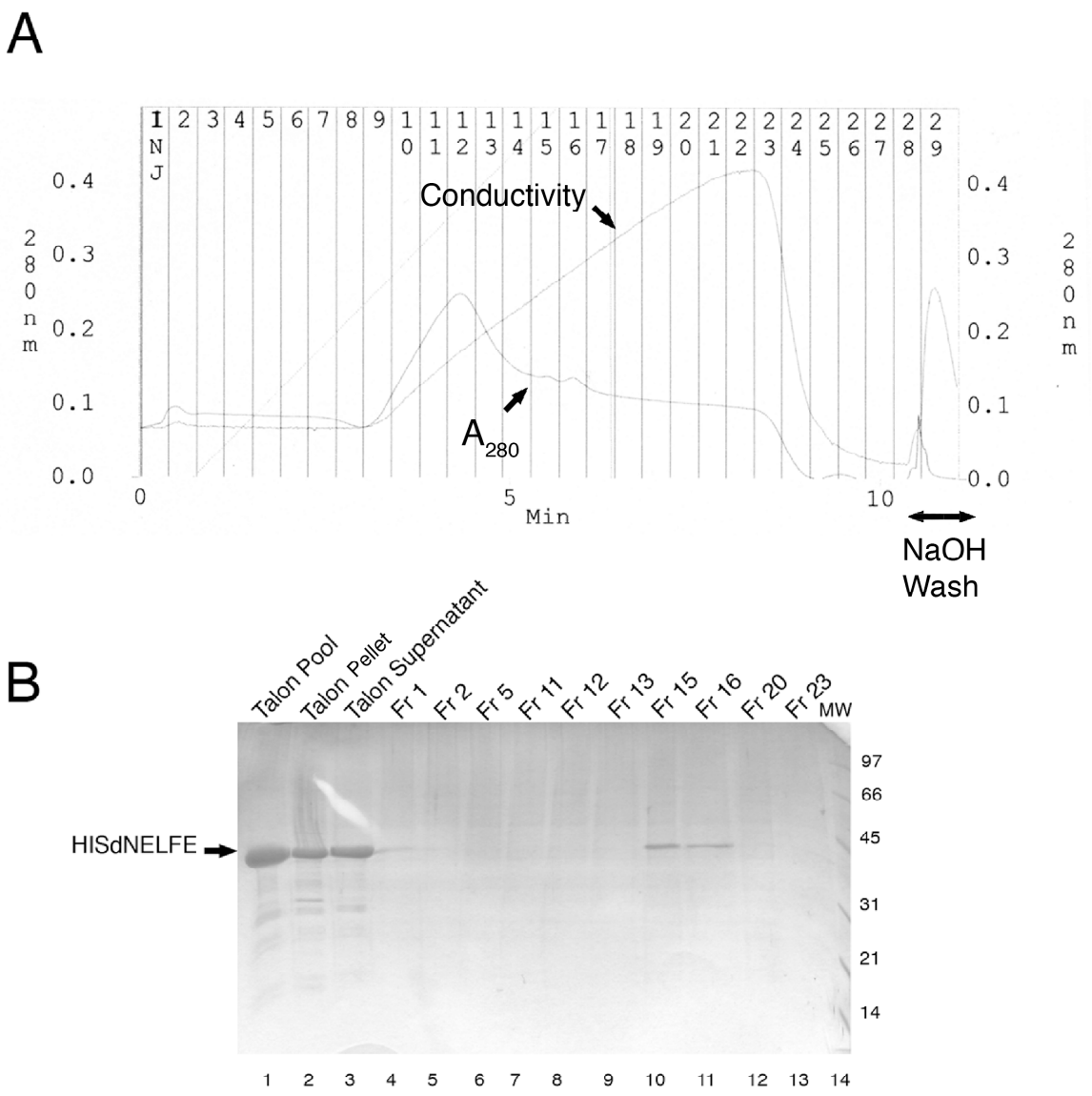

Figure 3.13: **Analytical Source S1 cation exchange chromatography purification of HISdNELFE. (A) Elution profile of HISdNELFE from the Source S1 column as represented by A_{280}. (B) SDS-PAGE analysis of selected elution fractions.**

Analysis with SDS-PAGE demonstrated that roughly 50% of the dialyzed HISdNELFE Talon pool was soluble, but that the majority of HISdNELFE that bound to the column did not elute during the salt gradient. Two small peaks, corresponding to elution fractions 15 and 16 revealed the presence of HISdNELFE.

The Talon pool indicated signs of precipitation in the dialysis tubing, so it was dialyzed against a higher salt buffer, H100 buffer [10 mM HEPES pH 7.5, 100 mM NaCl, 10 mM 2-mercaptoethanol], in an attempt to increase the solubility of HISdNELFE. Following dialysis buffer change and equilibration, SDS-PAGE analysis revealed that ~60% of HISdNELFE was soluble (results not shown).

The previous analytical Source S1 cation exchange chromatography purification step revealed only a small amount of HISdNELFE eluting from the column during the salt gradient in two small peaks. To improve the purification, a second analytical purification step was performed using the same column with a greater sample volume, a more gradual salt gradient, and H100 low salt buffer instead of H50 low salt buffer (Figure 3.14).

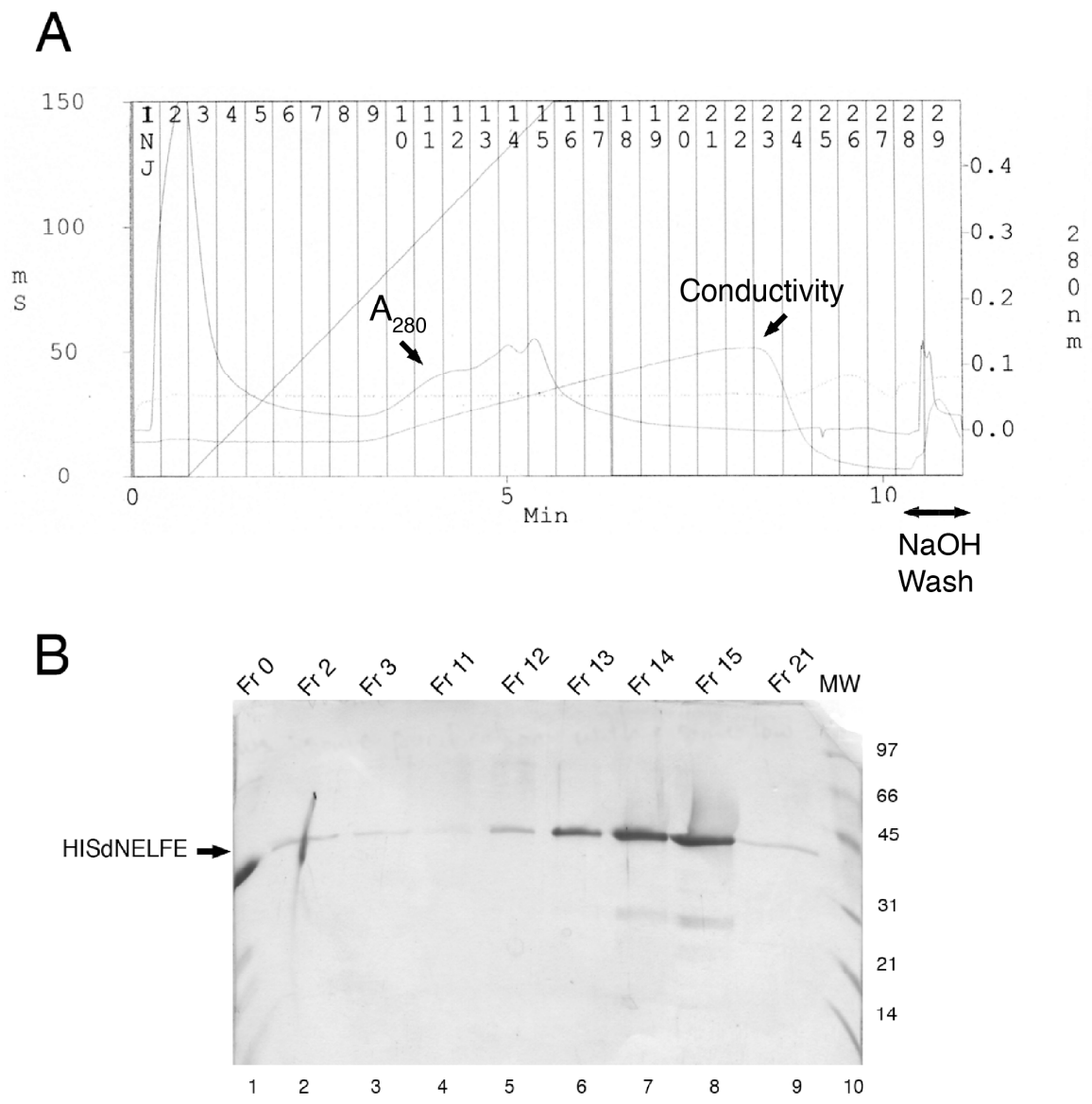

Figure 3.14: **Analytical Source S1 cation exchange chromatography purification of HISdNELFE. (A) Elution profile of HISdNELFE from the Source S1 column as represented by A_{280}. (B) SDS-PAGE analysis of selected elution fractions.**

Analysis by SDS-PAGE revealed slightly greater resolution of the two small HISdNELFE elution peaks (fractions 14 and 15). These fractions appeared to show most of the HISdNELFE input sample (fraction 0), but they indicated low molecular weight contaminants also present in fraction 0. Most of the sample appeared to be accounted for by elution samples during the salt gradient.

The next purification step attempted was analytical Source ISO1 hydrophobic interaction chromatography. To prepare for this, ammonium sulfate precipitation of the Talon pool was performed on an analytical scale to characterize the solubility of HISdNELFE in ammonium sulfate. HISdNELFE was found to precipitate at 1.2 M ammonium sulfate, so the high salt buffer for the hydrophobic interaction chromatography was chosen to be 1.0 M in ammonium sulfate. The Talon pool sample was brought to 1.0 M ammonium sulfate. Purification on Source ISO1 hydrophobic interaction column is displayed in Figure 3.15 A and C.

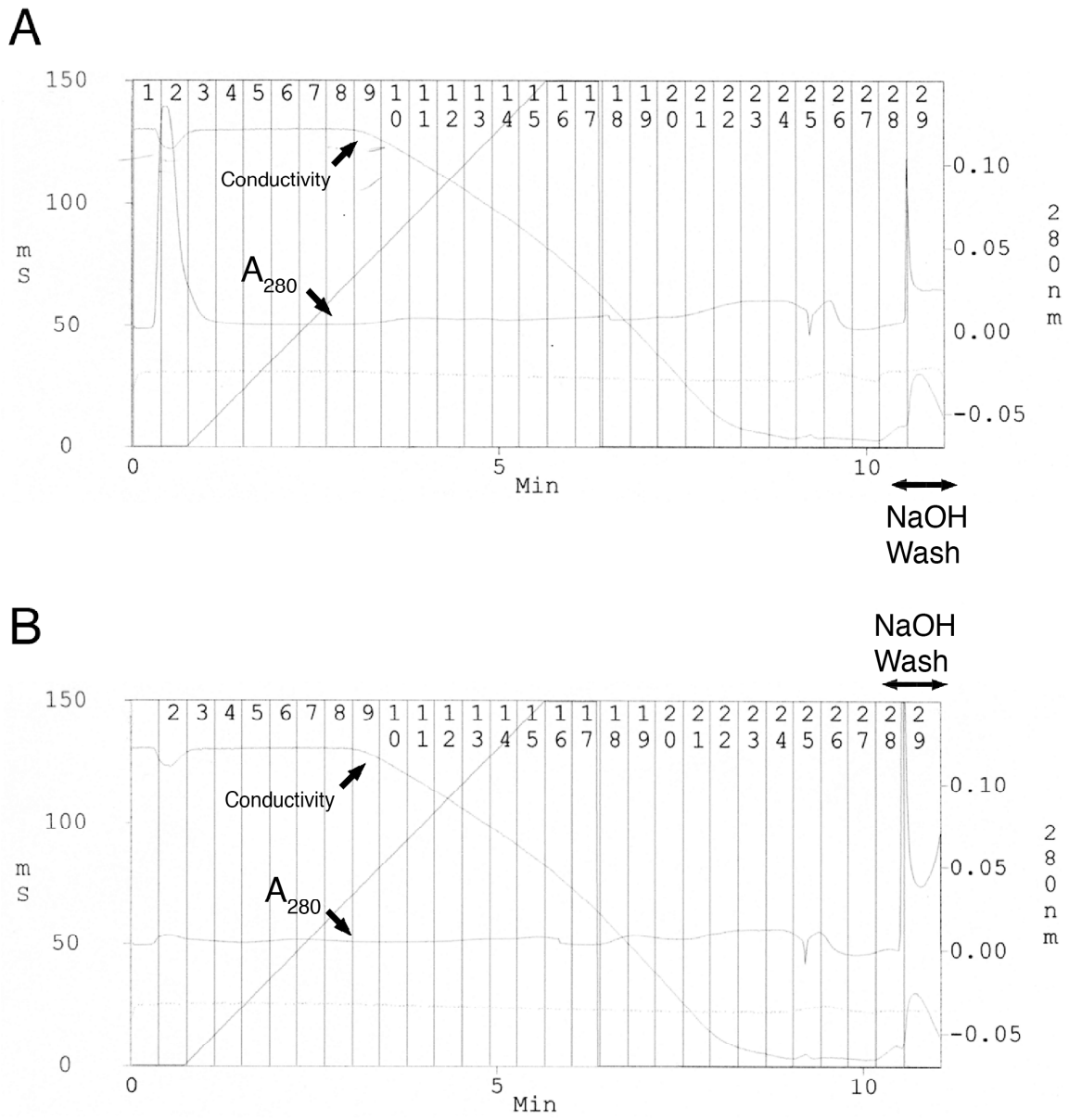

Figure 3.15 (A, B): **Analytical Source ISO1 and Source PHE1 hydrophobic interaction chromatography purification of HISdNELFE. (A) Elution profile of HISdNELFE from the Source ISO1 column as represented by A_{280}. (B) Elution profile of HISdNELFE from the Source PHE1 column as represented by A_{280}.**

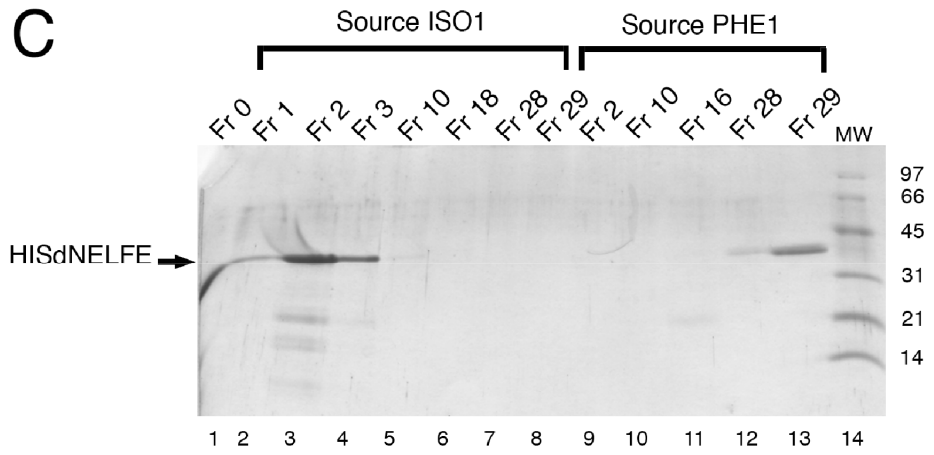

Figure 3.15 (C): **Analytical Source ISO1 and Source PHE1 hydrophobic interaction chromatography purification of HISdNELFE. (C) SDS-PAGE analysis of selected elution fractions from Source ISO1 and Source PHE1.**

The Source ISO chromatogram revealed a large protein peak in the flow through that did not bind to the column, and almost no protein in the decreasing salt gradient. Some protein was visible in the NaOH wash. Analysis with SDS-PAGE revealed that the large peak in the flow through (fractions 1 and 2) was mostly HISdNELFE with some low molecular weight contaminants. Thus, HISdNELFE did not appear to bind to the Source ISO1 column.

Since the hydrophobic interactions with the Source ISO resin were apparently not strong enough to bind HISdNELFE, analytical Source PHE1 hydrophobic interaction chromatography purification was attempted (the phenyl-based PHE resin provides a more hydrophobic surface for interactions). The same low and high salt buffers, ammonium sulfate sample concentration, and decreasing salt gradient were used as in the previous Source ISO1 trial purification. The chromatography results are shown in Figure 3.15.

Unlike the Source ISO purification trial, the chromatogram for analytical Source PHE1 revealed no protein in the flow through fractions (fraction 2). However, it also showed very little protein elution in the decreasing salt gradient. The NaOH wash contained a large protein peak. SDS-PAGE revealed no HISdNELFE in the flow through or salt gradient fractions, but that HISdNELFE eluted in the NaOH wash (fractions 28 and 29). In the decreasing salt gradient (fraction 16), there was a small protein peak on the chromatogram that SDS-PAGE revealed to be some of the low molecular weight contaminants seen in the input fraction.

In hopes that a change in pH may have allowed HISdNELFE to bind to the resin and elute during the salt gradient, an analytical test of solubility at different pH was performed and analyzed with SDS-PAGE to reveal that almost all HISdNELFE was soluble at pH 7.5, pH 6.5 and pH 5.5 (results not shown). Utilizing this information, analytical Source S1 cation exchange chromatography using low and high salt buffers of pH 5.5 was performed on a Talon pool sample dialyzed against pH 5.5 (Figure 3.16 A). As in previous Source S1 purification trials, HISdNELFE was visible in the flow through, in two small peaks in the increasing salt gradient and in the NaOH wash. A Western blot using anti-Histidine antibodies was performed on the elution fractions (Figure 3.16 C) that revealed that the low molecular weight contaminants eluting with HISdNELFE were truncations of the full-length HISdNELFE.

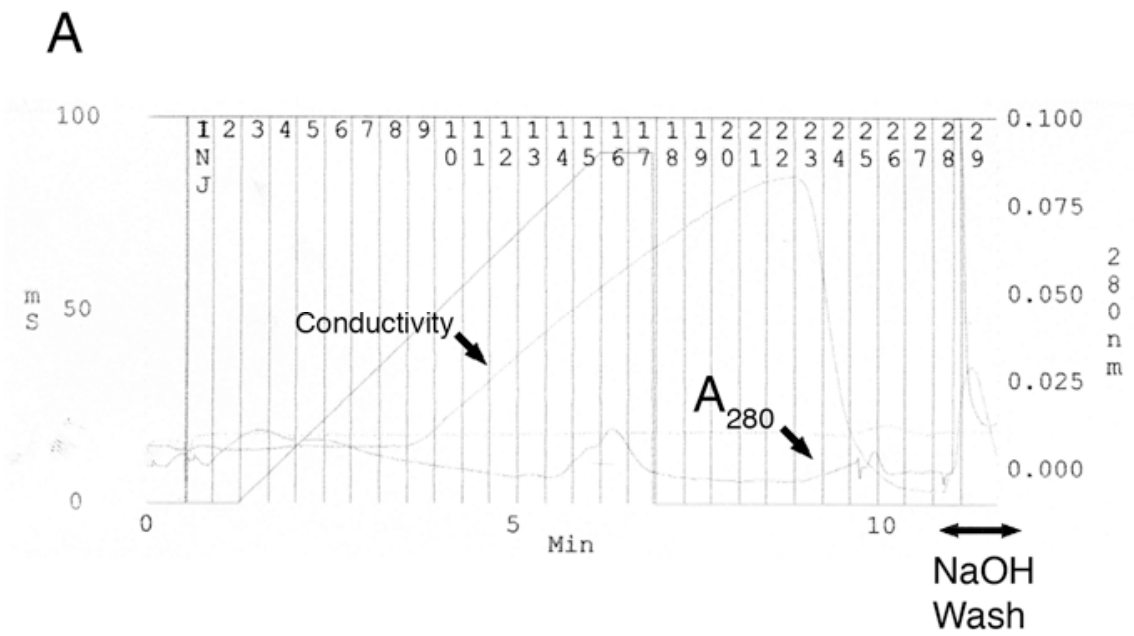

Figure 3.16 (A): **Analytical Source S1 cation exchange chromatography purification of HISdNELFE. (A) Elution profile of HISdNELFE from the Source S1 column as represented by A_{280}.**

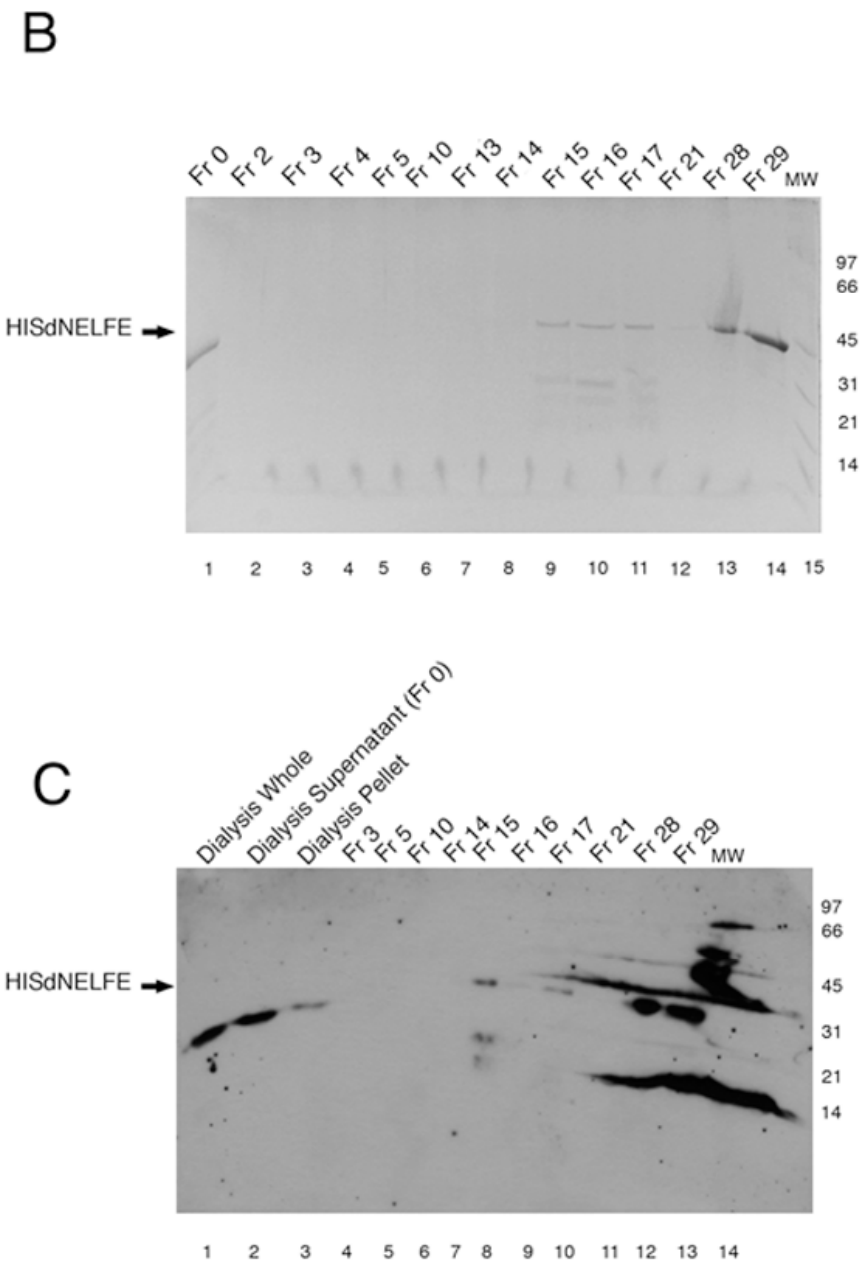

Figure 3.16 (B,C): **Analytical Source S1 cation exchange chromatography purification of HISdNELFE. (B) SDS-PAGE analysis of selected elution fractions from Source S. (C) Western blot analysis of selected elution fractions.**

Western blot analysis revealed that the majority of HISdNELFE bound to the column, but was not eluted in the increasing salt gradient. Fractions 28 and 29 demonstrated clear bands of full-length HISdNELFE in the NaOH wash that accounted for most of the input

sample. HISdNELFE appeared as a full-length protein in fractions 15, 16 and 17, corresponding to small, unresolved peaks in the chromatogram. Fraction 15 clearly showed low molecular weigh contaminants of less than 25 kDa. These contaminants were not visible in fraction 0 or in the NaOH wash fractions. These results suggested that the low molecular weight contaminants eluting with HISdNELFE were truncations of the full-length HISdNELFE that still contained the full N-terminus of HISdNELFE.

4. Discussion

4.1 Cloning of dNELF Subunits into the Polycistronic Expression Plasmid

Cloning of all four subunits of the dNELF complex into pST50 transfer cassette expression plasmids was successful. Further cloning of all four subunits into the pST44 polycistronic expression plasmid allowed the creation of the tricistronic expression plasmid pST44-dNELFD-dNELFB-HISdNELFE, but the dNELFAx1 subunit could not be subsequently cloned into this plasmid. It is likely that the large size of the dNELFAx1 gene is partially to blame, as very large plasmids are not well sustained by *E. coli* (Hansen and Olsen, 1978). Since the dNELFAx1 gene has successfully been cloned into some plasmids, and has been used to successfully transform TG1 competent cells, it is not likely that uninduced expression of dNELFAx1 is lethal to *E. coli.* It is possible that the complete dNELF complex could be lethal to *E. coli*.

Creation of pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE was attempted with use of HB101 cells as the host strain because they are deficient in RecA, which can remove large pieces of extrachromosomal DNA (Bi and Liu, 1994). So far, the use of HB101 cells has not fully alleviated all of the difficulties, but some transformant colonies have shown the desired DNA fragments produced by PCR screening with dNELFAx1-specific oligonucleotide primers, a result not found using TG1 cells. Of the transformant colonies that produced the desired PCR products, none have produced the desired restriction mapping DNA fragments. The continued use of HB101 cells may be useful in creating

pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE.

To circumvent the inability to create the pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE plasmid, an alternative strategy is to express the existing pST44-dNELFD-dNELFB-HISdNELFE plasmid in combination with another plasmid containing dNELFAx1. The pST44 vector confers ampicillin resistance, so the plasmid containing dNELFAx1 should contain resistance to another antibiotic to allow robust selection of host cells that have been transformed with both plasmids (McNally *et al.,* 1988). Attempts have been made to clone dNELFAx1 into the pST53 vector, which confers kanamycin resistance, but have so far been unsuccessful. Again, the large size of the dNELFAx1 gene could be partially to blame. It seems that the creation of a plasmid containing dNELFAx1 that confers resistance to an antibiotic other than ampicillin is more feasible than creating pST44-dNELFAx1-dNELFD-dNELFB-HISdNELFE, so there is merit in this avenue and it should be pursued.

The recombinant human NELF complex has been expressed and purified by multiply infecting insect Sf9 cells with recombinant baculoviruses (Narita *et al,* 2003), so perhaps there is something to be learned from the human version of the NELF complex that could aid in expression of the dNELF complex. This approach may have been more successful because dNELFA contains a large nonconserved region of over 800 amino acids between the N- and C-termini that is not present in hNELFA. To alleviate problems possibly caused by the large size of dNELFA, a version could be created lacking this nonconserved region. While this may aid in the expression of dNELF, it may not preserve

the functionality of the complex. Alternatively, expression of hNELFA with dNELFD, dNELFB and dNELFE could also be pursued. A clear advantage of expressing the NELF complex in insects is the presence of eukaryotic machinery that may impart necessary post-translational modifications. Zhang *et al*. proposed that some of the functions of NELF, such as dissociation from Pol II, might require post-translational modifications (2007).

4.2 Expression and Purification of dNELF Subunits and Subcomplexes

All of the plasmids that were created showed some degree of expression of proteins with the desired molecular weight with the possible exception of pST50Tr-dNELFAx1. Expression levels varied depending upon the dNELF subunit, the temperature of expression and the concurrent coexpression of other dNELF subunits. Talon affinity purification proved to be a valuable method for purifying the HISdNELFE subunit, although other subunits did not always copurify with it.

dNELFB was expressed at 37ºC, 28ºC and 18ºC, resulting in high levels of expression at all temperatures. At all temperatures, dNELFB appeared to be mostly insoluble. The optimal expression temperature was 28ºC due to the highest levels of expression, despite low solubility.

HISdNELFE was expressed at 37ºC, 28ºC and 18ºC, demonstrating high levels of expression at 18ºC and decreasing expression levels at higher temperatures. Talon

purification revealed that HISdNELFE bound the resin and was eluted with imidazole consistently at all temperatures of expression. HISdNELFE showed roughly two-thirds solubility at 18ºC expression and decreasing solubility at higher expression temperatures. Optimal expression conditions for Talon purification were 16+ hours at 18ºC.

dNELFD was expressed at 37ºC and 18ºC, but showed low levels of expression at both temperatures. Solubility tests revealed that dNELFD was ~50% soluble at both temperatures of expression.

dNELFAx1 was expressed at 37ºC, 28ºC and 18ºC, however the presence of dNELFAx1 could not be confirmed at any temperature. In some cases, a protein roughly the size of dNELFAx1 was present, but it was expressed at levels comparable to the background *E. coli* proteins. This protein in question showed partial solubility. No optimal temperature of expression was determined for dNELFAx1.

dNELFB and HISdNELFE were coexpressed at 37ºC, 28ºC and 18ºC, exhibiting high levels of HISdNELFE expression at all temperatures and a medium level of dNELFB expression at 37ºC. dNELFB expression decreased at lower expression temperatures. Talon purification showed that HISdNELFE bound the resin and was eluted with imidazole consistently at all temperatures of expression. Copurification of dNELFB with HISdNELFE could not be confirmed. Consistent with previous solubility tests, HISdNELFE displayed maximal solubility at 18ºC but even at 37ºC, slightly less than 50% was soluble. dNELFB was slightly less than 50% soluble at 37ºC, but expression

levels were low at lower temperatures of expression. The optimal expression conditions for Talon purification were 37ºC expression for 4 hours.

dNELFD-dNELFB-HISdNELFE was expressed at 37ºC, 28ºC and 18ºC, demonstrating medium levels of HISdNELFE expression and low levels of dNELFD and dNELFB expression at all temperatures of expression. Talon purification showed that HISdNELFE bound the resin and was eluted with imidazole consistently at all temperatures of expression. Both dNELFD and dNELFB were roughly 50% soluble at all temperatures. A very small amount of what appeared to be dNELFB copurified with HISdNELFE in the elution fractions at ~66 kDa. Optimal expression conditions for Talon purification were 16+ hours at 18ºC.

Coexpression of dNELFB and HISdNELFE showed that each protein was expressing, but upon Talon purification, dNELFB did not copurify. This could partially be due to a lower expression level and solubility of dNELFB than HISdNELFE, but literature would suggest that there would be an interaction between these two proteins. Upon studying NELF subunit interactions, Narita *et al.* found interactions between hNELFE and hNELFB (2003).

Coexpression of dNELFD, dNELFB and HISdNELFE revealed that all proteins were expressing and were partially soluble, yet there was only a small amount of copurification of dNELFB in the elution fractions. Again, this could have been due to higher levels of dNELFB or it could have been a genuine interaction with HISdNELFE. Narita *et al.*

discovered interactions between hNELFB and hNELFE within the complete NELF complex. However, with low amounts of expression of dNELFD and dNELFB, and because hNELFB has been shown to migrate ~3 kDa faster by SDS-PAGE than expected (Narita *et al.,* 2003), these experiments have not confirmed the identity of the small amount of coprecipitation observed at ~66 kDa.

The true goal of these copurification tests was to express all four subunits of the dNELF complex in *E. coli.* Previous studies have shown the ability to purify recombinant hNELF complex using only a N-terminal tag on hNELFE (Narita *et al.,* 2003). Instability of NELF subcomplexes is not indicative of the stability of the complete complex, and the complex may require all four subunits to become stabile. In the presence of the full complex, individual subunits could be aided in folding properly, thus decreasing insolubility. Post-translational modifications may yet be required that *E. coli* do not perform, which could affect the assembly and functionality of the complex.

4.3 Large-Scale Purification of HISdNELFE

Large-scale expression and purification of HISdNELFE was performed in hopes of developing a reliable source of purified, milligram-quantity protein for biochemical assays or structural studies. After multiple experimental steps of purification, HISdNELFE was partially purified, but these purification steps did not remove truncations of the full-length protein.

The crude extract from a 6-liter expression at 18ºC of the HISdNELFE protein was first purified over a Talon column via the engineered 6x Histidine affinity tag. The high isoelectric point of HISdNELFE then prompted purification by cation exchange chromatography, which was largely ineffective because of difficulty in eluting bound HISdNELFE during the salt gradient.

Hydrophobic interaction chromatography was attempted with Source ISO and Source PHE. Source ISO hydrophobic interactions did not prove strong enough to bind HISdNELFE. Source PHE hydrophobic interactions proved to bind HISdNELFE so strongly that HISdNELFE did not elute during the salt gradient.

Solubility tests revealed that HISdNELFE was almost entirely soluble at pH 5.5, so cation exchange chromatography was repeated at pH 5.5, but little HISdNELFE eluted during the salt gradient. The small amount that was eluted during the salt gradient was determined by Western blot analysis with anti-HIS tag antibodies to contain a high proportion of truncated HISdNELFE that was degraded at the C-terminus.

Truncated products of the full-length HISdNELFE ultimately proved too difficult to remove in the limited purification steps that were performed. Because the truncations shared many features with the full-length protein, including the 6x Histidine affinity tag, the proteins were not easily distinguishable.

The occurrence of truncated HISdNELFE is not entirely surprising. Narita *et al.* cited N-terminally FLAG-tagged hNELFE to appear as multiple bands, which they cited as resulting from proteolytic degradation at the C-terminus (2003). This could suggest that the degradation that I observed was not solely caused by experimental error and may have been caused by intrinsic properties of the polypeptide sequence. Despite the appearance of degradation products, Narita *et al.* were still able to copurify the entire NELF complex with FLAG-tagged hNELFE. It is likely that degradation of the C-terminus would be detrimental to NELFE because of the conserved C-terminal RRM motif that is essential for NELF function.

Some chromatography purifications did reveal that truncated HISdNELFE could possibly be separated from the full-length protein. Western blot analysis revealed that, while the majority of the HISdNELFE that was loaded onto the column did not elute during the salt gradient, a significant portion of the hexahistidine-tagged truncations were present in fractions that were eluted during the salt gradient.

Perhaps the best way to ensure the separation of full-length HISdNELFE from truncations would be to use a C-terminal affinity tag. Purification by N- and C-terminal tags would ensure that only full-length proteins would be purified. Affinity chromatography might also be useful when considering the difficulty experienced in attempting to bind HISdNELFE to a column and eluting it during the salt gradient.

Works Cited

Aiyar, S.E., Sun, J.L., Blair, A.L., Moskaluk, C.A., Lu, Y.Z., Ye, Q.N., Yamaguchi, Y., Mukherjee, A., Ren, D.M., Handa, H. et al. (2004). Attenuation of estrogen receptor alpha-mediated transcription through estrogen-stimulated recruitment of a negative elongation factor. Genes Dev., *18*, 2134–2146.

Bi, X. and Liu, L.F. (1994). *recA*-independent and *recA*-dependent Intramolecular Plasmid Recombination: Differential Homology Requirement and Distance Effect. Journal of Molecular Biology. *235*, 414-423.

Boehm, A.K., Saunders, A., Werner, J. and Lis, J.T. (2003). Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. Mol. Cell. Biol., *23*, 7628–7637.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T. and Peterlin, B.M., (2004). Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol. Cell. Biol., *24*, 787–795.

Gilchrist, D.A., Nechaev, S., Lee, C., Saikat, K.B.G., Collins, J.B., Li, L., Gilmour, D.S., Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev., *22*, 1921-1933.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Hansen, J.B. and Olsen, R.H. (1978). Isolation of Large Bacterial Plasmids and Characterization of the P2 Incompatibility Group Plasmids pMG1 and pMG5. J. Bacteriol. *135*, 227-238.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. Nature *436*, 876–880.

Mandecki, W., Hayden M.A., Shallcross, M.A., Stotland, E. (1990). A totally synthetic plasmid for general cloning, gene expression and mutagenesis in Escherichia coli. Gene. *94,* 103-107.

McNally, E. M., Goodwin, E. B., Spudich, J. A., and Leinwand, L. A. (1988). Coexpression and assembly of myosin heavy chain and myosin light chain in *Escherichia coli.* Proc. Natl. Acad. Sci. USA. *85***,** 7270–7273.

Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D.K., Hasegawa, J., Omori, M., Inukai, N. et al. (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Mol. Cell. Biol., *23*, 1863–1873.

Niiranen, L., Espelid, S., Karlsen, C.R., Mustonen, M., Paulsen, S.M., Heikinheimo, P, Willassen, N.P. (2007). Comparative expression study to increase the solubility of cold adapted *Vibrio* proteins in *Escherichia coli*. Protein Expr Purif. *52*, 210-218.

Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev. *10*, 2657–2683.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569–577.

Radonjic, M., Andrau, J.C., Lijnzaad, P., Kemmeren, P., Kockelkorn, T.T., van Leenen, D., van Berkum, N.L., and Holstege, F.C. (2005). Genome-Wide Analyses Reveal RNA Polymerase II Located Upstream of Genes Poised for Rapid Response upon S. cerevisiae Stationary Phase Exit Mol. Cell *18*, 171–183.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. Nat. Rev. Mol. Cell Biol. *7*, 557–567.

Tan, S. (2001). A Modular Polycistronic Expression System for Overexpressing Protein Complexes in *Escherichia coli*. Protein Expr Purif. *21*, 224-234.

Tan, S., R. Kern and Selleck, W. (2005). The pST44 Polycistronic Expression System for Producing Protein Complexes in *Escherichia coli*. Protein Expr Purif. *40*, 385-395.

Tolia, N. H. and Leemor, J.T. (2005). Strategies for protein coexpression in *Escherichia coli*. Nature Methods. *3*, 55-64.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F. et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev., *12*, 343–356.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D. and Handa, H. (1998). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. EMBO J., *17*, 7395–7403.

Wright, T. J., J. L. Costa, C. Naranjo, P. Francis-West, and M. R. Altherr. (1999). Comparative analysis of a novel gene from the Wolf-Hirschhorn/Pitt-Rogers-Danks syndrome critical region. Genomics *59*, 203–212.

Wu, C.H., Lee, C., Fan, R., Smith, M.J., Yamaguchi, Y., Handa, H., and Gilmour, D.S. (2005). Molecular characterization of *Drosophila* NELF. Nucleic Acids Res. *33*, 1269–1279.

Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H. and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in *Drosophila*. Genes Dev., *17*, 1402–1414.

Yamaguchi, Y., N. Inukai, T. Narita, T. Wada, and H. Handa. (2002). Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. Mol. Cell. Biol. *22*, 2918–2927.

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J. and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell, *97*, 41–51.

Yamaguchi, Y., J. Filipovska, K. Yano, A. Furuya, N. Inukai, T. Narita, T. Wada, S. Sugimoto, M. M. Konarska, and H. Handa. (2001). Stimulation of RNA polymerase II elongation by hepatitis delta antigen. Science *293*, 124–127.

Ye, Q., Y.F. Hu, H. Zhong, A. C. Nye, A. S. Belmont, and R. Li. (2001). BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. J. Cell Biol. *155*, 911–921.

Zhang, Z., Klatt, A., Gilmour, D.S., Henderson, A.J. (2007). Negative Elongation Factor NELF Represses Human Immunodeficiency Virus Transcription by Pausing the RNA Polymerase II Complex. Journal of Biological Chemistry. *282,* 16981-16988

Appendix A

These protocols were taken by protocols adapted and written by Dr. Song Tan.

Site Directed Mutagenesis with Pfu polymerase

Perform linear amplification of the plasmid with two mutagenesis oligonucleotides. Set up the following reaction in 0.2 ml PCR tubes on ice:

Water	17.7 µl
10xPfu Buffer	2.5 µl
2.5 mM dNTP	2.5 µl
10 ng/µl plasmid	0.5 µl
10 µM mutagenesis oligo 1	0.7 µl
10 µM mutagenesis oligo 2	0.7 µl
2.5 units/µl Pfu Turbo polymerase	0.4 µl
	25 µl

Preparation of Vector DNA

Digest the plasmid DNA with appropriate restriction enzymes. Examples of sample digests:

Water	19 µl	water	15 µl
1 µg/µl plasmid DNA	1 µl	0.2 µg/µl plasmid DNA	5 µl
10xNEBuffer	3 µl	10xNEBuffer	3 µl
1 mg/ml BSA	3 µl	1 mg/ml BSA	3 µl
100 mM DTT	1 µl	100 mM DTT	1 µl
10-20 units/µl restriction enzyme 1	1.5 µl	10-20 units/µl restriction enzyme 1	1.5 µl
10-20 units/µl restriction enzyme 2	1.5 µl	10-20 units/µl restriction enzyme 2	1.5 µl
	30 µl		30 µl

Preparation of Insert DNA

Digest the DNA with appropriate restriction enzymes to release insert. Examples of sample digests:

Water	17 µl	water	5 µl
1 µg/µl plasmid DNA	3 µl	PCR product in TE (10, 0.1)	15 µl
10xNEBuffer	3 µl	10xNEBuffer	3 µl
1 mg/ml BSA	3 µl	1 mg/ml BSA	3 µl
100 mM DTT	1 µl	100 mM DTT	1 µl
10-20 units/µl restriction enzyme 1	1.5 µl	10-20 units/µl restriction enzyme 1	1.5 µl
10-20 units/µl restriction enzyme 2	1.5 µl	10-20 units/µl restriction enzyme 2	1.5 µl
	30 µl		30 µl

Ligation (Sticky-ended DNA)

Set up (vector only) and (vector + insert) ligations as follows:

	A	B
water	5.5 µl	4 µl
10xT4 DNA ligase buffer	1 µl	1 µl
100 mM DTT	0.5 µl	0.5 µl
gel purified vector DNA (~15 ng/µl)	2 µl	2 µl
gel purified insert DNA		1.5 µl
40 units/µl T4 DNA ligase	1 µl	1 µl
	10 µl	10 µl

PCR Screening of Colonies

Prepare PCR reaction mix (volumes are in µl):

	6 samples	8 samples	12 samples	16 samples	18 samples
water	90.3	116.1	180.6	232.3	258.0
10xThermoPol buffer	14	18	28	36	40
2.5 mM dNTP	14	18	28	36	40
10 µM forward primer	7	9	14	18	20
10 µM reverse primer	7	9	14	18	20
2 units/µl Pfu pol	0.7	0.9	1.4	1.8	2.0
	133	171	266	342	380

For a 20 µl reaction, add 19 µl of PCR reaction mix to a 0.5 ml Eppendorf tube and add 1 µl of cell suspension.

Academic Vita of Nathan Hazi

126 North Barnard St. Rear
State College, PA 16801
nathan.hazi@gmail.com

EDUCATION

Pennsylvania State University, University Park, PA
B.S. in Biochemistry and Molecular Biology, May 2010
Minor in Business/Liberal Arts
Honors in Biochemistry and Molecular Biology

University of Queensland, Australia, St. Lucia, Queensland
Study Abroad with a focus on marine biology,
February 2008-June 2008

EXPERIENCE

Biochemistry and Molecular Biology Laboratory, Dr. Song Tan
Undergraduate Researcher *May 2007-May 2010*
•Expressed and purifying the *Drosophila* Negative Elongation Factor complex
•Experienced in a wide variety of laboratory techniques including molecular cloning, protein expression and HPLC purification

Environmental Biology Intern, U.S. Fish and Wildlife Service
Habitat Specialist *May 2009-August 2009*
•Created and enhanced habitat for native fish and wildlife in Pennsylvania

Coral Reef Ecology Research, Dr. Lexa Grutter
Undergraduate Researcher *February 2008-June 2008*
•Examined the interactions of Gnathiid Isopods and the changing Great Barrier Reef ecosystem

Action Potential Science Experience
Camp Mentor *June 2006-August 2006*
•Guided grade K-8 children through a wide variety of scientific subjects
•Supervised 6-8 children at all times throughout the day