PROTEIN-PROTEIN INTERACTIONS INVOLVED IN THE BIOGENESIS OF EUKARYOTIC SMALL RIBOSOMAL SUBUNITS

by

Cathy Lynn Castle

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

December 2008

©COPYRIGHT

by

Cathy Lynn Castle

2008

All Rights Reserved

APPROVAL

of a thesis submitted by

Cathy Lynn Castle

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Mensur Dlakić, PhD

Approved for the Department Microbiology

Michael Franklin , PhD

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Cathy Lynn Castle

December 2008

DEDICATION

I would like to thank the members of my committee, Dr. Mensur Dlakić, Dr. Sandra Halonen, and Dr. Valérie Copié for their guidance and support throughout my graduate school career. Specifically, to Mensur for giving me the opportunity to work in his lab on this project, to Sandra for assistance with all things related to cell culture and microscopy, and to Valerie for providing encouragement and forever drawing out the "shy people" in the classroom.

I also would like to thank my family for their encouragement and understanding during the last several years, especially my father who always believed, "It can be done." Thanks Dad.

TABLE OF CONTENTS

1. OVERVIEW OF RIBOSOME BIOGENESIS 1

2. RIBOSOME ASSEMBLY IN EUKARYOTES 5

Overview of Eukaryotic Ribosome Assembly 5
Synthesis Begins in the Nucleolus 7
Ribosomal Gene Structure 8
Modifications and Cleavage of the Primary Transcript 9
Processing Factors Involved in Ribosome Biogenesis 12
Final Processing in the Cytoplasm 17

3. SUMMARY OF RIBOSOME ACCESSORY PROTEINS 19

Proteins of the Small Subunit Processome and Early Cleavage Steps 20
Emg1p (Essential for Mitotic Growth) 21
Enp2p (Essential Nuclear Protein 2) 21
Fcf1p (Faf1p Copurifying Factor) 22
Fcf2p (Faf1 Copurifying Factor 2) 23
Gar1p (Glycine and Arginine Rich) 24
Krr1p (KRR-R motif) 24
Utp4p (U3 associated Protein) 25
Utp23p (U3 associated Protein) 26
Proteins Involved In Both Early and Late Processing Steps 26
Enp1 (Essential Nuclear Protein) 26
Dim1p (Dimethylase) 27
Rrp20p (Required for RNA Processing) 28
Proteins Involved in Export of Subunits to the Cytoplasm 29
Ltv1p 29
Proteins Involved with D Cleavage and Maturation in the Cytoplasm 29
Nob1p (Nin1 Binding protein) 29
Tsr1p (Twenty S rRNA accumulation) 30
Rio2p 31
Hrr25p 31

4. PROTEIN-PROTEIN INTERACTIONS 33

Previous Methods Utilized To Determine Protein-Protein Interactions 33
Predictive Methods 33
Yeast Two-Hybrid Analysis 34
Tandem Affinity Purification 34
Microarrays 35

TABLE OF CONTENTS - CONTINUED

Disadvantages of High-Throughput Methods 36
Detecting Protein-Protein Interactions with BiFC 37
Fluorescent Proteins 37
Variations of Green Fluorescent Protein 39
Principle of the BiFC Assay 40

5. MATERIALS AND METHODS 43

The Gateway Cloning Method 43
Obtaining DNA for Each Gene of Interest 51
Cultivating Plasmids 51
Amplifying Genes of Interest from a Genomic DNA Library 51
Preparing the Gene of Interest for Incorporation into Gateway Vectors 52
Incorporating PCR-Amplified Genes into Gateway Vectors 52
Creating Donor Constructs 52
Transformation of E. coli cells 53
Generating Destination Constructs 54
Screening Plasmids 55
Preparing DNA for BiFC Experiments 58
Cell Culture Methods 58
Cos-7 Cells 58
Cell Culturing Conditions 59
Transfection Protocol 59
Microscopy 60

6. RESULTS 61

7. SUMMARY AND CONCLUSIONS 80

8. FURTHER STUDIES 92

REFERENCES CITED 95

APPENDIX A: PCR Primers 106

LIST OF TABLES

Table Page

3.1: Human and yeast homologs of non-ribosomal protein genes 20

6.1: Localization signal of full-length Venus constructs 62

6.2: Constructs created with the Gateway system 63

6.3: Non-interacting pairs tested using the BiFC assay 64

LIST OF FIGURES

Figure Page

1.1: Computerized model of a functioning ribosome 1

1.2: Simplified gene organizational structure in prokaryotes and eukaryotes 4

2.1: Schematic diagram of the early cleavage events involved in biogenesis of eukaryotic ribosomes 10

2.2: Cleavage and processing sites involved in biogenesis of small and large ribosomal subunits 12

4.1: Model of a fluorescent protein with fragmentation points indicated. 39

4.2: Principle of the BiFC assay 42

5.1: Gateway cloning system showing generation of donor and destination constructs 44

5.2: Plasmid map for the Gateway donor vector, pDONR201. 47

5.3: Plasmid map for the Gateway destination vector pDEST14 47

5.2: Plasmid maps of Venus, VC-155, and VN-173 destination vectors 49

5.3 Modular nature of Gateway cloning vectors 50

5.3: Examples of PCR method used to screen plasmids 57

6.1: Full-length Venus constructs 65

6.2: Complementation of proteins involved in early cleavage steps 65

6.3: Full-length Venus constructs showing localization of individual accessory proteins involved in cleavage at site D 66

6.4: BiFC results involving Nob1p 67

6.5: BiFC results involving Rrp20p 68

6.6: BiFC results involving Enp1p 69

LIST OF FIGURES - CONTINUED

Figure Page

6.7: BiFC results involving Rio2p 70

6.8: BiFC results involving Tsr1p..... 71

6.9: BiFC results involving Dim1p..... 72

6.10: BiFC results involving Ltv1p 73

6.11: Interaction grid for proteins tested..... 74

6.12: Localization of full-length Venus, Cerulean and RFP constructs. 75

6.13: Comparison of localization of Venus and Cerulean constructs. 76

6.14: Evidence supporting trimeric complex formation 78

6.15: Evidence against trimeric complex formation..... 79

6.16: Conflicting evidence for trimeric complex formation 79

7.1: Assembly diagram showing localization of interactions. 90

ABSTRACT

Ribosome biogenesis is a complicated process involving numerous proteins and modification factors. The process has been well-documented in prokaryotic cells where it is much less complex than the process involved in eukaryotic cells. In eukaryotes, much of what is known about ribosome biogenesis has been learned from studies in the yeast *Saccharomyces cerevisiae*. Far less has been learned about higher eukaryotes such as humans. However, among organisms in all three domains of life, ribosome structure and function is well conserved. The biogenesis of ribosomal subunits is dynamic, complicated, and, in *S. cerevisiae,* requires over 200 *trans*-acting factors for synthesis to occur. The focus of this study is synthesis of the small ribosome subunit in eukaryotes. In order to study this process in higher eukaryotes, a mammalian cell culture method was used. This method involves cloning human ribosomal accessory genes using the Gateway system, a rapid and efficient cloning method that permits parallel construction of numerous plasmids in a modular type of system, each containing the desired gene of interest in a variety of vectors. Proteins were selected based on their activity in yeast with emphasis on the final processing step of the small subunit and the proteins that are involved at that step. These proteins are Nob1p, Enp1p, Tsr1p, Rio2p, Rrp20p, Dim1p, and Hrr25p. The corresponding genes were cloned into vectors containing either the coding region for a full-length fluorescent protein or the C-terminus or N-terminus half of the fluorescent protein. The bimolecular fluorescent complementation assay (BiFC) was then utilized to detect protein-protein interactions. Results of this assay demonstrate binary interactions among pairs of this group of proteins and the location within the cell where these interactions take place.

OVERVIEW OF RIBOSOME BIOGENESIS

Ribosomes in all three domains of life are composed of a small and large subunit (Figure 1.1). Processing sites are present in the area where the large and small subunits join. In this region, messenger RNA (mRNA) enters the ribosome. The mRNA contains the code which designates the sequence of amino acids that will be produced by the ribosome. Transfer RNA molecules (tRNA) are bound at three locations within the ribosome, sites A, P, and E. In the aminoacyl (A) site, a tRNA carrying an amino acid associates with the mRNA. The sequence of the codon of the mRNA dictates which tRNA and amino acid will be incorporated into the chain. The amino acid and tRNA molecule moves through the peptidyl (P) site as a peptidyl-tRNA complex. Finally at the exit (E) site, the empty tRNA passes out of the ribosome as the growing polypeptide chain continues to be made. As processing ensues at the P site, the next tRNA molecule moves into the A site. This process continues with the joining of amino acids ultimately producing a chain, the polypeptide, which is subsequently folded to produce a mature protein [1]. In order to accomplish this, the ribosome performs two separate functions, decoding the messenger RNA (mRNA) to ensure correct polypeptide sequence assembly and formation of peptide bonds which joins each amino acid to the preceding chain. The

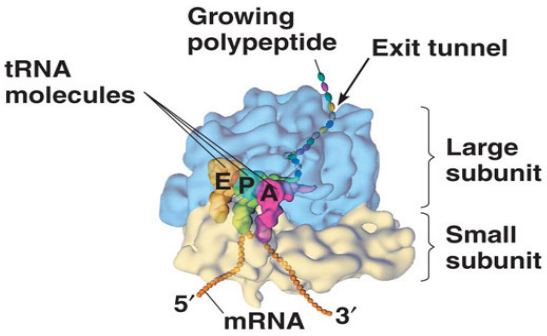


Figure 1.1 Computer model of a functioning ribosome.
© Pearson Education, Inc

decoding function is performed by the 30S subunit in bacteria and 40S subunit in eukaryotes. Peptide bond formation is catalyzed by the 50S or 60S subunit in bacteria and eukaryotes, respectively [1].

The structure and function of ribosomes are conserved throughout prokaryotes, eukaryotes, and archaea. Many of the ribosomal proteins are also conserved as is the overall process of ribosome assembly. The process is well understood in prokaryotes, where *Escherichia coli* is the model organism. In eukaryotes, much of what is known about ribosome biogenesis has been learned from studies in *Saccharomyces cerevisiae*. Much less is known about the process in archaea, although similarities exist between archaea and each of the other two kingdoms. In eukaryotes, prokaryotes, and archaea, the small subunit ribosomal RNA (rRNA) is located 5' to the large subunit rRNA. The rRNA segments for the two subunits are cotranscribed and coregulated. Once formed, the rRNA molecules fold into similar, highly condensed structures[1].

Ribosome assembly in prokaryotes has been demonstrated *in vitro*, however this has not been accomplished for the eukaryotic ribosome which requires numerous *trans*-acting proteins and small nucleolar RNA molecules (snoRNAs). Similarities exist among prokaryotic, eukaryotic, and archaeal organisms with respect to a common origin of ribosomal RNA processing. Archaeal rRNA modification is similar to that in eukaryotic organisms and also involves the use of snoRNAs to guide modification enzymes to the appropriate site of the pre-rRNA molecule [2, 3]. In archaeal species, ribosomal proteins and rRNA are more similar to what is found in eukaryotes than bacteria, however 32 ribosomal proteins have been shown to be conserved among all

three domains of life [2]. In the archaeon *Sulfolobus acidocaldarius,* a functional homologue of eukaryotic U3 snoRNA has been discovered which supports a common origin for pre-rRNA processing and snoRNA as an ancestral form [4, 5]. It has been suggested that an RNA world existed in the early evolution of life. According to this hypothesis the snoRNAs are among the ancestors of the ribosome [6, 7].

Bacteria transcribe rRNA as a single precursor molecule which is subsequently modified to produce the mature rRNA. Bacterial rRNAs are composed of 23S and 16S rRNA molecules. A tRNA molecule is located within the spacer region between the 16S small subunit rRNA and the 23S subunit. The sequences flanking the 16S rRNA gene and the 23S rRNA gene are complementary to each other and are able to base pair in the 30S precursor rRNA to form a duplex structure. This structure is then cleaved by RNAse III which results in production of the 17S and 23S pre-rRNA molecules. The remaining spacer sequences are then cleaved by exonucleolytic and endonucleolytic enzymes [8, 9]. Processing of pre-rRNA in prokaryotes appears to involve only proteins and is less complicated than in eukaryotes. In prokaryotes, all of the steps of ribosome synthesis occur in the same compartment of the cell. There is no nucleus and no separation of components or processing steps.

In eukaryotes, ribosome biogenesis begins in the nucleolus, a specialized area within the nucleus of the cell, after which pre-ribosomal RNA molecules move into the nucleoplasm and then through the nuclear pores into the cytoplasm where maturation and assembly into functional ribosomes takes place. The primary eukaryotic transcript is processed into mature 18S, 23S, and 5.8S rRNA molecules. While the tRNA coding

region present in bacterial transcripts is not present on the eukaryotic transcript, there is conservation between the 5.8S rRNA in eukaryotes and the 5' end of the 23S rRNA in bacteria [9].

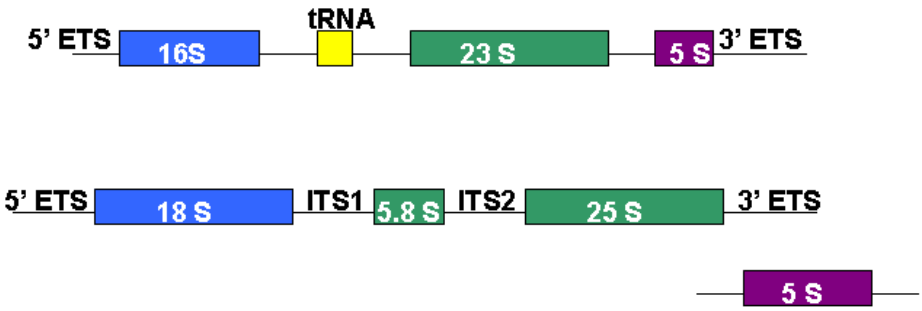


Figure 1.2: Simplified gene organizational structure in prokaryotes (top) and eukaryotes (bottom) showing the major polycistronic transcript of the large and small ribosomal units and location of internal and external spacer regions. A separate transcript produces the eukaryotic 5S rRNA.

Ribosome assembly can be carried out in bacteria in the absence of *trans*-acting factors, albeit more slowly, however this is not true for yeast ribosome synthesis in which *trans*-acting factors are essential for ribosome synthesis. In eukaryotes, most newly synthesized ribosomal proteins are imported into the nucleus before associating with pre-ribosomes. In bacteria, the ribosomal proteins bind to these pre-rRNAs immediately after being translated. Ribosomal proteins and accessory factors bind to and modify the rRNA which will ultimately become part of the 40S and 60S ribosomal subunits in a process described below.

RIBOSOME ASSEMBLY IN EUKARYOTES

Overview of Eukaryotic Ribosome Assembly

Most of what is known about eukaryotic ribosome assembly has been learned from studies carried out with *Saccharomyces cerevisiae*, or baker's yeast, although this process has been studied in a variety of diverse eukaryotic systems including plants [10-12], nematodes [13], fish [14, 15], frogs [16], and rodents [17-19] as well as numerous mammalian cell lines [20]. However, much less is known about ribosome biogenesis in higher eukaryotes and *S. cerevisiae* remains the model organism for studying this process. Upon determining the sequences of mammalian ribosomal proteins and comparing these with the genome of *S. cerevisiae*, it was discovered that close yeast homologues exist for all but one of the mammalian ribosomal proteins [21]. In addition, numerous studies have provided evidence for conservation of ribosomal processing components and pathways among eukaryotes, which gives further credibility for using this organism as a model [22, 23]**.** While differences are certain to exist among eukaryotes, ribosome biogenesis will be explained in the context of what is known about this process in *S. cerevisiae*.

The cell makes ribosomes and delivers them to the cytoplasm at a rate of about 2000 ribosomes per minute [24-26]. Ribosome synthesis is an extremely energy intensive process with more than half of the cellular resources dedicated to transcription consumed in this process [27]. Ribosomal RNA and rRNA processing is coordinated and involves a shared group of proteins [28]. Transcription of rRNA and mRNA for the

ribosomal proteins are coordinated so that equal amounts of the necessary components are produced. Transcription factors regulate rRNA expression, ribosomal proteins, and ribosome biosynthesis. The synthesis of ribosomes is highly regulated and, in addition to this, it is coordinated with other processes within the cell [29-32]. Ribosome biosynthesis is linked with transcription, recycling of mRNA, proteasome biogenesis, cell growth, and cell cycle regulation. For example, in addition to involvement in biogenesis of the small ribosomal subunit, two accessory proteins, Nop1p and Rrp20, are required for functional proteasomes [33] and depletion of *trans*-acting proteins involved in formation of the 40S pre-ribosome has been shown to inhibit progression of the cell cycle at the G1 phase [29]. Most cells grow and divide; they increase in both size and number. Only a few cell types do not grow and divide. These include cells such as neurons which grow in size but do not increase in number and cells in developing zygotes which increase in number but not size. For most cell types, however, an increase in cell size and subsequent increase in cell number occur in concert with each other. Cells require proteins in order to grow and, in order to make proteins, cells need ribosomes. Control of cell growth, therefore, is tightly linked to control of ribosome synthesis.

Eukaryotic ribosome synthesis is a complicated process involving numerous modification and cleavage events of the primary transcript produced by the cell and requires participation from numerous *trans*-acting proteins. Eukaryotic ribosomes are composed of four ribosomal RNAs and 80 ribosomal proteins along with approximately 200 accessory proteins that are required for ribosome maturation and assembly [34]. The proteins involved in ribosome assembly include enzymes that modify rRNA such as the

endo- and exonucleases, pseudouridine synthases, and methyltransferases; enzymes that are involved in folding and modifying ribonucleoproteins such as RNA helicases; proteins involved in the assembly and disassembly of protein complexes, such as GTPases and AAA-ATPases; transport molecules; and additional proteins, the functions of which have yet to be determined [25]. The ATPases, GTPases, and helicases drive and monitor the reorganization involved in synthesis of eukaryotic ribosomes. The maturation steps require modifications by enzymes that carry out exonucleolytic and endonucleolytic cleavage, pseudouridinylation, and methylation. Transport molecules relocate components to different portions of the cell and assist with export of the complexes through the nuclear pores and into the cytoplasm where final maturation and assembly of ribosomal subunits occur. Many proteins with as-of-yet unknown functions are possibly involved in assembly and transport of pre-ribosomal subunits.

Assembly is an ordered process with processing and export of the large and small subunits relying on different factors. Processing and maturation of the small subunit involves fewer steps and occurs at a more rapid rate than does that of the large subunit [20]. Ongoing ribosome biogenesis depends upon efficient recycling and shuttling of proteins within the cell to ensure that components are available in the necessary amounts for processing.

Synthesis Begins in the Nucleolus

As mentioned earlier, ribosome synthesis in eukaryotes begins in the nucleolus, a specialized subcompartment of the nucleus of the cell. The nucleolus coordinates

synthesis and assembly of the ribosomal subunits. Synthesis begins first in the dense, fibrillar region and then moves to the outer, more granular portion of the nucleolus [35]. The function of the nucleolus is linked to cell growth, proliferation, and regulation of the cell cycle in addition to synthesis of ribosomal subunits [36, 37]. Nucleoli form around repeated ribosomal gene clusters that are produced [20]. These structures are not membrane bound and much is unknown about how they are organized and assembled [38, 39]. It has been reported that approximately 90% of the yeast nucleolar proteins that have human homologues are located in the nucleolus in HeLa cells, a human cervical cancer cell line, demonstrating that the nucleolus is highly conserved among eukaryotic organisms [20].

Ribosomal Gene Structure

Ribosomal DNA (rDNA) is arranged in a consistent manner among eukaryotes [1, 40-42]. The first step in generating a ribosome is transcription of the approximately 8 kb primary transcript by RNA polymerase I. The gene for the small ribosomal subunit as well as two of the large subunit genes are cotranscribed by RNA polymerase I into a single large RNA precursor molecule, the 35S pre-rRNA. The 35S pre-rRNA associates with ribosomal proteins and *trans*-acting factors to produce a large ribonucleoprotein particle, or RNP, which is the precursor to the mature ribosome. Numerous factors are required for these processing steps, in addition to the ribosomal proteins that ultimately become part of the mature ribosome. The resulting RNP undergoes processing steps leading to formation of the mature ribosomal units, the 18S,

5.8S, and 25S rRNA molecules. Small nucleolar RNA molecules (snoRNAs) and more than 150-200 accessory proteins are needed for this processing and assembly of the ribosomal subunits. The 18S rRNA becomes part of the small 40S subunit while the 5.8S and 25S fragments are components of the large 60S ribosomal subunit [24, 25, 43]. An additional rRNA, the 5S rRNA of the large subunit, is transcribed separately by RNA polymerase III. The third RNA polymerase, RNA polymerase II, is required for transcription of the snoRNAs as well as the mRNA in the cell [42, 44]**.** Following transcription of the rDNA into rRNA, the rRNA assembles with ribosomal proteins and numerous *trans*-acting proteins to produce functional ribosomal subunits.

Modifications and Cleavage of the Primary Transcript

The rRNA coding regions in the long, primary transcript are flanked by two externally transcribed spacers (ETS), one located at the 5' end of the transcript and denoted as the 5' ETS and the other at the 3' end, or 3' ETS. The coding sequence is further divided by two internally transcribed spacers. The internally transcribed spacer 1 (ITS1) is located between the coding region for the 18S subunit and the 5.8S component of the large subunit, whereas the internally transcribed spacer 2 (ITS2) is located between the 5.8S and 25S coding regions for the large ribosomal subunit. Two subsequent cleavage events referred to as the A_0 and A_1 cleavage events occur at the 5' end in the region of the 5' ETS . These cleavages remove the 5' ETS from the coding region of what will become the 18S rRNA. The first cleavage, A_0, modifies the pre-rRNA from a 35S pre-rRNA molecule to a 33S molecule [45]. The A_1 cleavage further reduces this to

the 32S pre-rRNA molecule. A subsequent cleavage referred to as A_2 then occurs which separates the pre-rRNA into the segments that will become the large and small subunits. These segments are the 20S pre-rRNA and the $27SA_2$ pre-rRNA which, following additional cleavage and processing steps, ultimately become the small and large ribosomal subunits, respectively [46] (see figure 2.1 below). Recruitment of processing components for the large subunit occurs following cleavage at site A_2 [24].

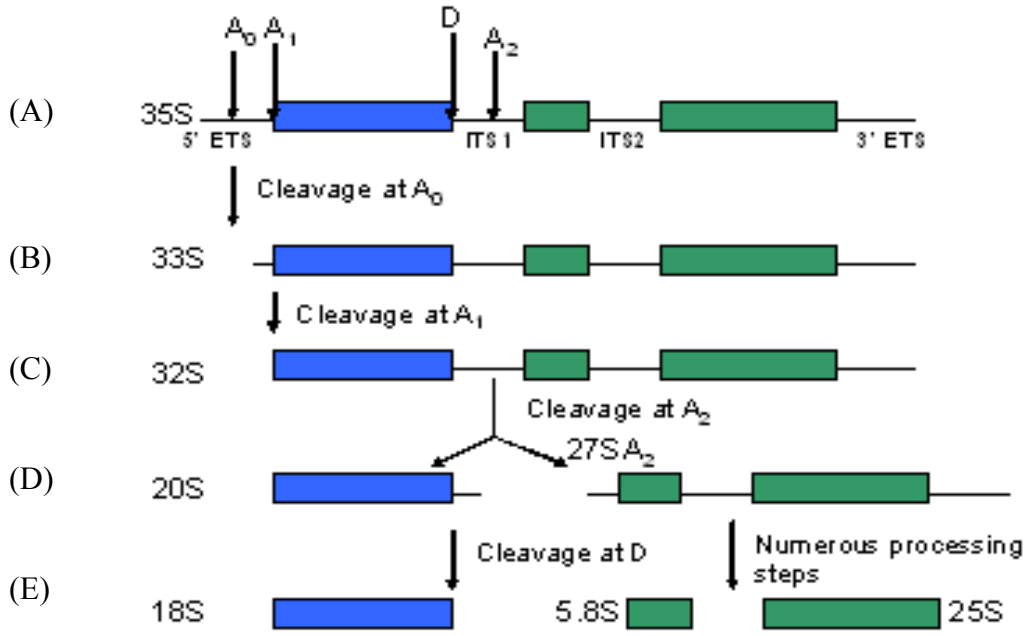


Figure 2.1: Schematic diagram of the early cleavage events involved in biogenesis of eukaryotic ribosomes. (A) Diagram showing early cleavage sites within the primary transcript. (B) Cleavage at site A_0 removes a portion of the 5' ETS followed by cleavage at site A_1 (C) where the remainder of the 5' ETS is removed. Cleavage at site A_2 (D) separates pre-rRNA of the small and large ribosomal subunits. Final processing for the small subunit pre-rRNA occurs at site D (E).

Following separation of the 35S pre-rRNA into the 20S and $27SA_2$ pre-rRNAs, precursors to small subunit contain several *trans*-acting factors including the proteins

Rio2p, Enp1p, Tsr1p, Dim1p,Rrp20p, Nob1p, Rrp12p, Hrr25p, and Ltv1p which, with the exception of Rrp12, were included in the current study and will be described in greater detail.

From this point forward in ribosome biogenesis, synthesis of the large and small subunits occur in parallel and separate from each other. For example, when components necessary for 18S synthesis are removed, mature 18S rRNAs are not produced, however 5.8S and 25S rRNA synthesis necessary for synthesis of the large subunit continues. Deficiencies in processing components for one subunit does not preclude formation of the other and synthesis of either the large or small subunit can continue unimpeded in the absence of synthesis of the other [34, 47]. In addition, small subunit components have been shown to exit the nucleoli of HeLa cells with different kinetics compared to the large subunit, supporting the evidence of independent biogenesis of the two subunits [20].

Following separation of the pre-RNA molecules for the large and small subunits, the small subunit precursor, the 20S pre-rRNA, is cleaved at site D which removes the remainder of the ITS1 from the coding region and produces the 18S rRNA (Figure 2.1-E above). This final cleavage and maturation step results in production of the 18S rRNA and occurs in the cytoplasm [48].

Several additional cleavage and processing steps occur on the $27SA_2$ pre-rRNA, which contains what will become the 5.8S and 25S rRNA molecules of the large ribosomal unit. Primary and secondary processing pathways exist for processing the $27SA_2$ pre-rRNA with cleavages at sites A_3, B_1, C_1, C_2, and E [1, 26, 34, 45, 49-52]. The

processing steps that lead to formation of mature large ribosomal subunits are more numerous and complicated than those that produce the small subunit.

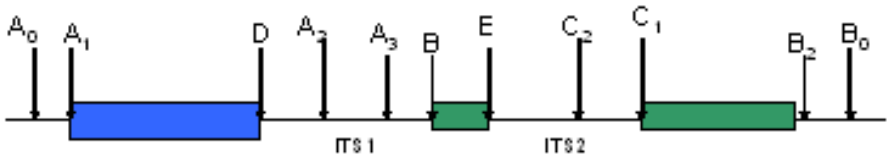


Figure 2.2: Cleavage and processing sites involved in biogenesis of small and large ribosomal subunits

Following synthesis in the nucleus, pre-rRNAs are exported through the nuclear pore and out into the cytoplasm where they undergo final maturation steps. Transport of the 40S and 60S subunits requires export factors including Xpo1p/Crm1p [42] and, although the synthesis of the two subunits are separate from one another, some export proteins are shared such as is the GTPase Ran [53]. The focus of this project, however, is biogenesis of the small subunit in eukaryotes. Accordingly, the proteins and processing steps involved in the synthesis of the small subunit will be discussed and emphasized. Proteins involved in the cleavage and modification steps to produce the mature 18S rRNA of the small subunit will be further discussed below.

Processing Factors Involved in Ribosome Biogenesis

In most organisms, rRNA is the largest constituent of the nucleic acid within the cell. Most of the nucleic acid in a growing yeast cell is noncoding RNA that has been conserved over evolution and is important for basic cellular processes [6]. Noncoding RNA molecules in the cell include the rRNA and snoRNAs involved in ribosome

biogenesis as well as tRNA. These noncoding RNAs are important for guiding modification enzymes and other proteins to the sites of the developing rRNA where they are needed to process rRNA molecules [34, 54].

As previously mentioned, numerous *trans*-acting factors are necessary for pre-rRNA processing in addition to the snoRNAs. The snoRNAs associate with these proteins to form ribonucleoprotein particles (snoRNPs). Numerous ribonucleoprotein complexes (RNPs) have been shown to be involved in rRNA processing and ribosome assembly [1, 26, 34, 49, 52, 55]. In fact, the ribosome is the cell's major RNP, accounting for approximately 10% of the cell's mass [56]. Many of the *trans*-acting factors necessary for synthesis of the small subunit are present on the 35S primary transcript, however factors involved in processing steps for the large subunit are absent at this stage [26, 34]**.** Within seconds of completion of rRNA transcription, the 40S subunit processing proteins associate with it and begin to compact the rRNA into a dense knob-like structure, the small subunit processome [50, 57, 58].

There are two kinds of snoRNPs necessary for ribosome synthesis, the C/D box snoRNPs and the H/ACA snoRNPs. The C/D box snoRNP is composed of two copies of a methyl transferase protein, Nop1p, along with three additional proteins (Nop56p, Nop58p, and Snu13p) and an RNA molecule. The H/ACA snoRNP consists of eight proteins (2 copies of Nhp2, Nop10p, Gar1p, and Cbf5p) and one RNA [54].
One function of snoRNAs is to identify nucleotides where the 2-hydroxyl position of ribose becomes methylated during maturation of the ribosome. Methylation occurs five

nucleotides from the paired region between the rRNA and snoRNA and may permit tighter packing of the ribosome by making it more hydrophobic [59, 60].

In electron micrographs, newly formed pre-rRNA molecules in yeast have large complexes located at the 5' end of the transcripts [50, 61]. These complexes contain pre-rRNA processing factors which include both pre-rRNA and U3 snoRNPs. These complexes assemble onto the transcripts creating the terminal knobs at the 5' end of the pre-rRNA molecule [62, 63]. Binding of U3 snoRNPs to the pre-rRNA molecule is required for both early cleavage events and assembly of the small ribosomal subunit [26]. U3 snoRNPs are composed of the U3 snoRNA and several proteins including Mpp10 which is specific to the 35S pre-rRNA [44].

The U3 snoRNA is complementary in sequence to the pre-rRNA found in the 5'ETS. Base pairing of these two regions facilitates the formation of the small subunit processome which is observed as the knobs at the ends of the pre-rRNA transcripts [64, 65] Proteins associated with these complexes include those necessary for cleavage at sites A_0, A_1, and A_2 and include Emg1p, Krr1p, Enp1p, Ltv1p [34]. Components of the small subunit processome have been shown to be strongly dependent upon rRNA expression in order to accumulate in the nucleoli. They are most likely recruited for biogenesis of the small subunit and are released when rRNA is not being transcribed [20]. The U3 snoRNPs have been shown to be only weakly associated with the 20S pre-rRNA and do not associate with components of the large ribosomal subunit [66].

SnoRNPs are required for processing at the A_1 site in the 5' ETS region and for processing at site A_2 [67]. Cleavage at these sites is prevented when snoRNPs are

depleted from the cells. Depletion of snoRNPs prevents synthesis of the 18S rRNA, however synthesis of 5.8S and 25S rRNA continues [68]. The snoRNPs are present in the nucleolus and associate with pre-ribosomal particles by forming hydrogen bonds between the snoRNAs and pre-rRNA molecules [65, 69]. It has been suggested that snoRNPs interact with sequences in ITS1 where they form a complex that brings 5' ETS and ITS1 close together, leaving the 18S rRNA in a loop configuration between them [7, 65]. The snoRNP complexes may bring the ends of the rRNA molecules close together in a manner which substitutes for the base pairing interactions that have been shown to exist in ribosome biogenesis in bacteria. It has also been suggested that the yeast pre-RNAs are prevented from folding completely in order to allow binding of the snoRNP particles [70]. The snoRNPs bind to the pre-rRNA which is then modified. Following modification, the snoRNPs dissociate and the pre-rRNAs are permitted to refold. Once folded, *trans*-acting proteins are able to bind to and then subsequently dissociate from the pre-rRNA [34].

Direct base pairing interactions between U3 snoRNA and the 5' ETS have been shown to be required for processing at all three of the early cleavage sites, A_0, A_1, and A_2 [71]. When deletions are made in the 5' ETS region, cleavage is also inhibited at both site A_1 and site A_2 [68]. Furthermore, accessory proteins involved at one site of cleavage may also function at other cleavage sites. For example, cleavage events at A_0 and A_1 are thought to occur with the aid of many of the same proteins. An example of this is Nsr1p which has been implicated in cleavage at both the A_1 and A_2 sites [72] and the snoRNP complex described above.

In addition to *trans*-acting factors, *cis*-acting elements are also required for 18S synthesis. These have been identified in the 5' ETS region and include the binding site for the U3 snoRNA [73, 74]. Two signals direct cleavage in the A_1 site. First, a conserved sequence is recognized upstream of the A_1 site in the 5' ETS. Secondly, cleavage occurs at a set distance from structures present within the 18S rRNA. These two signals direct cleavage to occur at the same nucleotide [74]. At the A_2 site there is a second conserved sequence, ACAC, located 3' to the A_2 site that may represent another *cis*-acting element [75]. The signals required for accurate cleavage of site A_2 are located in the 3' flanking sequence, deletion of which inhibits processing. This evolutionarily conserved sequence is located at the site of cleavage [76]. Signals necessary for cleavage at the A_2 site include the U3 binding site in the 5' ETS and at site A_1 as well as the 5' end of the mature 18S [76]. For both A_1 and A_2 cleavages, it appears that two recognition signals are present for each and these direct cleavage to occur at the same nucleotide. It has been suggested that the A_1 and A_2 sites are cleaved within snoRNP complexes and involve numerous interactions between the components of the snoRNPs and the sequences near the sites of cleavage [7, 76]. After the U3 snoRNP is released from the complex after the early cleavage steps at sites A_0, A_1, and A_2, the 20S pre-rRNA is transported to the cytoplasm where it undergoes final maturation and processing before it is folded into the final 18S RNA of the small subunit.

Final Processing in the Cytoplasm

It has been reported that 32 accessory factors are assembled with the pre-rRNA in its 35S form. Following the cleavage events at A_0, A_1, and A_2, most of these factors dissociate from the pre-rRNA and the resulting 20S pre-rRNA is exported to the cytoplasm where it undergoes the final cleavage and maturation steps necessary to generate the 18S rRNA [34]. Several processing factors are required at the final processing step to generate the 18S rRNA. These include Dim1p, a dimethylase which is also involved in earlier steps of ribosome synthesis [77, 78]; Dim2p, a protein associated with Dim1p and Nob1p that is required for small subunit biogenesis and export [79, 80]; Enp1p, an essential protein involved in small subunit biogenesis both at earlier stages and at the time of cleavage at site D; Rio1p and Rio2p, kinases which may help regulate assembly and disassembly of proteins to and from the complex through its phosphorylation activity and actions as a shuttling protein [81, 82]; Tsr1p, which is similar to Bms1p, a putative GTPase [83]; Nob1p, a nuclease that has been shown to bind to Dim2p [32, 84, 85]; Ltv1, a protein implicated in export of the small subunit [86]; and Rrp12p, a protein associated with the pre-rRNA that is also required for export [87]. These proteins have been shown to associate with each other based on tandem affinity purification studies where a protein such as Dim1p is used as a "bait" protein to find potential interacting partners, however this type of study does not demonstrate the dynamic nature of ribosome assembly. The precise order of assembly and disassembly of these various proteins has not been determined.

The proteins just described are involved in the final maturation pathway of the small ribosome subunit, which is the focus of this project. Descriptions of each of these proteins are provided in the paragraphs below. Also included are descriptions of several proteins involved in the early cleavage steps at A_0, A_1, and A_2.

SUMMARY OF RIBOSOME ACCESSORY PROTEINS

As described previously, numerous accessory proteins are involved in biogenesis of the small ribosomal subunit. It is beyond the scope of this paper to define and describe each of these and their known or suspected interactions and functions. The focus of this project has been to determine protein-protein interactions involved in the assembly of the small ribosomal subunit with an emphasis on those proteins involved in the final maturation and processing step at site D, to demonstrate the location within the cell where these interactions take place, and to further elucidate the order of assembly of the various proteins involved in this process. The genes encoding the proteins studied during the course of this project vary in size from approximately 600 to 2400 base pairs. Homologs exist in yeast however the nomenclature differs between yeast and humans. Table 3.1 lists the genes and approximate sizes of each along with the name of the yeast homolog.

Table 3.1: Human and yeast homologs of non-ribosomal protein genes and the approximate size in base pairs of each.

Gene	Aliases	Yeast Homologue	Approximate Gene Size (bp) in Yeast Cells	Approximate Gene Size (bp) in Human Cells
Dim1		YPL266W	957	828
Emg1		YLR186W	759	735
Enp1		YBR247C	1452	1314
Enp2		YGR145W	2124	2067
Fcf1		YDR339C	570	597
Fcf2		YLR051C	654	2271
Gar1		YHR089C	618	654
Hrr25		YPL204W	1485	1251
Krr1		YCL059C	951	1149
Ltv1		YKL143W	1392	1835
Nob1		YOR056C	1380	1239
Pno1	Dim2 Rrp20	YOR145C	825	759
Rio2		YNL207W	1278	1659
Tsr1		YDL060W	2367	2415
Utp4		YDR324C	2331	2061
Utp23		YOR004W	765	750

Proteins of the Small Subunit Processome and Early Cleavage Steps

The proteins of the small subunit processome are involved in processing RNA precursors to the small ribosomal subunit rRNA molecules. These proteins include the U3-associated, or Utp proteins, a group of at least 17 proteins involved in early rRNA processing. In addition, numerous other proteins are involved in the small subunit processome, including Emg1, Krr1. Some proteins, such as Enp2, are involved in early cleavage steps but do not appear to be part of the small subunit processome.

Emg1p (Essential for Mitotic Growth).

Emg1p is a member of the knot fold methyltransferase family and is conserved among eukaryotic organisms [88] as well as between eukaryotes and archaea [89]. It has been shown to be required for maturation of 18S rRNA and production of the small ribosomal subunit. Mutations in *EMG1* results in decreased production of small ribosomal subunits, but not large subunits[88, 89], further demonstrating the independent nature of processing steps for these two subunits. Emg1p interacts with RNA and in addition has been shown to form homodimers. It is involved in the endonucleolytic cleavage step in the 5' ETS and the endonucleolytic cleavage in the ITS1 where separation of the precursors for the large and small subunit rRNA occurs. Yeast two-hybrid screening has shown that Emg1p interacts with Nop14p, a nucleolar protein. Biochemical and cytological analyses have shown Emg1p to be located throughout the cell, however localization of Emg1p in the nucleus depends on its interaction with the protein Nop14p [88]. Both *EMG1* and *NOP14* are members of the repressed environmental stress response genes. This group of genes also includes genes that encode ribosomal proteins and proteins involved in both protein synthesis and RNA metabolism [30]. Emg1p is required for cell viability and synthesis of the small ribosomal subunit [88, 89]. It co-immunoprecipitates with both Mpp10p and the U3 snoRNA, suggesting that it is a component of the small subunit processome [58]

Enp2p (Essential Nuclear Protein 2).

Enp2p was shown to colocalize with components of the small subunit processome and localizes predominantly in the nucleolus [58]. Overexpressed Enp2p

accumulates in the nucleoplasm but when expressed from its own promoter, it is found predominantly in the nucleolus [58]. Enp2p, however, does not appear to be a component of the small subunit processome since it does not co-immunoprecipitate with Mpp10p, a known, primary component of the SSU processome, or the U3 snoRNA [58]. U3 snoRNA has been detected in 90S sedimentation fractions as well as fractions containing the SSU processome [90]. Further evidence that Enp2p is a component of the 90S pre-ribosome early during processing of the 35S pre-rRNA but not part of the SSU processome was obtained from a study combining sucrose gradient fractionation and immunoprecipitation experiments along with mass spectrometry in which Enp2p coprecipitated with early processing proteins but not those of the processome [90]. Furthermore, when Rrp5p, a protein required for cleavage of the 35S pre-rRNA precursor, is depleted from cells, Enp2p coprecipitates with both Pwp2p and Nan1p [90]. Depletion of Rrp5p prevents cleavage of the 35S pre-rRNA at the early processing sites of A_0, A_1, and A_2. Enp2p assembles onto the 35S pre-rRNA in conjunction with Pwp2p and a group of accessory proteins which includes Krr1p as well. This experiment demonstrates the association of Enp2p with factors necessary for processing of the large, primary 35S pre-rRNA transcript.

Fcf1p (Faf1p Copurifying Factor).

Fcf1p also known as Utp24p, has been shown to localize predominantly in the nucleolus of the cell [91, 92] and contains within it two nuclear localization signals [92]. This protein contains a PINc domain and therefore, based on homology to other proteins that contain this region, is believed to be a nuclease [91]. The PIN domain has been

shown to be required for pre-rRNA cleavage and has been further characterized as an RNAse domain [33]. Fcf1p has been shown to be involved in the early cleavage steps at A_0, A_1, and A_2 and has been suggested as the endonuclease required for cleavage of the pre-rRNA at A_1 and A_2 since a point mutation within the active site prevents cleavage at A_1 and A_2 but not A_0 [91]. When Fcf1p is depleted from the cell, accumulation of the 35S pre-rRNA occurs and an associated lack of 20S and $27SA_2$ pre-rRNA of the small and large ribosomal subunits [91, 92].

Fcf2p (Faf1 Copurifying Factor 2).

Fcf2p is a nucleolar protein [93] that has been shown to coimmunoprecipitate with the known small subunit accessory protein Faf1p and also Fcf1p [92]. When Fcf2p is depleted from the cell, like Fcf1p, decreased synthesis of pre-rRNA precursors and mature 18S rRNA occurs [91, 92]. Fcf2p is a member of a co-regulated set of genes known as the RRB regulon that are involved in synthesis of rRNA and, as a result of its association with this group, is believed to be involved in ribosome biogenesis [92]. It has been shown to interact with both ribosomal proteins and RNA-binding proteins [94]. Analysis of polyribosomes by sedimentation gradients has shown that in the absence of functional *FCF2*, mature 40S subunits are not produced [92]. Copurification of Fcf2p with both Fcf1p and Faf1p along with localization within the nucleolus suggests that Fcf2p has a role in the early modifications of the 35S pre-rRNA transcript [91, 92].

Gar1p (Glycine and Arginine Rich).

Gar1p is a component of the H/ACA snoRNPs [82, 95-97] and has been shown to localize in the nucleolus [96-98]. In addition, the nucleolar localization of proteins associated with H/ACA snoRNAs has been shown to require binding between Gar1p and the endonuclease Rnt1p [97]. Gar1p is involved with the isomerization reaction of uridine to pseudouridine on the 35S pre-rRNA and is necessary for processing at the early cleavage sites A_1 and A_2. In the absence of Gar1p, the isomerization of uridine to pseudouridine does not occur [96-99]. Gar1p requires interaction with Rnt1p in order for efficient isomerization reactions to take place [97]. While some snoRNAs, such as the U3 snoRNA, are necessary for early processing steps during processing of pre-rRNA and ribosome subunit biogenesis [45, 67], other members of this group function as guide molecules through base-pairing interactions [60, 99]. In addition to Gar1p, H/ACA snoRNPs involved in rRNA processing include Cbf5p, Nhp2p, and Nop10p [98, 100].

Krr1p (KRR-R motif).

Krr1p has been previously shown to localize within the nucleolus of yeast cells [66, 101]. It is essential for survival of the cell as well as for synthesis and assembly of small ribosomal subunits [101-103]. It has been shown to associate with Faf1p in yeast two-hybrid screening, which is also known to be involved in early processing stages [104]. Krr1p has been shown to be present in purified 90S pre-ribosomes [102] and co-immunoprecipitates with both Mpp10p and the U3 snoRNA [58], suggesting that it is a component of the small subunit processome. In addition, when Rrp5p is depleted from cells and the 35S pre-rRNA is prevented from being cleaved at the early processing sites,

like Enp2p discussed above, Krr1p coprecipitates with Pwp2p and Nan1p [90] providing evidence for the presence of Krr1p early during processing and association with the 35S pre-rRNA. In the absence of functional *KRR1*, 18S rRNA is not produced and pre-40S subunits accumulate in the nucleolus [66]. However, the 25S rRNA for the large subunit is correctly synthesized, providing further evidence for the separate nature of large and small ribosomal subunits once the 35S primary transcript has been separated into the small and large subunit components.

Utp4p (U3 associated Protein).

Utp4p is one of the numerous Utp proteins involved in processing the 35S pre-rRNA molecule. U3 snoRNAs and 28 proteins, including nucleolar Utp proteins, Utp1p through Utp17p, form a large ribonucleoprotein complex. This group of proteins can be further divided into subcomplexes. Utp4p is among the t-Utp subcomplex which also contains the protein Nan1p [66]. Utp4p has been shown to bind to Nan1p and other proteins that bind to Nan1p including Pwp2p, proteins which are linked during processing of the 35S pre-rRNA [66, 90]. Several Utp proteins including Utp6p, Utp7p, Utp9p, Utp10p, and Utp22p as well as the RNA binding protein Krr1p, have been shown to associate with Utp4p [66, 90]. Utp4p coprecipitates with Mpp10p, a marker for nucleolar localization[63]. This large ribonucleoprotein complex which contains the U3 snoRNA and proteins appears to correspond to the terminal knobs present on the pre-rRNA molecules and has been described as the small subunit processome.

Utp23p (U3 associated Protein).

Utp23p has been shown to localize primarily to the nucleolus in immunofluorescence studies [91]. This protein contains a PINc domain that is homologous to that of Fcf1p, however, unlike Fcf1p, this domain is not required for function of Utp23p [91]. Like Fcf1p, Utp23p is involved in processing of the primary 35S pre-rRNA transcript and is required for cleavage steps at the A_0, A_1, and A_2 processing sites [91]. Depletion of Utp23 results in accumulation of 35S pre-rRNA with a corresponding lack of the 20S and $27SA_2$ pre-rRNAs [91]. Further evidence for involvement of Utp23p as part of the SSU processome are experimental results showing coprecipitation of Utp23p with Rrp5p and Utp22p, both known components of the SSU processome [63, 66]

Proteins Involved In Both Early and Late Processing Steps

Enp1 (Essential Nuclear Protein).

Enp1 is conserved among eukaryotes [105, 106] and has been shown to be localized within the nucleus and concentrated in the nucleolus [66, 105, 106]. Enp1p is important for ribosome biogenesis and assembly, associating both with U3 and U14 snoRNAs early in the synthesis process of the small subunit and also during later processing steps. It is required for pre-rRNA processing and 40S ribosomal subunit synthesis [66, 106, 107]. Enp1 is involved with the endonucleolytic cleavage step in the ITS1 where large and small subunit pre-rRNAs are separated. Mutations in *ENP1* have been shown to cause processing defects at A_0, A_1, and A_2 [58, 106]. Enp1p has been

shown to associate with dimethylated 20S pre-rRNA molecules [24]. It is also phosphorylated by the protein kinase Hrr25p. Once phosphorylated, Enp1p dissociates from the pre-ribosome [108]. Coimmunoprecipitation experiments show association between Enp1p with early processing proteins, U3 and U14 RNAs, as well as with Nop1p which is involved later in processing [106].

Dim1p (Dimethylase).

Dim1p, as its name implies, is responsible for methylating the pre-rRNA during processing. One of the rRNA modifications that is conserved between bacteria and eukaryotes is dimethylation at the 3' end of the pre-rRNA of the small subunit [77, 109-111], a function which in yeast is performed by the Dim1p [77]. This enzyme was first described in *E. coli* as a member of the KsgA family of proteins [112] and has subsequently been described in eukaryotic organisms [78, 113]. Pre-rRNA processing in the nucleolus and methylation at the 3' end of the pre-rRNA has been shown to be inhibited in the absence of Dim1p with resulting accumulation of non-dimethylated pre-rRNA [114]. Cells lacking Dim1p are not processed properly at the A_1 and A_2 processing sites which generate 20S pre-rRNA, however the two dimethylated adenine residues are not themselves required for cleavage to take place [114]. Dim1p has been shown to be part of both the 90S and 40S pre-particles in nucleolus. It is important for early and late processing steps [77, 114, 115] and has been implicated as a part of a quality control mechanism during rRNA processing [77].

Rrp20p (Required for RNA Processing).

Rrp20p is also referred to as Dim2p and Pno1p and is closely linked to Dim1p. It was originally named Pno1p for its association with Nob1p (partner of Nob1p). It is an RNA-binding protein with strong similarity to other potential RNA-binding proteins in eukaryotic organisms based upon BLAST sequence analysis and comparison of genetic sequences [116]. It contains a KH domain which is an evolutionarily conserved sequence found in numerous RNA-binding proteins [117]. Rrp20p also shares homology with another eukaryotic assembly protein, Krr1p [102]. This protein is present in pre-rRNA particles [80] and binds to early pre-90S particles and remains on the pre-40S particle until final maturation in the cytoplasm. It has been shown to be associated with nucleolar and late cytoplasmic processing steps [80] and therefore shuttles between the nucleus and the cytoplasm [80, 107]. Depletion of this protein causes defects in early nucleolar cleavages at sites A_0, A_1, and A_2 [80, 116]. It has been shown in large-scale affinity-tagging experiments to associate with U3 components and 20S processing factors, Tsr1p and Rio2p [118]. In addition, Dim1p and Rrp20p have been shown to interact. Rrp20p might recruit Dim1p to the 90S pre-ribosome where it functions in the early processing steps [80]. Originally named for "partner of Nob1" as Pno1, it has been proposed that the acronym be changed to the more descriptive "Protein Needed for Ordinary A_1-A_2 cleavage" which is more descriptive of its function in ribosome biogenesis [119].

Proteins Involved in Export of Subunits to the Cytoplasm

Ltv1p.

Ltv1p is associated with pre-40 subunits and binds pre-40S rRNA [120]. Cells lacking *LTV1* produce approximately half of the number of 40S ribosomal units produced when *LTV1* is present [120]. Ltv1p contains a nuclear export sequence and is necessary for export of the pre-40S subunit into the cytoplasm. It localizes in both the nucleus and the cytoplasm and shuttles between these two cellular compartments [86]. Ltv1p binds the same region of the 40S subunit as Enp1p, Rps3p, and Rps15p [108]. Binding and phosphorylation cause the pre-40S to change shape, enabling it to fit through the nuclear pore. [86].

Proteins Involved with D Cleavage and Maturation in the Cytoplasm

Nob1p (Nin1 Binding protein).

Nob1p is a putative nuclease containing a PIN domain which is homologous to that of 5' exonucleases and is believed to be the endonuclease necessary for cleavage of the pre-rRNA at site D [32, 33]. It has also been reported to be involved in the formation of the proteasome [84]. Nob1p is essential for synthesis of the 40S ribosome subunit and is required for cleavage of the 3' end of the 18S rRNA[32]. It copurifies with the 40S pre-ribosome particles in the cytoplasm [107] and processes the pre-rRNA at site D during maturation of the small subunit rRNA [32]. Due to its presence in both the nucleus and the cytoplasm, Nob1p is believed to accompany the pre-rRNA of the small ribosomal subunit as it is exported into the cytoplasm for final processing [32]. Based on

homology modeling, a mutation was created in the active site thought to be responsible for the enzymatic activity of Nob1p. This resulted in lack of cleavage of the pre-rRNA at site D which provides evidence that Nob1p is the nuclease responsible for this cleavage [33]. Nob1p has been identified in complexes with other ribosome accessory proteins involved in late processing of small subunit rRNA including Rrp20p, Rio2p, Dim1p, and Tsr1p [66, 83, 114, 116]

Tsr1p (Twenty S rRNA accumulation).

Tsr1p and Tsr2p are required for processing the 20S pre-rRNA to the 18S rRNA. Tsr1p is related to Bms1p, a putative GTPase. Both Tsr1p and Bms1p localize in the nucleolus although they function at different steps during maturation of the small ribosomal subunit rRNA. Furthermore, these proteins are not functionally equivalent, even though they share similar sequences in both amino- and carboxy-terminal regions [83]. Whereas Bms1p blocks processing at an early step, Tsr1p is involved in final maturation leading to the mature 18S rRNA [83]. Tsr1p has been suggested to be involved with assembly of nearly mature pre-rRNA particles late in the assembly process. Tsr1p was named for its importance in the late stage of ribosome biogenesis and its role in accumulation of the 20S rRNA species (Twenty S rRNA accumulation). In the absence of Tsr1, 20S pre-rRNA molecules accumulate in the cell. It is possible that this protein also functions during transport of the pre-rRNA to the cytoplasm [83].

Rio2p.

Two protein kinases, Rio2p and Hrr25p are associated with purified pre-40S subunits [107]. Rio2p is a member of a family of atypical protein kinases. Both Rio1p and Rio2p are required for 40S biogenesis. Phosphorylation reactions must play a role in 40S biogenesis. Rio2p has been reported to be involved in processing 20S pre-rRNA [82, 121] and late maturation and export of the 40S subunit [107]. In the absence of Rio2p, dimethylated 20S pre-rRNA accumulates in the cytoplasm. Pre-40S particles were isolated by tandem affinity purification using Rio2p as bait followed by gel filtration in the presence of a high salt concentration of 100 mM $MgCL_2$. Under these conditions, Rio2p, Tsr1p, Ltv1p, Enp1p, Nob1p, Hrr25p, Dim1p, and Dim2p all dissociated from the 40S subunit, however most of the small subunit proteins did not dissociate. Mature 40S subunits were similarly not disrupted by this increased salt concentration. Enp1p and Ltv1p eluted together along with Rps3p indicating that they form a complex. Tsr1p and Rio2p, however, were eluted in later fractions, indicating that they became monomeric [108].

Hrr25p.

Hrr25p is involved in late remodeling steps of the pre-40S subunit. Hrr25p-dependent phosphorylation of a complex composed of Enp1p, Ltv1p, and Rps3p leads to dissociation of the subcomplex from the pre-40S subunits *in vitro*. The kinase activity of Hrr25p likely regulates the maturation of 40S ribosomal subunits [108]. If the phosphorylation/dephosphorylation cycle is required for nuclear export of pre-40S particles, phosphorylation should occur in nucleus and dephosphorylation in the

cytoplasm. Hrr25-depleted cells show defects in 18S rRNA maturation. In cells depleted of Hrr25p, Enp1p and Rps3p were not phosphorylated and released from pre-40S subunits [108]. When Rio2p was depleted instead, Enp1p, Rps3p, and Ltv1p were phosphorylated and dissociated. Therefore, phosphorylation of Enp1p and Rps3p, and their subsequent dissociation from the 40S pre-ribosome, is regulated by Hrr25p [108].

PROTEIN-PROTEIN INTERACTIONS

Previous Methods Utilized To Determine Protein-Protein Interactions

It is important to study protein-protein interactions since few proteins act alone. Most proteins function within complex cellular pathways and interact with other proteins either in pairs or as members of larger complexes and networks. It has been estimated that an average of five interaction partners exist for any given protein [122]. By understanding protein-protein interactions on a pair-wise basis, insight can be gained into the processes of more complicated pathways and networks of proteins.

Predictive Methods.

There are numerous methods for studying protein-protein interactions. One method to gain insight into protein-protein interactions is through the use of predictive methods. However, many predictive methods are less than 50% accurate on the basis of how precisely they identify known RNA processing proteins [119, 123, 124]. Predictive methods, however, do provide a basis for carrying out additional experiments. For example, by predicting function of a protein based on its homology to other known proteins, hypotheses can be made and experiments designed to test the hypotheses. In the absence of predictive methods, many hypotheses would likely not be developed and tested. This type of method gives researchers a “roadmap” and direction as far as interactions to test experimentally. They provide a starting point or frame of reference where one may not otherwise exist.

Yeast Two-Hybrid Analysis.

The yeast two-hybrid system is a commonly used *in vivo* system for studying protein-protein interactions [94, 125, 126]. This system makes use of the structure of transcription factors which contain a DNA binding domain and transcription activation domain. When the binding domain and transcription activation domain are separated, the transcription factor becomes inactive. Function is restored by fusing these two domains to two interacting proteins. One protein, labeled as the "bait" is fused to a DNA binding domain while the other protein, labeled as the "prey" protein, is fused to the transcription activation domain. If the "bait" and "prey" proteins interact, function of the transcription factor will be restored. While widely used, this technique has limitations for detecting weak interactions due to limited growth of yeast cells when working with low-abundance proteins [127] and is also best suited for studying interactions that occur within the nucleus [122]. It does not, however, provide information about protein interaction as they occur in real time, nor does it provide information about the localization of the interactions within living cells. This method has also been adapted to permit high-throughput examination of protein interactions between two proteins on a pair-wise basis. More than 4000 protein-protein interactions have been examined with this technique [94, 128].

Tandem Affinity Purification.

Tandem affinity purification methods have been used to determine protein-protein interaction *in vitro*. This technique can be employed in large-scale studies to discover interactions among large groups of proteins. Individual proteins are tagged and

used as bait for possible interacting proteins. This method is based on an affinity tag used in two steps of the process. A bait protein is again used as in the yeast two-hybrid method. An affinity tag is genetically added to the bait protein, and proteins bound to the bait are subsequently precipitated from the mixture [129]. All of the proteins present in the cell can be tested for potential interaction with the bait protein. The protein complexes are then purified and subjected to gel electrophoresis to separate them according to mass. They are excised from the gel, digested with trypsin, analyzed by mass spectroscopy, and identified. This technique permits analysis of an organism's complete proteome and has been used to study proteins involved in ribosome biogenesis. Hundreds of proteins have been characterized in this manner [118, 130]. However, this type of analysis does not provide information about where the protein interactions occur within the cells and problems with the reporting of both false-positive and false-negative results occur.

Microarrays.

An additional high-throughput method to analyze interactions is the use of microarrays [131]. Microarray analysis allows expression levels of hundreds or thousands of genes to be measured in a single experiment. In this type of study, genes of similar function exhibit similar expression patterns. Proteins with functions that are unknown or not well understood can be compared to more well-characterized proteins. This comparison can provide insight as to the function of some of the less-understood proteins and their functions [132]. Proteins are fixed on a slide or microarray chip. Binding interaction partners for the proteins are then detected for each protein present on

the chip. This method has also been used to study components involved in ribosome assembly [133-135]**.** Difficulties with this method include the cost of microarray experiments and the time required to process and analyze the enormous amount of data that can arise from such a study. In addition, it is less common and more specialized, requiring additional training and equipment compared to methods performed "at the bench".

Disadvantages of High-Throughput Methods.

These high-throughput approaches identify numerous protein complexes and hundreds if not thousands of potential interactions, however they do not provide information about the location of these interactions within a living cell. Protein structures can be altered, and with the altered structure comes altered function, when they are removed from their physiological environment. In addition, these high-throughput studies are unable to provide information about the order of assembly of the complexes that form. An additional complication is that these approaches tend to have a high false-positive rate when identifying interactions and fail to identify some known interactions [118, 130]. In one study it was estimated that as much as 30% of the interactions detected were false [118]. In a DNA microarray hybridization experiment which conducted 20 experiments comparing 5 functional classes, 25 genes were misclassified in at least 19 out of the 20 experiments [124]. While these 25 genes represented a small fraction of the total number of genes studied, it is still important to note that misclassifications can occur with these large-scale methods. Large-scale studies

are extremely useful, however their results must be integrated with data obtained from other types of studies to obtain more accurate information.

Detecting Protein-Protein Interactions with BiFC

A system to study binary, and in some cases ternary, interactions of ribosomal proteins involved in the biogenesis of the small ribosomal subunit was developed to further examine protein-protein interactions involved in this process. The goals of this project are to provide information about protein-protein interactions *in vivo,* under normal physiological conditions, and to observe the location within the cell where these interactions take place. By identifying the location of the interactions, it is hoped that it will be possible to gain insight as to the order of assembly and disassembly of the processing factors involved in synthesis of the small subunit of the ribosome. Previously published data and known or suspected interactions among ribosomal proteins were considered when designing the individual experiments. This system makes use of the bimolecular fluorescent complementation assay and fluorescent microscopy which will be described below.

Fluorescent Proteins.

Green fluorescent protein was first discovered in the jellyfish, *Aequorea* in 1962 [136]. It was later cloned and the DNA of the protein, aequorin, was expressed in the 1980s making it possible to use recombinant aequorin in living cells [137]. The crystal structure was later determined in 1996 [138]. Red fluorescent proteins similar to GFP were discovered in corals in 1999 [139]. In the last decade countless studies have

been done utilizing fluorescent proteins. More recently, variations of GFP have been developed that are more suited for studies in living cells at physiological conditions [140-143]. The basic structure of all known fluorescent proteins is an 11-stranded β-barrel wrapped around a central helix to form a cylinder with the fluorophore protected in the center of the structure [138]. The fluorophore that develops is formed by three residues located at amino acids Ser65, Tyr66, and Gly67. Cyclization, oxidation, and dehydration reactions at the site of these amino acids are required for development of the fluorophore. Proper folding of the protein is also required in order for fluorescence emission to occur. All fluorescent proteins share this basic structure, the difference being increased conjugation of bonds around the amino acids that make up the fluorophore. The development of fluorescent protein variants of GFP has made assays such as BiFC possible.

The BiFC assay permits visualization of protein-protein interactions in an intact cell. It is based on observing the fluorescence that is generated when two non-fluorescent halves of a fluorescent molecule are brought into contact with each other through interactions of proteins fused to each half of the fluorescent molecule [141, 144]. GFP and yellow fluorescent protein (YFP) were originally used for this type of assay. The fluorescent protein can be split into two portions and upon recombination of the halves, the protein is reformed and the fluorescent signal can be detected. Two locations where the protein can be split into two portions that can later recombine occur at amino acids 155 and 173.



Figure 4.1: Model of a fluorescent protein with fragmentation points at amino acids 155 and 173 indicated.

A nomenclature system has been developed for describing these fluorescent fragments. The fragment EYFP-N173 begins with the first amino acid at the N-terminus and ends with amino acid 173 of the fluorescent protein. The fragment EYFP-C155 begins with amino acid 155 and ends at the C-terminus. A small segment of amino acids overlap between the two fragments when using this complementation pair. Other successful combinations include the N-terminus half up to amino acid 154 with the C-terminus half beginning at 155 (N-154 and C-155) or the N-terminus half up to amino acid 172 with the C-terminus half beginning at 173 (N-172 and C-173).

Variations of Green Fluorescent Protein.

Recently new fluorescent proteins have been developed for use in the BiFC assay. Different fluorescent proteins other than GFP and EYFP were desired. Reasons for this include the fact that GFP does not fold optimally at 37°C, a quality necessary for conducting studies in living cells at physiological conditions. EYFP requires a preincubation step at lower temperatures prior to visualizing the BiFC signal [141], which also is suboptimal for observing living cells since cells must be incubated at 30°C

for up to 2 hours. Fluorescent protein fragments developed in the laboratory of Chang-Deng Hu were utilized for this study [140]. These include both Venus and Cerulean constructs. Venus shows a 12- to 15-fold increase in BiFC efficiency and 2-fold increase in specificity when compared to EYFP. In addition, Venus is less sensitive to higher temperatures, such as 37°C, and the fluorophore matures more rapidly (12 hours versus 48 hours) when compared to EYFP [140]. These factors make Venus a superior choice to EYFP. Cerulean constructs also behave superiorly to EYFP with a 2.5-fold increase in fluorescent intensity compared to ECFP [145]. Constructs containing the Venus fluorescent protein are named VN-173 which begins at the N-terminus and ends at amino acid 173 and VC-155 which begins at amino acid 155 and proceeds to the C terminus. Cerulean constructs were also made, CN-173 and CC-155. Emission and excitation wavelengths for Venus are 515 nm and 528 nm. Corresponding values for Cerulean are 439 nm (excitation) and 479 nm (emission) . Combining the N-terminal half from amino acid 1 through 172 (N-173) with the C-terminal half from amino acid 155 to the end of the protein has shown the greatest complementation efficiency compared to other combinations and Venus was identified as the most useful yellow fluorescent protein [140]. It was for those reasons that VC-155 and VN-173 constructs of the fluorescent protein Venus were chosen for this project.

Principle of the BiFC Assay.

In the BiFC assay, a gene of interest is fused to the gene encoding a full-length fluorescent protein or to either the C-terminus half or N-terminus half of the gene. The chimeric genes are then transfected into cells where the chimeric construct is expressed.

Using a gene that is fused to the full-length fluorescent protein gene permits visualization of the location in the cell where the protein encoded by the gene being studied will localize when it is overexpressed on its own within the cell. This serves two purposes. First it allows determination of the primary location of the protein within the cell and second, it acts as a control for designing the individual experiments and allows the researcher to optimize experimental conditions for the group of proteins under study.

The primary purpose of the BiFC assay is to detect protein-protein interactions within living cells. This is done by fusing the gene of interest to one half of the fluorescent protein. For example, when the gene encoding protein "A" is fused to the C-terminal half of the fluorescent protein and the gene encoding protein "B" is fused to the terminal half, interactions between protein "A" and protein "B" can be detected. If protein "A" and protein "B" interact within the cell, this interaction will bring the two fluorescent protein halves together and reconstitute the fluorescent molecule and hence the fluorescent signal. An interaction between protein "A" and protein "B" will be detected by the fluorescence observed within the cell.

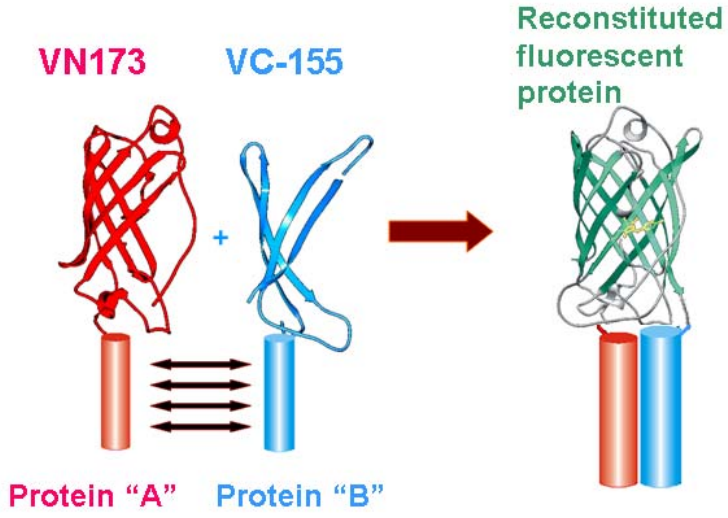


Figure 4.2: Principle of the BiFC assay. Diagram showing N-terminus and C-terminus fluorescent protein fragments fused to interacting proteins "A" and "B".

MATERIALS AND METHODS

The Gateway Cloning Method

The Gateway cloning method was a key step in the success of this project. As such, significant background information about this method is necessarily included. The Gateway cloning method can be used for more rapid and large-scale cloning compared to other methods, such as restriction enzyme cloning methods. It is a site-specific, recombination method based on bacteriophage lambda in *E. coli* [146, 147]. The Gateway cloning system allows DNA to be transferred from a donor vector to multiple destination vectors rapidly and efficiently. Restriction enzyme digests and re-ligation steps are not required. Instead, cloning is accomplished through site-specific recombination utilizing the attachment sites of bacteriophage lambda and *E. coli*. This also ensures that the open reading frame of the gene is maintained in the final product. The system recombines a gene of interest into a donor vector following which it can be easily shuttled into numerous destination vectors. There are several variations of the procedure depending upon the starting material and requirements of the individual project. The design used in this project will be discussed below. Figure 5.1 gives a graphical representation of this system.

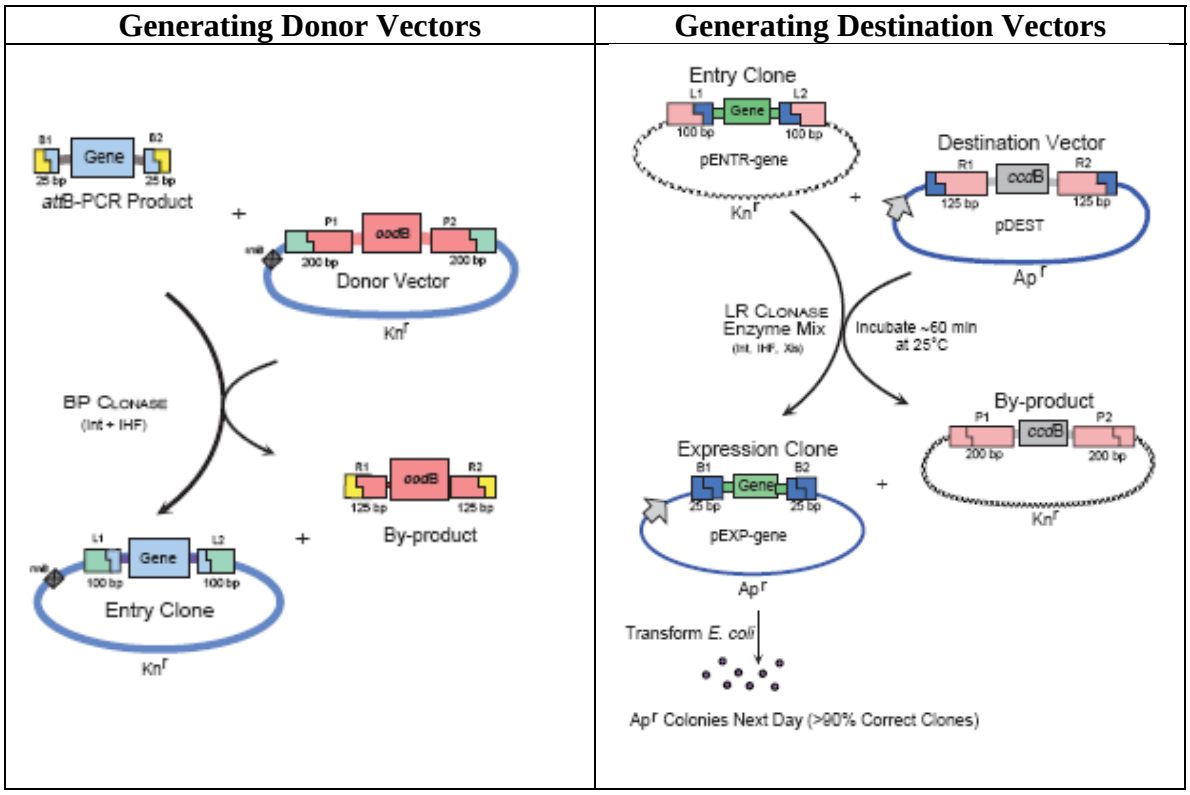


Figure 5.1: Gateway cloning system showing generation of donor and destination constructs. Figure adapted from Invitrogen Technologies.

The first step in the Gateway cloning method is to amplify the gene of interest with gene-specific primers for the gene with additional sequences designed into the ends of the primers that are complementary to the attachment sites in the vectors. This is done via the polymerase chain reaction (PCR) with a high-fidelity polymerase. A second PCR reaction is then carried out with primers specific for the attachment sites. This produces a slightly longer transcript that is then ready to be recombined into the donor vector. Following initial steps where the gene of interest is amplified by PCR, subsequent steps can be done in parallel with multiple constructs being produced at the same time.

The recombination steps are based on the recombination of bacteriophage lambda into its bacterial host, *E. coli* at specific attachment sites. These attachment sites are designated *attB* and *attP* for creating the donor construct and *attL* and *attR* which are used to create destination constructs. Two recombination proteins, Int and Xis, are also required along with the integration host factor encoded by *E. coli*.

To create the donor construct containing the gene of interest, the PCR product is combined with the donor vector. The PCR product contains the *attB* site following the two rounds of PCR amplification as a result of the primers used in the second amplification step which contain sequence regions complimentary to the *attB* site. The donor vector contains the *attP* site along with a counterselectable gene which, in the case of pDONR201, is chloramphenicol. What is referred to as the BP reaction is carried out, using the donor vector and the PCR product. The BP Clonase enzyme is added to pDONR201 and the PCR product after which the mixture is incubated at room temperature. Incubation time can be as short as a few hours. Recombination takes place at the *attB* and *attP* sites. Products of the BP reaction are then transformed into *E. coli* cells. Because the donor vector contains a kanamycin-resistance gene, successfully cloned and transformed cells will also be resistant to kanamycin. Therefore, cells are grown on media containing kanamycin. The counterselectable gene encoding chloramphenicol resistance is exchanged for the gene of interest contained in the PCR product. Colonies are then screened by PCR to ensure the gene of interest is contained within the bacterial genome. Positive colonies are then grown in kanamycin-containing

media. DNA is then harvested from these cultures and used for the next step in the Gateway cloning method.

Once a successful donor construct has been made, a second reaction is carried out. This is very similar to the BP reaction. The gene of interest is now in the donor vector and can be easily recombined into numerous destination vectors. The donor construct contains the *attL* site which recombines with the *attR* site present in the destination vector. DNA from the positively screened donor construct is combined with the destination vector along with the LR Clonase enzyme. The reaction is then incubated for several hours at room temperature. Destination vectors are designed with an ampicillin resistance gene. Also present is the counterselectable *CCDB* gene, the product of which is lethal to *E. coli* cells. However, the *CCDB* cassette is located in a portion of the vector that, once recombination has taken place, will not part of the final construct; it will exchange places with the gene of interest from the donor vector. Products of the LR reaction will be used to again transform *E. coli* cells, however ampicillin-containing media is now used due to the ampicillin-resistant nature of the plasmids. Colonies present on the plate will likely contain the gene of interest for the following reasons. First, colony growth is possible only if the *CCDB* gene has been eliminated, and the gene of interest is designed to exchange places with that gene. Second, growth on ampicillin-containing media indicates that the destination vector is otherwise intact. Colonies are screened by PCR to ensure the gene of interest is contained within the plasmid. Following positive screening and sequencing, larger quantities of the plasmid are

cultivated and used experimentally. Plasmid maps for the donor vector, pDONR201, and the destination vector, pDEST14, are shown below.

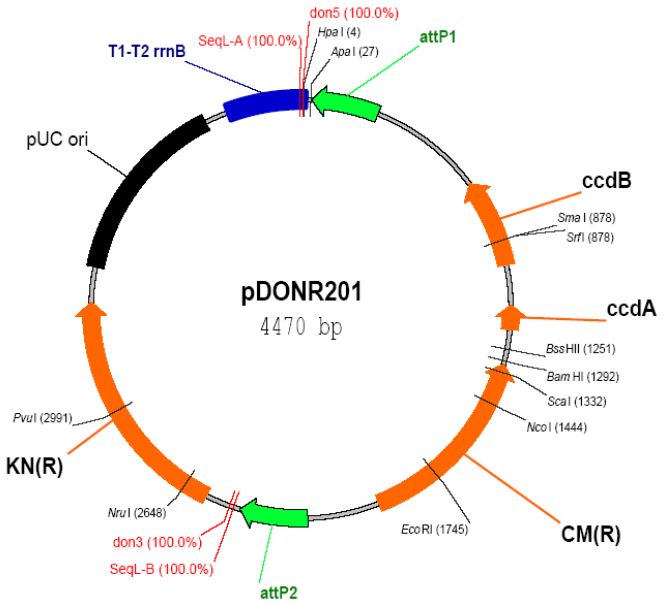


Figure 5.2: Plasmid map for the Gateway donor vector, pDONR201.

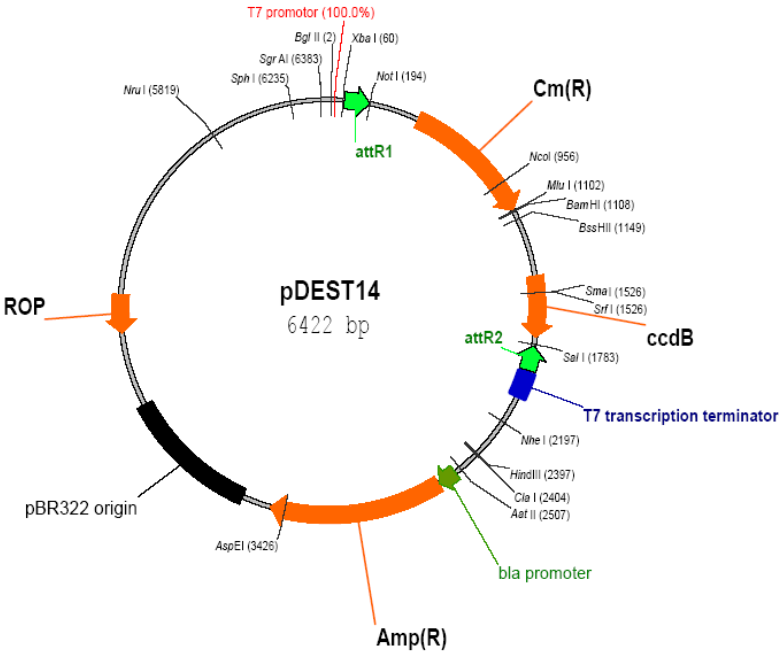


Figure 5.3: Plasmid map for the Gateway destination vector pDEST14

The destination vector above, pDEST14, was used to create constructs that were to be used for *in vitro* studies, however these experiments were not carried out. The focus of this project was on *in vivo* protein-protein interactions using the BiFC assay. For this purpose, additional destination vectors were required. These vectors contain either the coding region for the full-length fluorescent protein, the C-terminal half of the protein, or the N-terminal half of the protein. Several destination vectors were utilized in recombination reactions for each gene of interest. These included a full-length Venus construct and corresponding C-terminus and N-terminus fragments, VC-155 and VN-173. In addition, Cerulean constructs were produced for more than half of the proteins studied. These were designated Cerulean (for the full-length construct) and fragments designated CC-155 and CN-173. For a small subset of the proteins, red fluorescent constructs (mRFP) were also created. The best complementing pair was not certain, so five constructs were made for each gene that was moved into mRFP vectors. These constructs were mRFP, mRC-155, mRN-156, mRC-169, and mRN-170. Plasmid maps of destination vectors containing the gene for the full-length Venus fluorescent protein as well as the two complementation halves, VC-155 and VN-173, are shown below. The Cerulean and RFP destination vectors are similar, the only difference being in the coding region for the fluorescent protein.

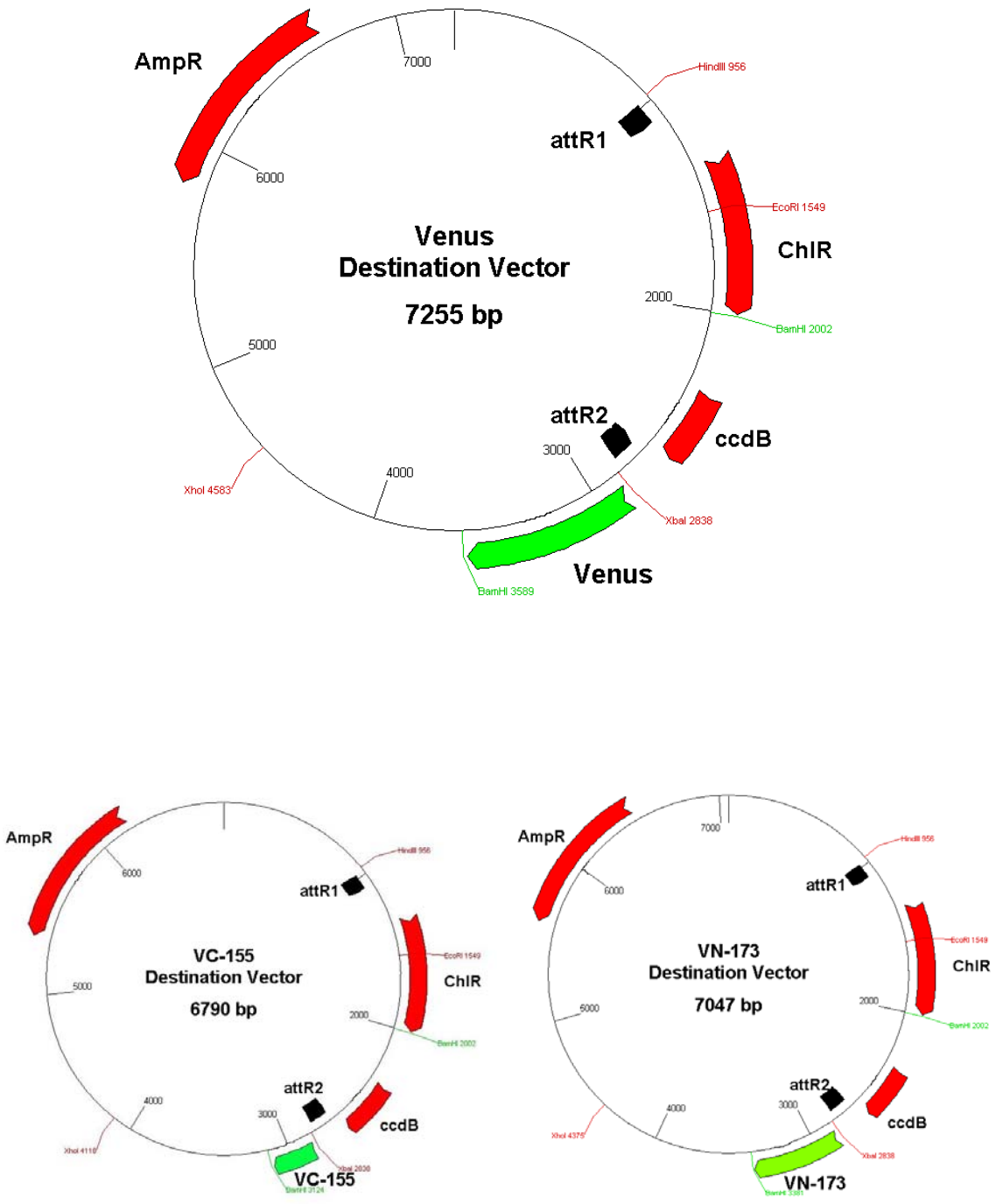


Figure 5.2: Plasmid maps of Venus, VC-155, and VN-173 destination vectors.

The Gateway system permitted multiple destination vectors to be created rapidly once the gene of interest had been cloned into the donor vector. Figure 5.3 illustrates the modular nature of this system and the destination vectors utilized once the donor vector had been created.

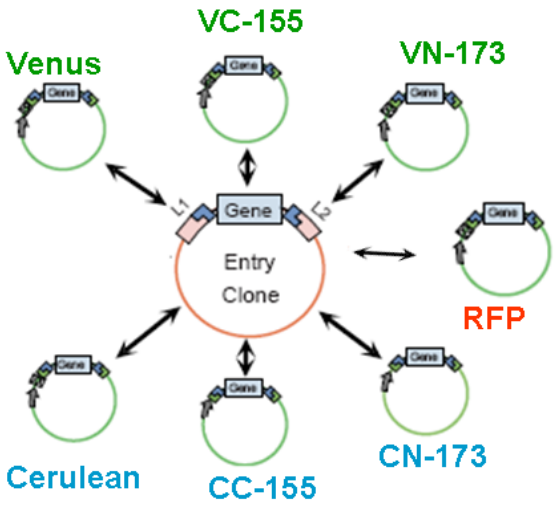


Figure 5.3 Modular nature of Gateway cloning vectors showing the entry clone in the center and destination vectors that were used during the course of this project.

Recombination from the donor vector into any of the destination vectors resulted in a construct containing the gene of interest fused to the gene for the fluorescent protein or fluorescent protein fragment. These chimeric genes were then expressed in mammalian cells, fluorescence observed, and both interactions and localization of the interactions determined. The Gateway cloning method permits expression of functionally active recombinant proteins in living cells that can be used for *in vivo* assays.

Obtaining DNA for Each Gene of Interest

Cultivating Plasmids.

Several genes of interest were purchased as plasmids from Open Biosystems. These plasmids were grown in Luria-Bertani Broth (LB) with the appropriate antibiotic for the plasmid as designated by Open Biosystems. Cultures were incubated in a shaking platform incubator at 37°C for 16 hours. Following incubation, cultures were centrifuged to pellet cells and DNA was harvested from the cell pellets using the Promega plasmid mini-prep kit according to the manufacturer's guidelines. The resulting DNA was then used as a template for amplifying the gene of interest using PCR and primers specific for the DNA sequence of the gene.

Amplifying Genes of Interest from a Genomic DNA Library.

The remaining genes were amplified from genomic DNA. Primers were designed specific for the DNA sequence of each gene and ordered from Integrated DNA Technologies (IDT). Gene-specific primers used to amplify each gene are listed in Appendix 1. PCR reactions were carried out using the commercially available enzyme, Phusion, according to the manufacturer's guidelines. Cycling parameters were adjusted for length of constructs being amplified and the T_M of the associated primers. The subsequent PCR product was run on a 1% agarose gel at approximately 90 volts until bands were resolved on the gel. Bands of DNA that were approximately the size of the expected gene product were excised from the gel and purified with either the Promega gel

purification kit or a similar kit manufactured by Qiagen. For each method, the directions published by the manufacturer were followed.

Preparing the Gene of Interest for Incorporation into Gateway Vectors.

The resulting PCR product, once purified from the gel, was then used in a second round of PCR. The primers used in the initial PCR reactions previously carried out, in addition to being complementary to the gene of interest, contained extensions outside of the gene of interest that were complementary to regions necessary for incorporation into the Gateway vectors. This provided a universal sequence on all of the initial PCR products that could be further amplified with a set of universal primers designed for their complementarity to the Gateway vectors. The enzyme Phusion was again used according to the manufacturer's guidelines. Universal primer sequences are also listed in Appendix 1. The PCR product obtained was again run on a 1% agarose gel as before, excised from the gel, and purified.

Incorporating PCR-Amplified Genes into Gateway Vectors

Creating Donor Constructs:

Once the gel-purified PCR product was obtained, it was subsequently moved into the donor vector, pDONR201, for all genes of interest. The Gateway protocol for generating donor clones was followed with the following exception. The Gateway method recommends using 75-150 ng of the entry clone, 2 µL of the proprietary BP Clonase enzyme, and water to equal 8 µL. Using half-reaction volumes proved to be sufficient to generate positive clones, therefore these reactions were set up with 75 ng of

the entry clone, 1 µL of BP clonase and water to equal 4 µL. Reactions were left at room temperature overnight (generally 14-16 hours) after which time 1 µL of Proteinase K was added and the reactions incubated at 37°C for up to 1 hour. The BP reaction mixture was then used to transform competent *E. coli* cells.

Transformation of E. coli cells:

NEB5-α (DH5α) competent cells were originally purchased from New England Biolabs. Subsequent batches of competent cells were prepared according to the protocol in Short Protocols in Molecular Biology, Fifth Edition [148]. NEB5-alpha cells were transformed using 30 µL aliquots of cells with 2 µL of the BP reaction , plated on LB agar plates with kanamycin, and incubated at 37°C overnight. Colonies were subsequently screened by PCR to verify that the gene of interest was contained within the donor vector. Since the donor vector, pDONR201, has a kanamycin resistance gene, growth on the agar plates containing kanamycin was evidence that the donor vector was intact. Colonies that screened positive for the gene were grown in ZYP-0.8G media with kanamycin. A volume of 5 mL was used for each and cultures were grown at 37°C for 16 hours in a shaking incubator. Glycerol stocks were prepared from the culture and the remainder centrifuged and harvested using the Promega miniprep kit as previously described. The resulting miniprep DNA was then used for preparing final destination clones.

Generating Destination Constructs.

Destination constructs were created in a manner similar to that used to create the donor constructs described above The miniprep DNA from the positively screened donor constructs was combined with a destination vector containing either the coding sequence for a fluorescent fragment or whole-length fluorescent protein. These vectors contain an ampicillin resistance gene, therefore subsequent culturing steps were carried out with the presence of ampicillin in the culture media. In addition to the ampicillin resistance gene, the destination vectors also contain the *CCDB* gene, the product of which inhibits growth of most *E. coli* strains [149]. During recombination, the donor cassette recombines into the donor at the *att* sites, producing a product containing the chimeric gene of interest and fluorescent protein gene (or fluorescent fragment) in a plasmid with ampicillin resistance. A byproduct of the reaction is a plasmid containing the lethal *CCDB* gene.

To produce the final clones used for subsequent experiments, 75 ng of the previous donor construct DNA was combined with 75 ng of the destination vector and 1 µL of LR Clonase with addition to water to make 4 µL per reaction. Again the Gateway protocol was followed with the exception that reactions were carried out in half volumes as described. The reactions were left at room temperature overnight (approximately 16 hours) after which proteinase K was added and the reaction incubated at 37°C for 15 minutes to 1 hour. The resulting mixture was used to transform NEB5-alpha cells as previously described for the BP reactions and generation of donor plasmids, with the exception that ampicillin was used in place of the kanamycin that was used for

the donor plasmids. Again colonies were screened using PCR and primers specific for the genes of interest. Positive colonies were grown in ZYP-0.8G media with ampicillin overnight after which glycerol stocks were again made and the DNA harvested using the Promega plasmid DNA miniprep kit.

Screening Plasmids:

Numerous plasmids were created during the course of this project. A reliable and efficient screening method was needed. Colonies were picked and resuspended in 20 µL of sterile deionized water. From this 2 µL were used as the template DNA for PCR screening; the remainder was set aside to be used to grow in LB broth should PCR screening be positive. Colonies were screened with the 3' primer specific to the gene of interest and a universal 5' primer for either the donor vector or destination vector. PCR reactions were then set up with Taq polymerase and other components as recommended by New England Biolabs for use with Taq polymerase. Following amplification, the PCR products were separated on a 1% agarose gel until discrete bands were observed. These were compared to the expected size for each gene of interest. Positive clones were grown in LB with the appropriate antibiotic and a miniprep of the DNA was made. This was screened by PCR again to verify the presence of the gene in the miniprep DNA solution. Once a gene had been moved into a destination vector, the final round of PCR screening included three separate PCR reactions. The first utilized 5' and 3' primers specific to the gene of interest. The second utilized 5' and 3' primers specific for the vector and would produce a band slightly larger than the gene of interest. The third reaction utilized the 5' universal primer for the vector and the 3' primer specific for the

gene of interest. This produced a stair-step pattern that would, in most cases, prove to be positive screening once the construct was verified by sequencing as described below. An example of the PCR screening pattern can be seen below in figure 5.3

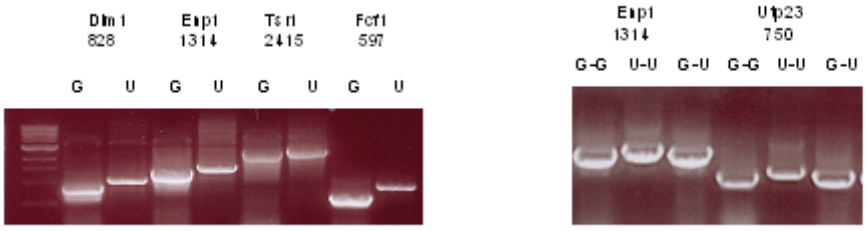


Figure 5.3: Examples of PCR method used to screen plasmids. *Left*: PCR products from reactions using 5' and 3' gene-specific primers, designated as "G" or 5' and 3' universal primers for BiFC destination vectors designated as "U". *Right*: Pattern observed when three PCR reactions were carried out for a gene of interest. "G-G" denotes 5' and 3' gene-specific primers; "U-U" denotes 5' and 3' universal primers for the BiFC destination vector; "G-U" denotes a 5' gene-specific primer and a 3' BiFC universal primer specific for the BiFC destination vector.

Once the destination clone had been verified by PCR, it was subsequently sent to the Nevada Genomics Center in Reno, Nevada, to verify the DNA sequence for each construct. Sequences were verified by two methods. Sequences obtained from the Nevada Genomics Center were uploaded to the NCBI BLAST site where they were searched against the BLAST database and examined for any mutations present in the sequence. Secondly, sequences obtained from the Nevada Genomics Center were compared against virtual cloning results using the computer program Gentle. Briefly, for each gene of interest, a virtual computer clone was created by using this software program. Donor and destination clones were created and the products of the Gateway reactions could be predicted and mapped. The sequence from the Nevada Genomics Center and the sequence obtained from virtual cloning utilizing Gentle were aligned against each other and compared for mutations. Clones containing the genes of interest and the 5' and 3' portions of the destination vector flanking the gene were then cultured and used for BiFC experiments.

Preparing DNA for BiFC Experiments:

Positively sequenced and verified clones were grown to produce a sufficient amount of DNA for subsequent experiments. Glycerol stock cultures from the positive clones were used to start 2 mL cultures in LB broth with ampicillin. All cultures were grown at 37°C in a shaking platform incubator. Following 8 hours of growth, cultures were added to 50 mL of ZYP-0.8G media containing ampicillin and incubated overnight. The resulting cultures were then harvested for their DNA using the Promega DNA Midiprep kit. The DNA concentration was determined for each and a working concentration of 100 ng/μL was created for each to be used in subsequent experiments.

Cell Culture Methods

Cos-7 Cells.

All experiments were done using Cos-7 cells, a green African monkey kidney cell, fibroblast-like cell line. Cells were initially obtained from the laboratory of Dr. Michele Hardy at Montana State University, Bozeman. This cell line was originally derived from the CV-1 cell line and is similar to Cos-1 cells which contain an integrated copy of the early region of the SV40 genome. Maintenance of the Cos-7 cells was carried out according to guidelines provided by the American Type Culture Collection (ATCC) [150]. Cos cells are suitable for transfection by vectors that require the expression of the SV40 T antigen. The presence of the SV40 promotor drives gene expression in constructs prepared for experimental use as will be discussed below.

Cell Culturing Conditions.

Specific cell culture methods include the following. Cells were grown at 37°C with 5% CO_2 and were maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum and 1% antibiotic/antimycotic mixture containing penicillin, streptomycin, and amphotericin B. Cells were passaged approximately every 3 days and were split at a ratio of approximately 1:4 or 1:5 to maintain cells in the exponential growth phase. Approximately 24 hours prior to transfection, cells were trypsinized and resuspended in fresh medium to reach a concentration of 1×10^5 cells/mL. Tissue culture-treated 12-well plates were used for most experiments. A small number of experiments were carried out in larger, individual glass-bottom dishes to permit examination using higher magnification and oil immersion lenses on the inverted Nikon microscope. Cells were plated with 1 mL per well in the 12-well plates and 2 mL in the glass-bottom dishes to reach approximately 80% confluence at the time of transfection.

Transfection Protocol.

Cells were transfected using the transfection reagent *Trans*IT®-Cos supplied by Mirus. This transfection reagent has been specifically designed to work with Cos cell lines. Guidelines supplied by Mirus [151] were followed with the exception that less DNA per transfection reaction was used. The reduced amount of DNA was used based on experiments carried out previously in which higher concentrations of DNA were not advisable for BiFC experiments due to resulting high false positive results and detection of nonspecific interactions [140]. DNA was transfected into Cos-7 cells at a

concentration of 500 ng per well of a standard 12-well plate. For co-transfections, 500 ng of each construct was added to the well. This resulted in a total amount of DNA that was still within the guidelines prescribed by Mirus.

Microscopy.

At 24 hours post transfection, cells were examined microscopically for the presence of fluorescence signal in the individual wells of the 12-well plates using a Nikon inverted microscope. The camera set-up utilizes a black-and-white camera for improved resolution over some color cameras. Prior to imaging, culture medium was removed from the cells and replaced with PBS to reduce defraction caused by the culture medium that was observed when it remained in the wells during imaging. Digital photographs were taken of the cells using a camera attached to the microscope and the software program Metamorph at 24 hours post transfection. Plates were returned to the incubator and imaged a second time at 36-48 hours post transfection. Images were artificially colored using the Metamorph software, Adobe Photoshop, and/or Irfanview, a freeware image processing software package available on-line. Images are representative of cells in at least two different wells for each of the combinations shown.

RESULTS

Initially it was attempted to re-create experiments performed by C-D Hu and T. Kerppola [141] by recombining Fos-VC155 with Jun-VN173, however these experiments were unsuccessful. Other complementing pairs using YFP constructs, however were successful as were transfections with both full-length Venus and Cerulean constructs on their own without the addition of other DNA sequences, as was done in the experimental constructs. The well-documented interaction between Nob1p and Rrp20p was ultimately used as an internal positive control. Rrp20p, which is also known as Dim2p, was originally labeled Pno1p for "partner of Nob1". Nob1p N-terminal or C-terminal Venus constructs combined with the complementing partner in an Rrp20p construct exhibited a strong fluorescent signal. Similarly, two known interacting partners involved in the early processing steps of ribosome assembly, Utp23p and Fcf1p, demonstrated a strong fluorescent signal in the nucleolus.

As expected, constructs containing a full-length fluorescent protein exhibited greater signal than that produced by complementation pairs reconstituting the intact fluorescent protein. Full-length Venus constructs for each protein and its cellular localization are listed in table 6.1. Several proteins were also transfected in full-length Cerulean or full-length mRFP constructs. As expected, the same localization pattern was observed as was seen with the Venus constructs.

Table 6.1: Localization signal of full-length Venus constructs

Protein	**Localization**
Dim1p	Primarily nucleolar with some cytoplasmic signal
Enp1p	Primarily nucleolar with some cytoplasmic signal
Hrr25p	Cytoplasmic and nuclear, diffusely and evenly throughout
Krr1p	Primarily nucleolar with weak cytoplasmic signal
Nob1p	Cytoplasmic signal only with no nuclear signal
Rio2p	Cytoplasmic signal only with no nuclear signal
Rrp20p	Nuclear and nucleolar, some cytoplasmic signal
Tsr1p	Nucleolar and nuclear
Gar1p	Nuclear and nucleolar, no cytoplasmic signal
Fcf1p	Nucleolar
Fcf2p	Nuclear and nucleolar
Utp23p	Nucleolar
Ltv1p	Nucleolar and speckling in cytoplasm
Enp2p	Nuclear and nucleolar
Emg1p	Nucleolar
Utp4p	Nucleolar

Numerous constructs were created during the course of this project for the purpose of determining binary protein-protein interactions using the BiFC assay. Table 6.2 summarizes the constructs that were made.

Table 6.2: Constructs created with the Gateway system

Gene	Venus	VC-155	VN-173	Cerulean	CC-155	CC-173	mRFP
Dbp4							
Dim1	X	X	X	X			
Emg1	X	X	X	X	X	X	
Enp1	X	X	X	X	X	X	X
Enp2	X	X	X	X	X	X	
Fcf1	X	X	X	X	X	X	
Fcf2	X	X	X				
Gar1	X	X	X				
Hrr25	X		X				
Krr1	X	X		X	X	X	
Ltv1	X	X	X	X	X	X	X
Nob1	X	X	X	X	X	X	X
Rio2	X	X	X	X	X	X	X
Rrp20	X	X	X	X	X	X	X
Tsr1	X	X	X	X	X	X	X
Utp4	X	X	X				
Utp23	X	X	X	X	X	X	

In addition to the table above, several of the genes were successfully moved into mRFP vectors. Two different C-terminal halves and two N-terminal halves were created since it was unclear which combination would optimally recombine to produce the fluorescent protein, mRFP. The vectors used for the red fluorescent protein included a full-length version, denoted mRFP, and the fragments RC-155, RN-156, RC-169, and RN-170. The fluorescent signal for the full-length proteins was not as strong as had been observed with either Cerulean constructs or Venus construct. Unfortunately, none of the complementary pairs combined to reform a detectable fluorescent protein in the combinations tested with the setup available at this time. Genes moved into the five RFP vectors were *RIO2*, *RRP20*, *ENP1*, and *NOB1*, all of which had been successfully cloned

into Venus vectors and emitted a fluorescent signal when combined in the VC-155 and VN-173 constructs.

Binary interactions were detected for numerous protein-protein pairs tested using the VC-155 and VN-173 constructs. Table 6.3 lists pairs for which an interaction was not detected. Beginning with figure 6.1, full-length Venus constructs and interacting complementation pairs used in the BiFC assay are shown. Figure 6.1 shows the cellular localization of full-length Venus constructs corresponding to proteins involved with early cleavage steps at sites A_0 through A_2. Figure 6.2 shows representative pictures of the signal observed for proteins involved at these early cleavage steps when combined in the BiFC assay. Figure 6.3 shows cellular localization of full-length constructs corresponding to proteins involved in cleavage and processing at site D. Figures 6.4 through 6.10 show representative photographs obtained in the BiFC assay for proteins involved at processing at site D. Finally, figure 6.11 is an grid showing interactions among the proteins involved at site D and indicates the cellular location where the interaction takes place.

Table 6.3: Non-interacting pairs tested using the BiFC assay

VC-155 Construct	**VN-173 Construct**
Enp1-VC155	Hrr25-VN173
Rio2-VC155	Hrr25-VN173
Dim1-VC155	Hrr25-VN173
Tsr1-VC155	Dim1-VN173
Dim1-VC155	Tsr1-VN173
Rrp20-VC155	Hrr25-VN173
Tsr1-VC155	Hrr25-VN173
Ltv1-VC155	Hrr25-VN173
Ltv1-VC155	Dim1-VN173
Dim1-VC155	Ltv1-VN173

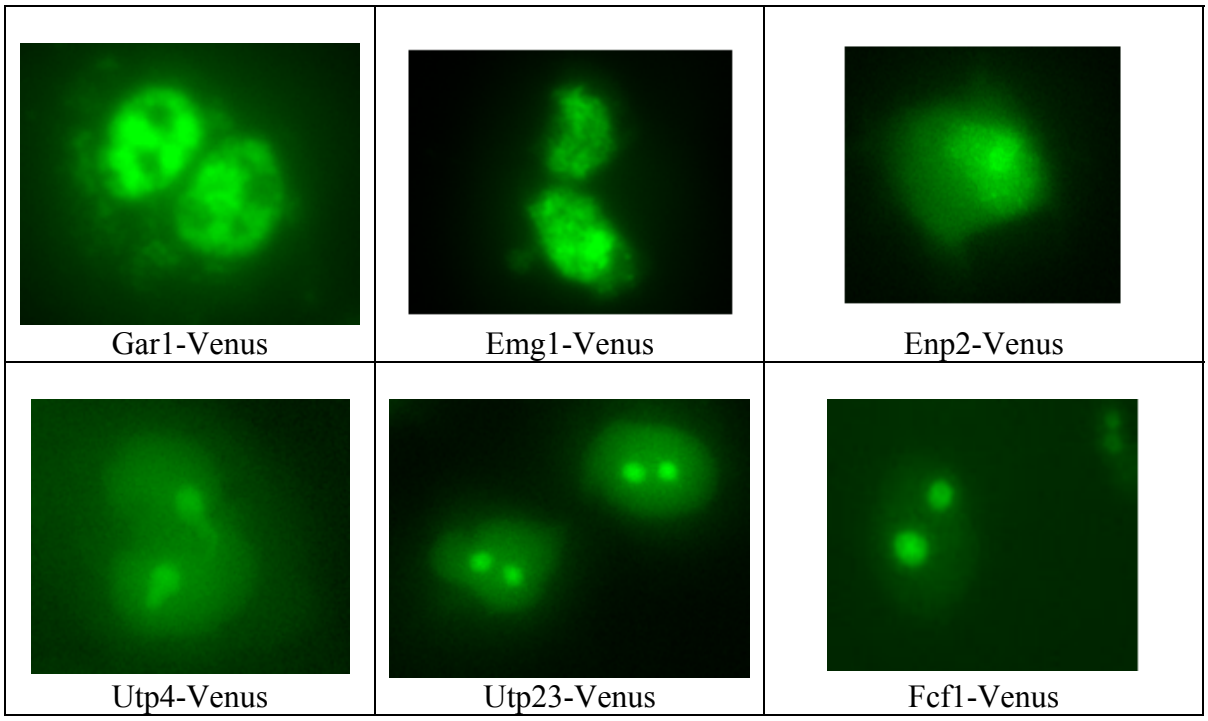


Figure 6.1: Full-length Venus constructs showing localization of individual accessory proteins involved in early cleavage steps at A_0, A_1, and A_2.

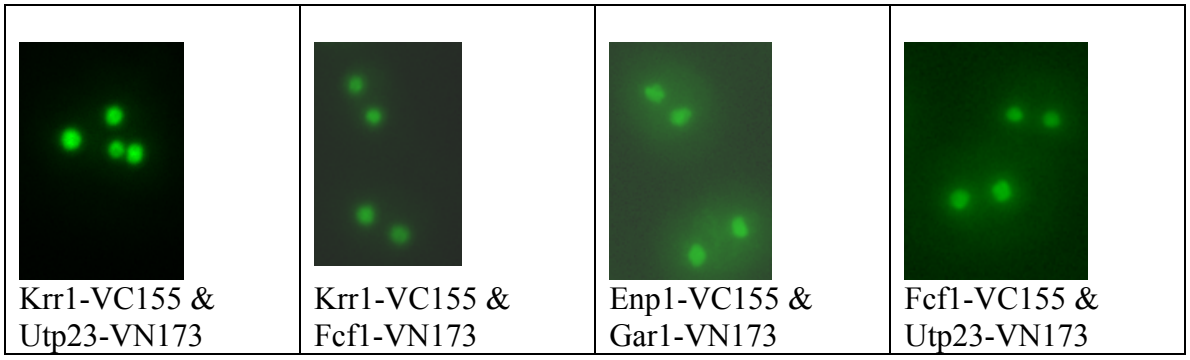


Figure 6.2: Complementation of proteins involved in early cleavage steps at A_0, A_1, and A_2. All pairs tested had a similar nucleolar localization pattern.

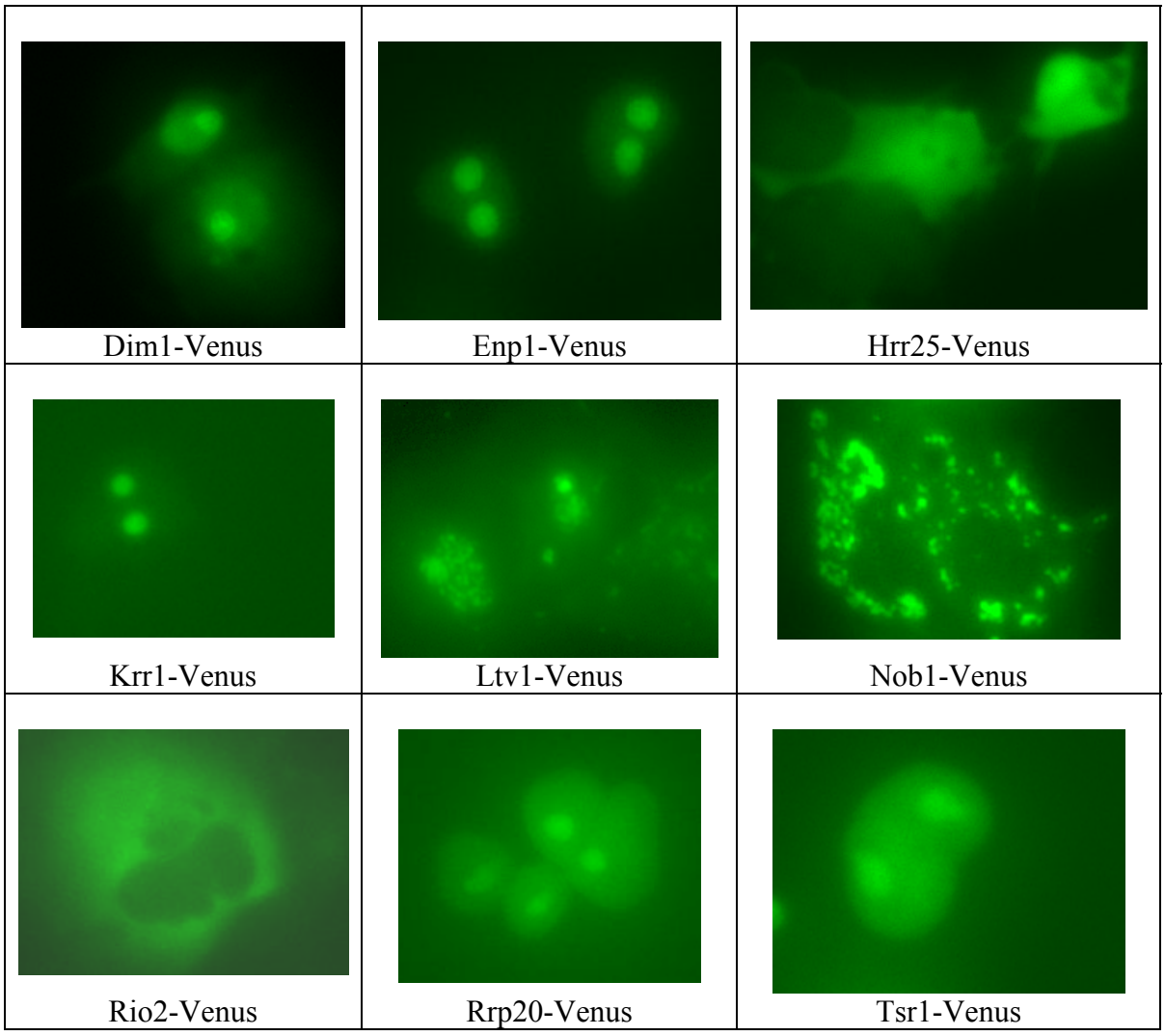
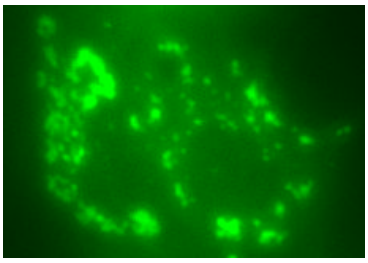
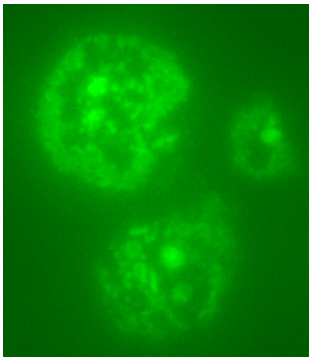


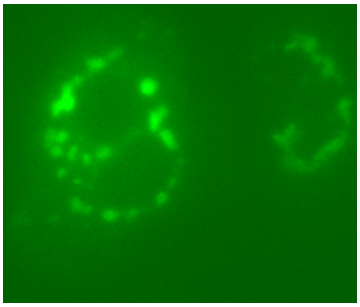
Figure 6.3: Full-length Venus constructs showing localization of individual accessory proteins involved in cleavage at site D



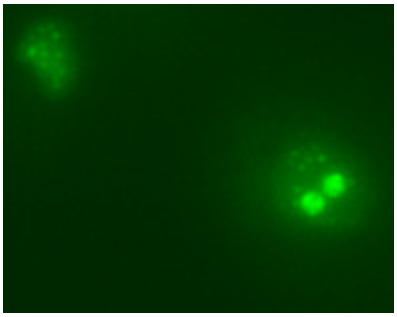
Localization of Nob1



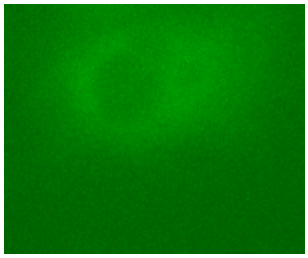
Enp1-VC155 &
Nob1-VN173



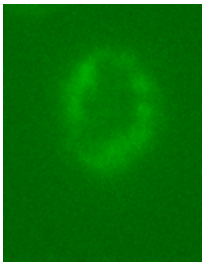
Nob1-VC155 &
Rio2-VN173



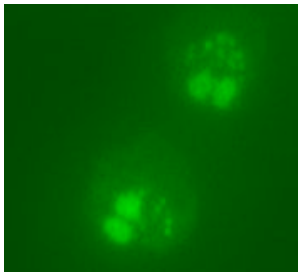
Nob-VN173 &
Tsr1-VC155



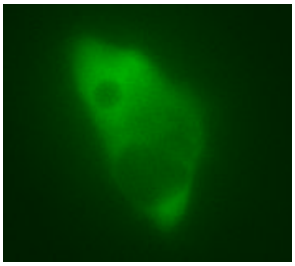
Nob1-VC155 &
Dim1-VN173



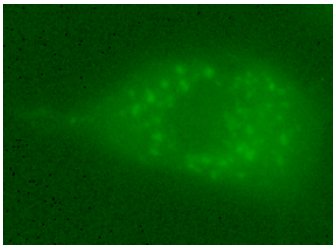
Nob1-VC155 &
Dim1-VN173



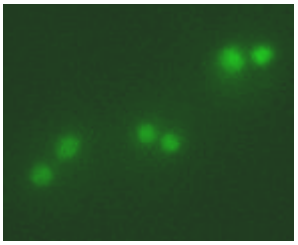
Nob1-VN173 &
Rrp20-VC155



Nob1-VC155 &
Hrr25-VN173

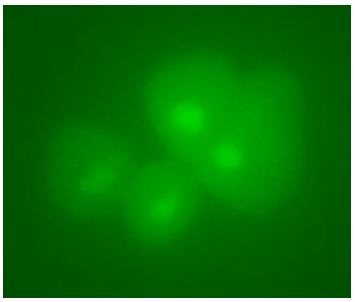


Nob1-VN173 &
Ltv1-VC155

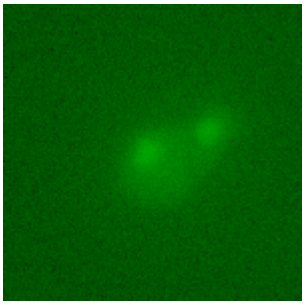


Nob1-VC155 &
Krr1-VN173

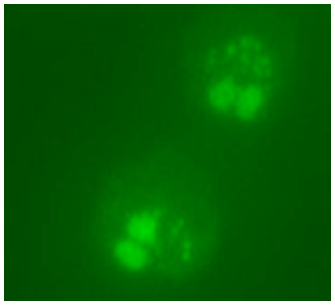
Figure 6.4: BiFC results involving Nob1p



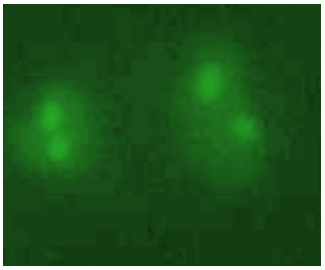
Localization of Rrp20



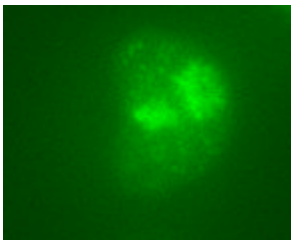
Dim1-VC155 &
Rrp20-VN173



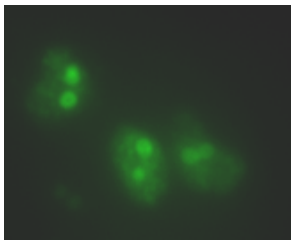
Rrp20-VC155 &
Nob1-VN173



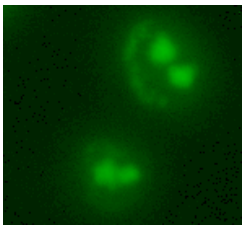
Rrp20-VC155 &
Enp1-VN173



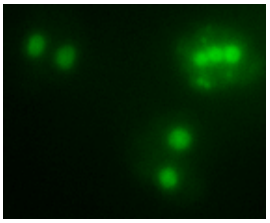
Rio2-VC155 &
Rrp20-VN173



Rrp20-VC155 &
Tsr1-VN173

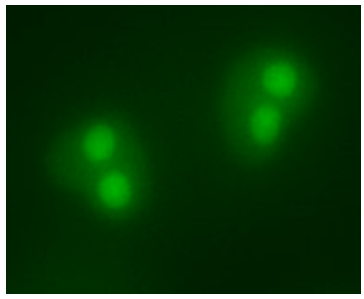


Ltv1-VC155 &
Rrp20-VN173

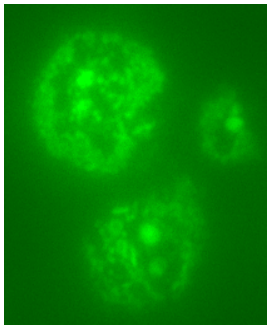


Krr1-VC155 &
Rrp20-VN173

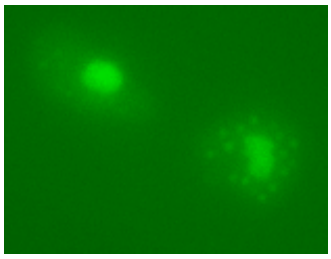
Figure 6.5: BiFC results involving Rrp20p (Dim2p/Pno1p)



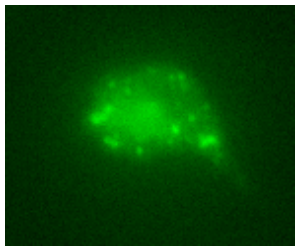
Localization of Enp1-Venus



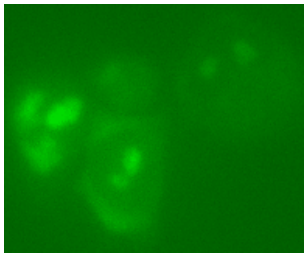
Enp1-VC155 &
Nob1-VN173



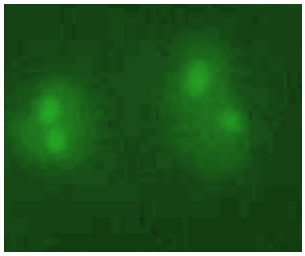
Enp1-VN173 &
Tsr1-VC155



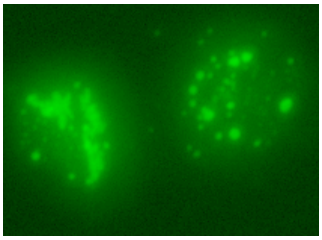
Enp1-VC155 &
Rio2-VN173



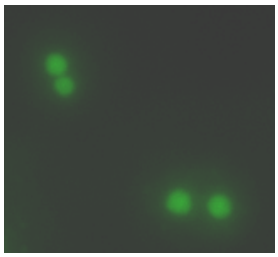
Enp1-VC155 &
Dim1-VN173



Enp1-VC155 &
Rrp20-VN173



Ltv1-VC155 &
Enp1-VN173



Enp1-VN173&
Krr1-VC155

Figure 6.6: BiFC results involving Enp1p

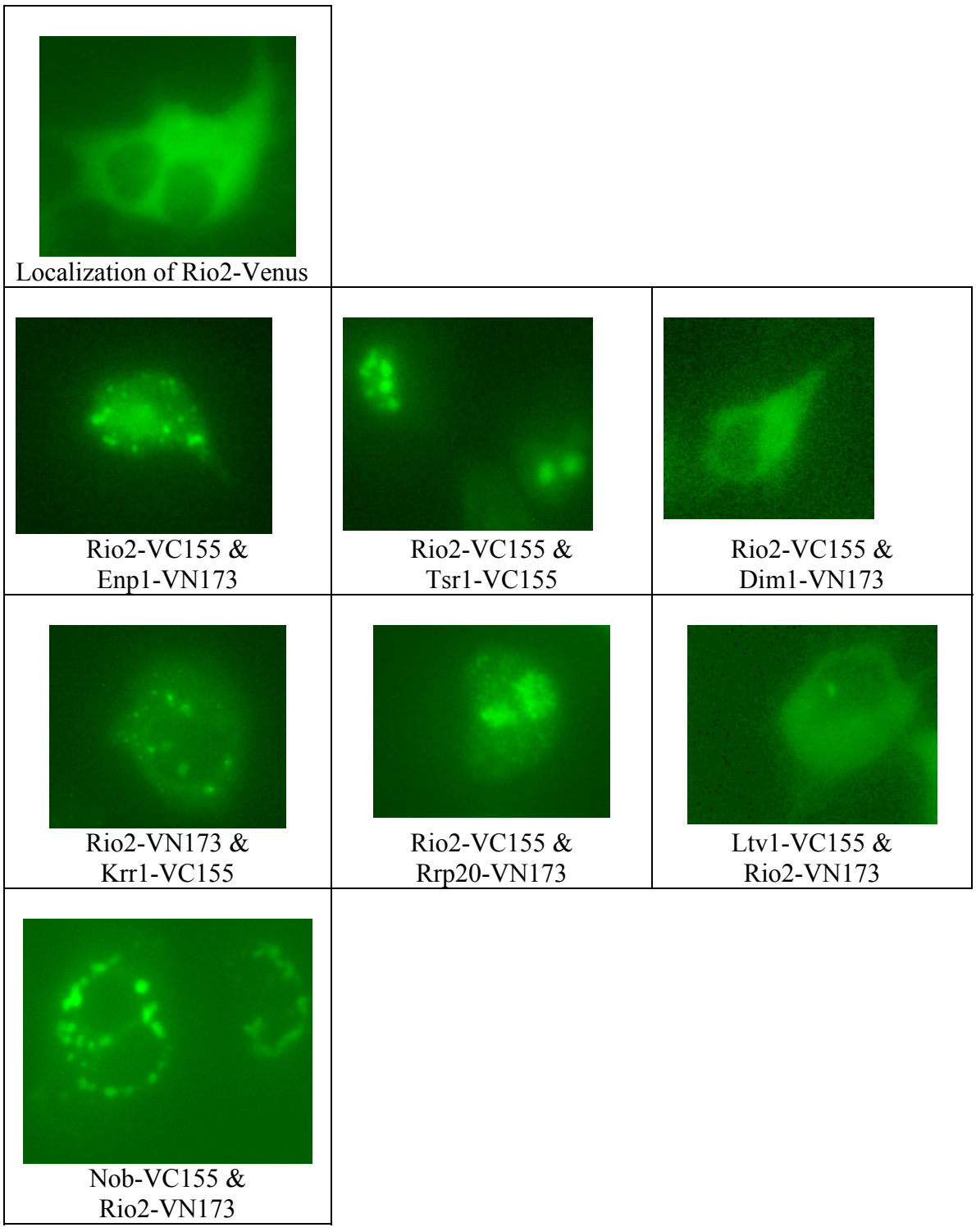
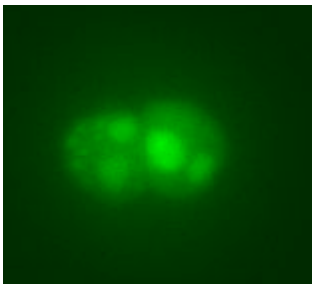
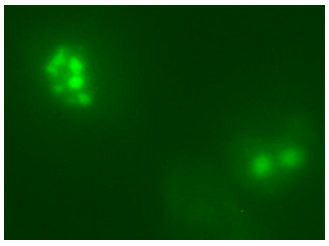


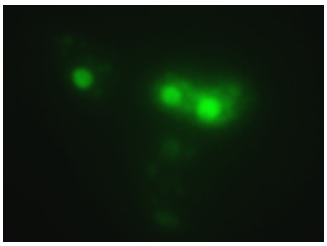
Figure 6.7: BiFC results involving Rio2p



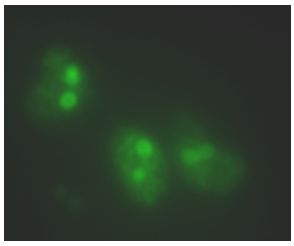
Localization of Tsr1-Venus



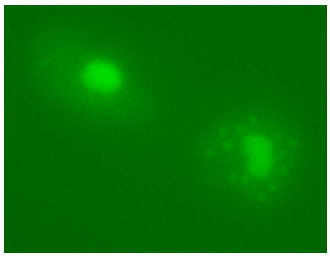
Rio2-VC155 &
Tsr1-VN173



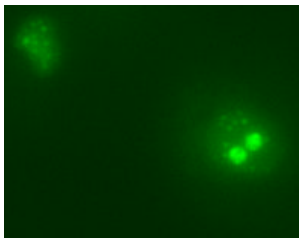
Tsr1-VC155 &
Nob1-VN173



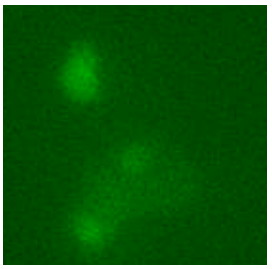
Rrp20-VC155 &
Tsr1-VN173



Tsr1-VC155 &
Enp1-VN173



Tsr1-VN173 &
Krr1-VC155



Tsr1-VC155 &
Dim1-VN173

Figure 6.8: BiFC results involving Tsr1p

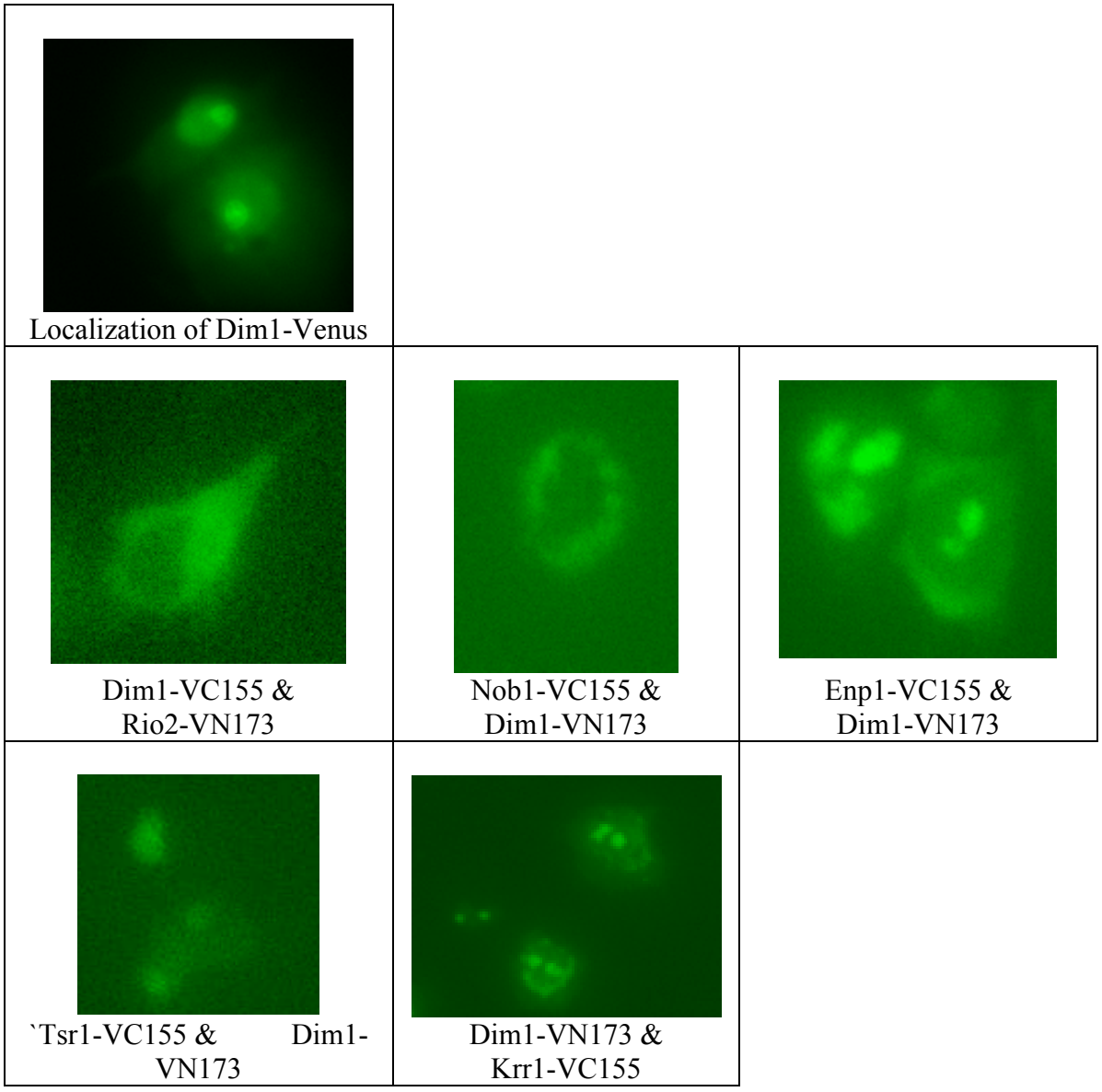


Figure 6.9: BiFC results involving Dim1p

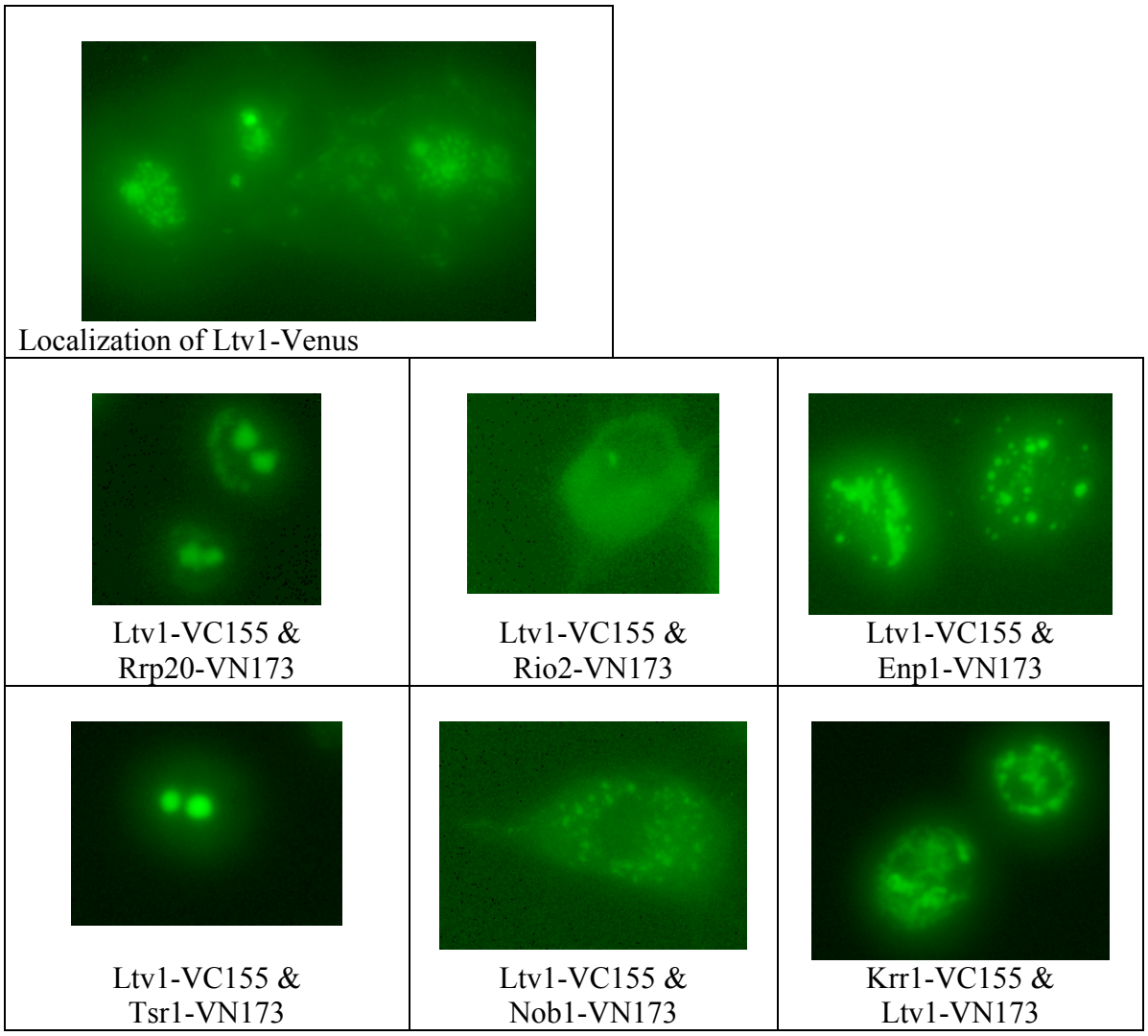


Figure 6.10: BiFC results involving Ltv1p

	Dim1	Enp1	Hrr25	Krr1	Ltv1	Nob1	Rrp20	Rio2	Tsr1
Dim1									
Enp1	**O/N**								
Hrr25	**none**	**none**							
Krr1	**R**	**O**	**none**						
Ltv1	**none**	**N/C**	**none**	**N**					
Nob1	**C**	**O/N/C**	**C**	**O**	**C**				
Rrp20	**O**	**O**	**none**	**O**	**O**	**O/N**			
Rio2	**C**	**N**	**none**	**N**	**C**	**C**	**O/N**		
Tsr1	**none**	**O/N**	**none**	**O**	**N/C**	**O/N**	**O**	**O/N**	

Figure 6.11: Interaction grid for proteins tested. Letters indicate the primary localization of the interaction. O=Nucleolus; N=Nucleus; C=Cytoplasm; none= no reaction.

While most of the experiments carried out during the course of this study were done with full-length Venus constructs and the complementing halves, VC-155 and VN-173, several constructs were also made and used in Cerulean and red full-length constructs as well as a few Cerulean complementation constructs using CC-155 and CN-173. As expected, localization for full-length constructs was the same for Venus, Cerulean, and RFP constructs (Figure 6.12). Similarly, complementation experiments using CC-155 and CN-173 fused to interacting proteins yielded the same results as was

observed in Venus constructs. A comparison of the three fluorescent proteins used for the same protein is presented below in figure 6.13.

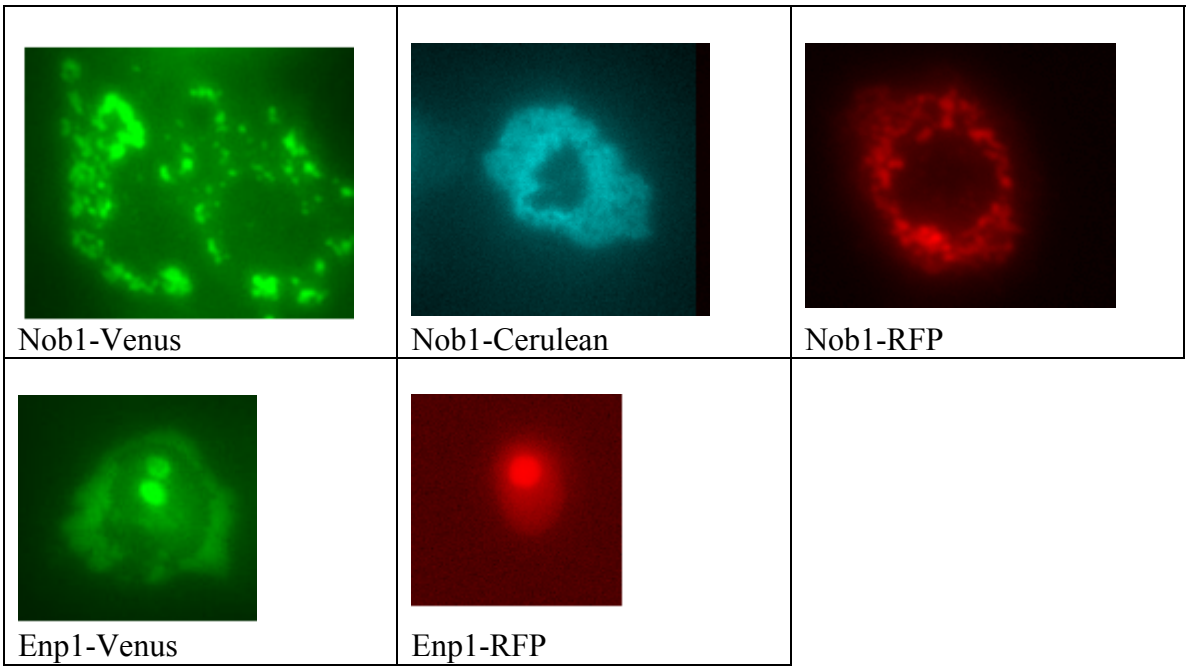
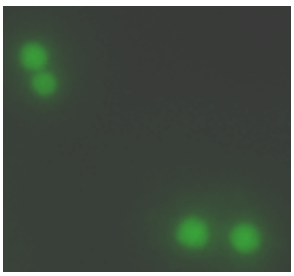
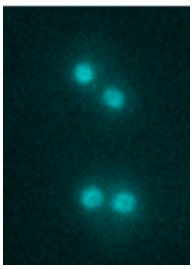


Figure 6.12: Localization of full-length Venus, Cerulean and RFP constructs. Localization occurs in the same cellular compartment when the gene of interest is fused to Venus, Cerulean, or RFP as in the case of Nob1p with fluorescent signal detected within the cytoplasm (top row) or with Venus and RFP in the case of Enp1p where signal is predominantly nucleolar (bottom row).



Enp1-VN173 & Krr1-VC155



Enp1-CN173 & Krr1-CC155

Figure 6.13: Comparison of localization of Venus and Cerulean constructs in BiFC assay. Localization occurs in the nucleolus for complementation pairs using Venus or Cerulean fragments for the interacting pair Enp1p and Krr1p.

Several combinations were tested using one pair of complementing constructs, such as Enp1-VC155 and Rrp20-VN173, along with a full-length red fluorescent construct, such as Nob1-RFP. For each set tested in this manner, protein "A" was shown to interact with protein "B", protein "B" interacted with protein "C", and protein "A" and protein "C" also interacted. For these combinations, three outcomes were observed. In most cases, the localization for the construct containing the full-length red fluorescent protein was the same as the localization for the BiFC pair composed of two proteins each fused to either VC-155 or VN-173 (Figure 6.14). In one case, the localization of the BiFC complementing pair was clearly different from the full-length RFP construct (Figure 6.15). Finally, for one set of three proteins, the localization pattern differed depending on which proteins made up the complementing pair and which was added in the full-length RFP construct (Figure 6.16). No bleed-through of signal from the green channel into the red channel was observed for any of the combinations tested. Essentially

no bleed-through of signal from the red channel into the green channel was observed. In some cases where a very strong red signal was present, an extremely faint, barely detectable signal was observed as bleed-through in the green channel. This was only noticeable in the case of full-length red constructs transfected on their own without a green fluorescent construct transfected in the same well. When co-transfections were carried out with both green and red fluorescent constructs, true signal was observed in the green channel that was identical to that observed when green constructs alone were transfected. In the presence of a true green signal, no bleed-through from the red channel was detected.

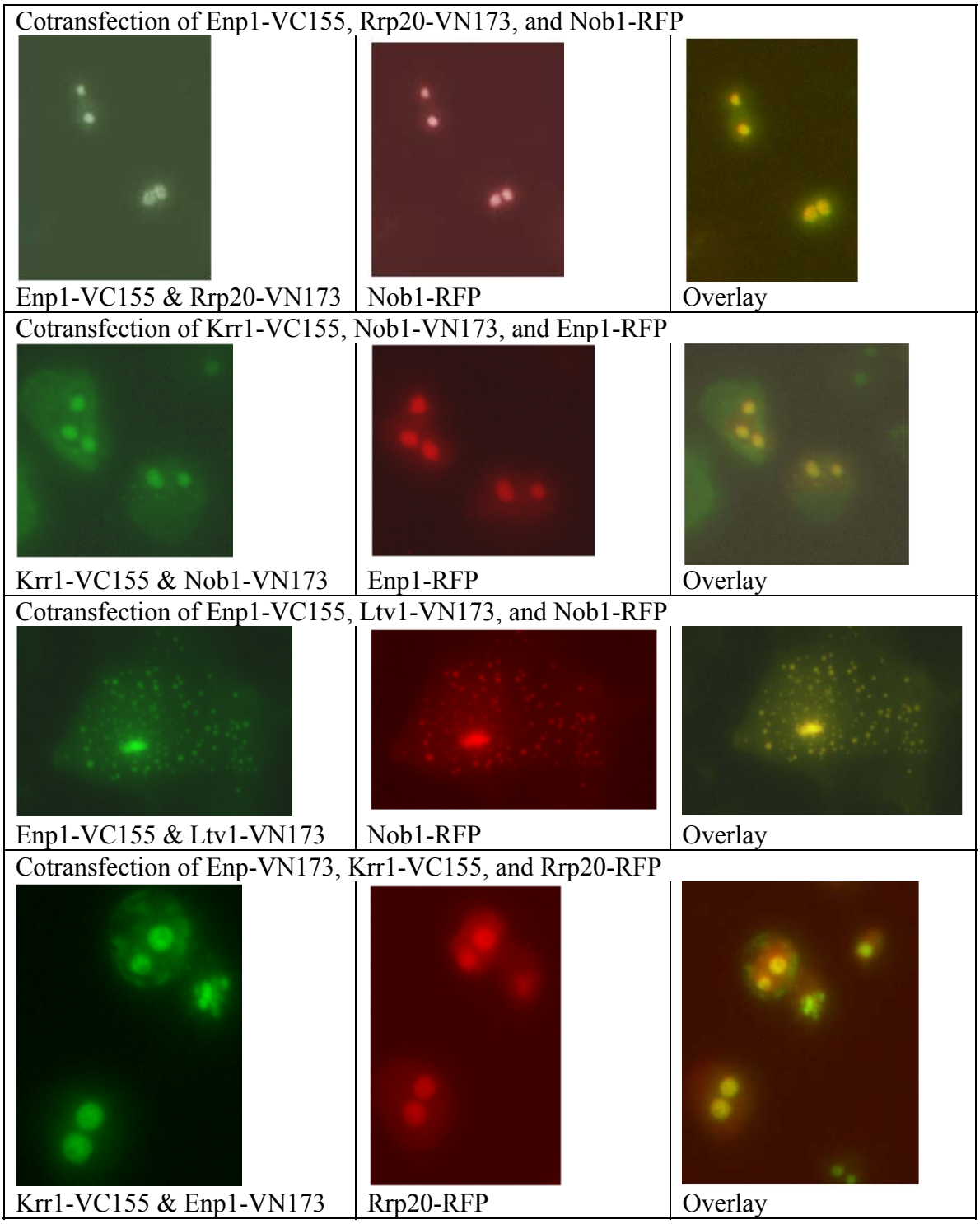


Figure 6.14: Evidence supporting trimeric complex formation

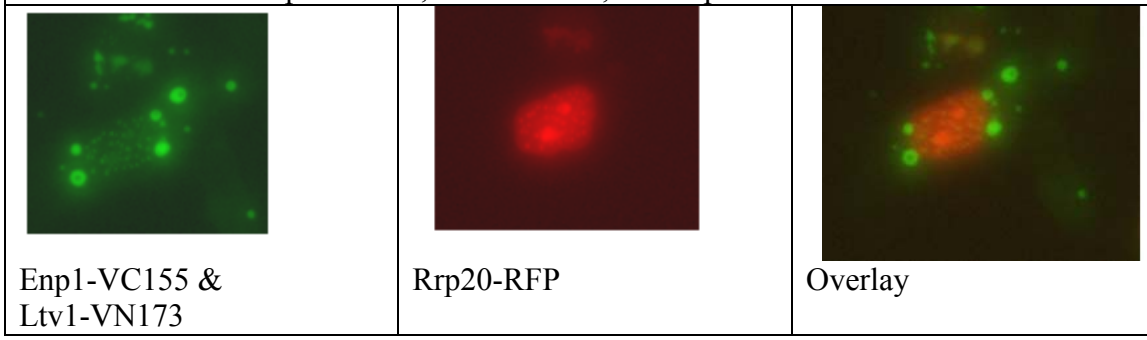


Figure 6.15: Evidence against trimeric complex formation

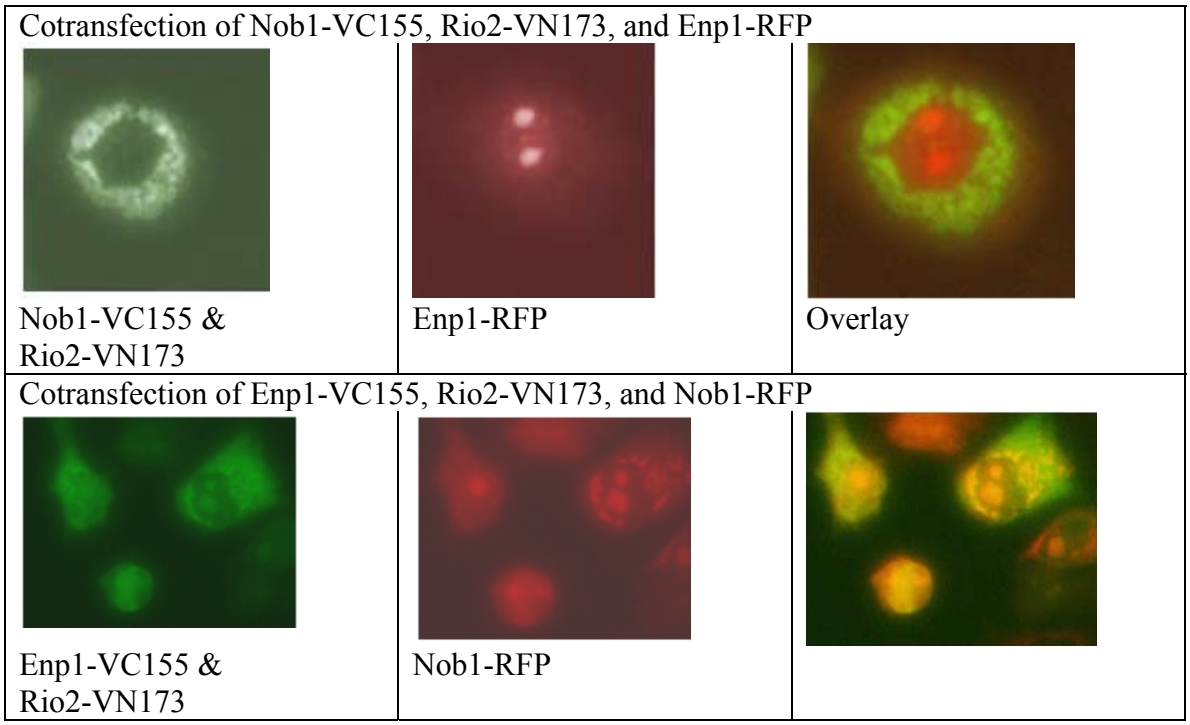


Figure 6.16: Conflicting evidence for trimeric complex formation among Nob1p, Rio2p, and Enp1p when co-transfected into Cos-7 cells

SUMMARY AND CONCLUSIONS

Ribosome biogenesis is a complicated, ordered process that, in eukaryotes, requires four different rRNA molecules, approximately 80 ribosomal proteins, and at least 200 *trans*-acting factors. Many of the participants in this process are known, however the assembly and disassembly process of the associated factors has not been determined. Complexes join and exit the developing pre-ribosome in an ordered fashion with each step catalyzed or influenced by the previous step, as is done in an assembly line. To better understand the entire process it is necessary to understand the smaller, individual steps and the complexes that are formed and dismantled as the ribosome is being developed.

By examining binary interactions that occur within the larger overall process, it is possible to better understand the complex network of interactions that takes place.

This project has sought to provide evidence for binary interactions among accessory proteins involved in synthesis of the small ribosomal subunit in eukaryotes, specifically using human genes, to further our understanding of this process in mammalian cells. While numerous high-throughput studies have been done which demonstrate the participants involved, these studies do not permit visualization in living cells and are further complicated by difficulties with both false positive and false negative results [56, 119, 127, 130, 152]. Studies carried out with *in vitro* systems need to be repeated *in vivo* to confirm that proteins perform in a similar manner in both

systems since removing a protein from its normal physiological setting can alter its function.

With the above considerations in mind, an *in vivo* system to study binary interactions was developed. This system utilizes the Gateway cloning method where the accessory ribosomal protein genes could be efficiently combined into plasmids. This system is based upon bacteriophage lambda in *E. coli.* Once the gene of interest has been amplified with flanking regions complementary to the *att* sites used by the Gateway system, it is easily recombined into donor vectors which can then be used to shuttle the gene into numerous destination vectors that can be used experimentally. These plasmids were designed to contain a chimeric gene containing the gene of interest and a fluorescent protein fragment. For each gene, multiple chimeric constructs were created. The fluorescent component of these constructs contained either the gene encoding the full-length fluorescent protein which was used to determine localization of the protein when it was expressed in mammalian cells on its own or a fluorescent fragment. Two different fragments were created for each fluorescent protein, one containing the C-terminus of the gene and the other containing the N-terminus. These halves have been previously shown to reconstitute the fluorescent protein when attached to interacting proteins. With this system, two potentially interacting accessory proteins could be co-transfected into mammalian cells where the proteins encoded by the genes were subsequently produced and, if an interaction occurred between the two proteins, the fluorescent signal of the reconstituted fluorescent protein could be observed using fluorescent microscopy. This

system permitted not only determination of binary interacting pairs, but also demonstrated the location within the cell where this interaction takes place.

An additional advantage of this method is the efficiency and speed with which it can be performed. Once DNA corresponding to the gene of interest has been obtained, two rounds of PCR are required to amplify the DNA and incorporate the attachment sites necessary for shuttling the gene into the various vectors. Following this, two sequential reactions, first the BP reaction followed by the LR reaction, are performed to move the gene of interest into a donor vector and then into the final destination vector. Broth cultures are then set up to produce an ample supply of DNA which is harvested and purified from the cells. This DNA is then transiently transfected into cells, incubated, and 24 hours later, observed using a fluorescent microscope. From start to finish, this can be completed within a week or less.

The most error-prone aspect of this assay has been amplifying the genes of interest from a cDNA library to generate the DNA to be used in the cloning steps. For numerous genes it was not possible to generate the desired piece of DNA. Rrp5p and Rrp6p are two of the proteins that fall into this category. Rrp5p is contains a coding sequence greater than 5 kb, so it was not surprising that PCR amplification proved difficult. Others, such as Rrp6p, however, are within the size range successfully amplified. Redesigning primers for these genes or attempting to amplify them from a new cDNA source could potentially allow them to be amplified and used in this study.

An additional problem encountered was that for several genes a PCR product was generated of the appropriate size that was expected based upon the sizes of yeast

homologues. However, upon sequencing these constructs it was discovered that instead of the gene of interest, either a similarly sized gene had been amplified or during recombination steps, a bacterial gene had been recombined into the destination vector rather than the gene of interest. An example of this was Dbp8p which, despite numerous attempts, failed to produce a construct containing the *DBP8* gene.

Image processing and analysis proved to be the most time-consuming step of this assay. Typically two or three 12-well plates were set up for each experiment. Analyzing cells and images in 24 to 36 wells was the most labor-intensive aspect of the assay. This could be improved by utilizing the CellProfiler software**.** This software package, which is freely available on-line, can be used to process and analyze images. The software is capable of examining factors such as intensity of fluorescent signal in the images and identifying cellular compartments, such as the nucleus in an automated fashion. Images are processed in what the developers refer to as a pipeline with modules added for various processing steps [153-155]. This study was qualitative in nature, observing whether or not interactions took place and, if they did, identifying the location in the cell where the interaction occurred. However, a quantitative approach could be undertaken by comparing fluorescent intensities for images photographed under the same exposure settings.

Transfection efficiency was approximately 50% to 60% for most experiments. However, variations in transfection efficiencies were observed. In general, this was due to either overly confluent cells prior to transfection, or a high passage number for the Cos-7 cells. These factors were remedied. Maintaining a relatively low passage number

not to exceed 50 is recommended for further experiments with this cell line. Additionally, cells should not be greater than 80% confluent at the time of transfection, as this has been shown to inhibit transfection efficiency.

It is difficult to assign an order to the various proteins involved in the assembly process of the small subunit of the eukaryotic ribosome, however some hypotheses can be made in this regard. Dim1p likely acts before Rio2p during the sequence of events leading to formation of mature ribosomal subunits since in the absence of Rio2p, dimethylated 20S pre-rRNA accumulates in the cytoplasm [121]. Accordingly, multiple images showed Dim1p to have a nucleolar localization signal as well as a weaker cytoplasmic signal whereas Rio2p, when expressed alone, was observed in the cytoplasm and was clearly absent from the nucleus.

Hrr25p has been reported to phosphorylate a complex composed of Enp1p, Ltv1p, and Rps3p, however during this study Hrr25p was not shown to interact with either Enp1p or Ltv1p. During the course of creating the various constructs used in these experiments, difficulty was encountered with cloning Hrr25p. Of note, only the full-length Venus construct and the N-terminal fragment, Hrr25-VN173, were created. When setting up transfection experiments to examine potential binary interactions, two different combinations were set up for each pair. For example, Rio2-VC155 was combined with Tsr1-VN173 in one well and in the adjacent well, Rio2-VN173 was combined with Tsr1-VC155. For most of the combinations tested, the localization and strength of the fluorescent signal was comparable between the two pairs. For a few pairs, however, one combination produced a stronger fluorescent signal than the other. This occurred with

the Rio2/Tsr1 pair just described. The signal was considerably stronger with the Rio2-VC155 and Tsr1-VN173 combination. The other pair showed markedly decreased signal in comparison and was barely detectable.

Detection of a fluorescent signal requires that the two proteins interact and bring the halves of the fluorescent molecule into contact with each other to reconstitute the fluorescent protein. It is possible in some cases that the two proteins interact in such a way that the fluorescent fragments are pointing in opposite directions and thus cannot reform the fluorescent protein. This scenario could explain the unexpected results obtained with the Hrr25p constructs. A second possible explanation is that phosphorylation of the Enp1p/Ltv1p/Rps3p complex by Hrr25p follows a prior modification event that was inhibited by the excess of Enp1p and Hrr25p in the cell. Perhaps the interaction of Enp1p and Ltv1p causes a conformation change on Enp1p which is then able to be phosphorylated by Hrr25p. The great excess of Enp1p in the cell would not be adequately bound by Ltv1p and as a result Hrr25p would not bind. This is, of course, speculation but it could explain this finding. To further examine the interactions between Enp1p and Hrr25p, however, the next step would be to repeat cloning attempts to produce the Hrr25-VC155 construct and repeat the experiment with the Enp1-VN173/Hrr25-VC155 pair to elucidate whether the issue observed with binding of the Rio2p/Tsr1p pair occurs with this pair as well.

Ltv1p is has been shown to be involved with transporting small subunit pre-rRNA from the nucleus to the cytoplasm. Accordingly, in this study Ltv1p interacts with proteins involved in late maturation steps of the small ribosomal subunit, including

Enp1p, Rio2p, and Tsr1p. Ltv1p and Enp1p were shown to interact only in the cytoplasm. The signal observed from the interaction of Ltv1p and either Tsr1p or Rio2p was observed in the both the cytoplasm and the nucleus. This suggests that the interaction with Ltv1p and Enp1p occurs prior to the interaction with Ltv1p and either Tsr1p or Rio2p. Tsr1p could be involved with export of the pre-rRNA to the cytoplasm, explaining its localization in the nucleolus and ultimate presence in the cytoplasm. In this scenario, Tsr1p would be exported to the cytoplasm and then reimported back into the nucleus. It is possible that binding of Tsr1p is required prior to export to the cytoplasm.

The co-transfected trios of constructs localizing to precisely the same location within the cell does not prove that an interaction is taking place between the BiFC pair of constructs and the full-length red fluorescent construct, however it does provide evidence that such an interaction is likely and possible since all three proteins are in the same place in the same cells at the same time. Figure 6.14 shows localization of the complementing constructs in the green channel (first column) and the full length construct in the red channel (second column) with an overlay of the two images in the third column. For one combination, Enp1p, Krr1p, and Rrp20p, while the localization of the complementing BiFC partners is the same as that of the full-length RFP construct, another possibility must be considered. Krr1p and Rrp20p are both RNA binding proteins. While it is possible that they exist in a complex concurrently with Enp1p, it is also possible that both Krr1p and Rrp20p are binding the same segment of rRNA that is being processed and

Enp1p is interacting with them. Further experiments would be necessary to determine if interactions among the trios of proteins discussed are truly taking place.

Finally, some evidence for and against trimeric complexes was presented. In one instance, co-transfection of three constructs provided evidence against complex formation for the three proteins involved. The BiFC complementing pair composed of Enp1-VC155 and Ltv1-VN173 exists in discrete compartments within the cytoplasm of the cell. However, Rrp20-RFP is shown to be localized to the nucleus and concentrated in the nucleolus. Based on the localization patterns of Rrp20p, Enp1p, and Ltv1p, it does not appear that these three proteins are present in the same complex at the same time, even though they interact on a pair-wise basis for each possible pair for these three proteins. Localization of these constructs are presented in figure 6.15**.** Finally, one set of three proteins co-transfected gives conflicting results. When Nob1-VC155 and Rio2-VN173 are co-transfected, localization within the cell occurs at the nuclear rim. Enp1-RFP co-transfected along with the BiFC pair, however, is present predominantly in the nucleolus. It would appear that these three proteins, therefore, do not form a trimeric complex. However, when the BiFC pair was changed to Enp1-VC155 and Rio2-VN173 and Nob1p was co-transfected as Nob1-RFP, localization for all three proteins occurred together in the cell, with signal in both the cytoplasm and nucleolus. One explanation for this result could be that the localization patterns reflect the binding order of the proteins. It can be envisioned that Enp1p and Nob1p, both with localization in the nucleolus, must first bind in the nucleolus of the cell. Then, later, as the complex begins moving outward toward the cytoplasm, Rio2p binds and the complex is exported to the cytoplasm. If,

however, Nob1p and Rio2p are allowed to form first, their preference for cytoplasmic localization excludes the possibility that Enp1p could bind. Localization patterns for the combinations tested for this trio of proteins can be found in figure 6.16.

This project shows evidence for binary protein-protein interactions for numerous proteins involved in the D cleavage step during biogenesis of the small ribosomal subunit. In the nucleolus, these interactions include Enp1p with Rrp20p, Nob1p with Tsr1p, Dim1p with Rrp20p, Enp1p with Dim1p, and Enp1p with Tsr1p. Enp1p, Rrp20p, and Nob1p, in addition to exhibiting nucleolar localization patterns as binary interacting pairs, also colocalize together as a trio. The localization pattern within the nucleolus suggests that Nob1p binds to Tsr1p relatively early on during the process, possibly before Nob1p interacts with many of the other proteins involved. Similarly, with respect to Enp1p, binding with Rrp20p appears to occur at an early stage. The complex that is formed in the nucleolus is envisioned to contain Rrp20p, an RNA binding protein bound to the pre-rRNA molecule, and Nob1p, Dim1p, Enp1p, and Tsr1p. This complex then migrates out of the nucleolus and into the nucleoplasm of the nucleus. At this point Rio2p becomes involved in the process. Binary interactions detected in the nucleus, but not in the nucleolus, include Rio2p with Rrp20p, Rio2p with Enp1p, and Rio2p with Tsr1p. It is possible that Rio2p joins the previously formed complex consisting in part of Tsr1p, Enp1p, Nob1p, Rrp20p, and Dim1p. This complex then migrates through the nucleus toward the nuclear rim and nuclear pore structures. Prior to export from the nucleus to the cytoplasm, Rrp20p likely dissociates from this complex. Notably absent from the cytoplasm are binary interactions that involve Rrp20p. It is thought that this

protein dissociates and remains in the nucleus. Ltv1p is found in the cytoplasm in binary interactions with Rio2p, Nob1p, Enp1p, and Rio2p. In addition, a trimeric complex of Enp1p, Nob1p, and Ltv1p is suggested. Similarly, another trimeric complex appears unlikely based on the localization patterns when co-transfected. Evidence has been presented that argues against a trimeric complex composed of Enp1p, Nob1p, and Rrp20p. Ltv1p is involved in nuclear export of the small ribosomal subunit, which is further confirmed by its localization pattern moving from the nucleus to the cytoplasm. It is possible that as the complex consisting of Rio2p, Dim1p, Nob1p, Enp1p, and Tsr1p move through the cytoplasm that Rrp20p dissociates after which Ltv1p binds to the complex and assists in exporting the complex into the cytoplasm. Hrr25p is a kinase that plays a regulatory role during the assembly and disassembly process. The phosphorylation of Enp1p likely causes Enp1p to dissociate from the complex, and in addition, may be responsible for dissociation of numerous accessory proteins after which they are recycled by the cell, either degraded in the proteasome or shuttled back into the nucleus to repeat the process and produce additional small subunit precursors that will in turn be exported into the cytoplasm. A proposed order of assembly for the accessory proteins involved in cleavage and processing at site D during synthesis of the small ribosomal rRNA is provided below in figure 7.1.

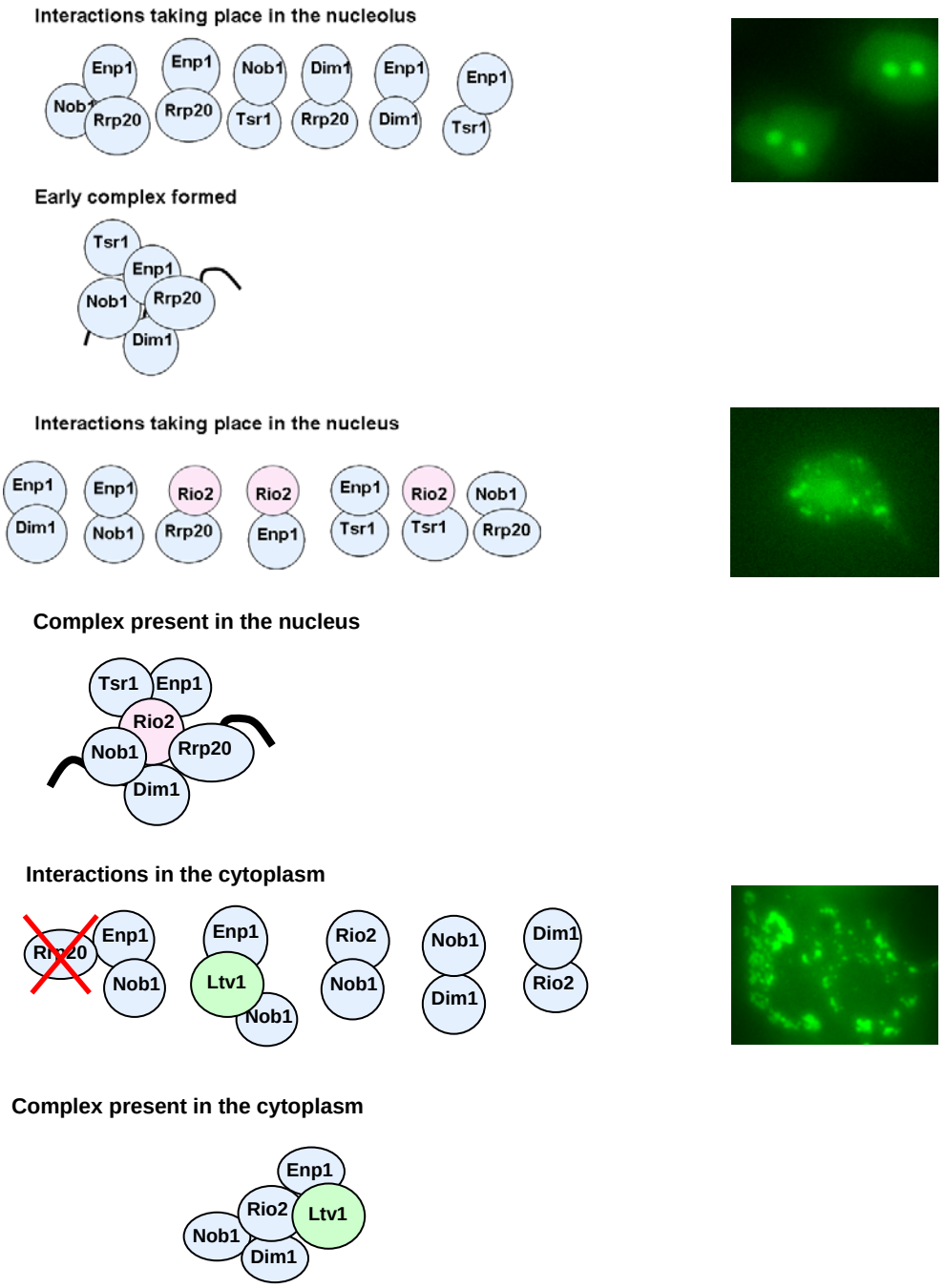


Figure 7.1: Assembly diagram showing localization of interactions.
Protein interactions in the nucleolus form an early complex. As the complex moves through the cell, Rio2p joins the complex. The complex moves through the nucleus and out into the cytoplasm where Ltv1 becomes involved. Representative pictures showing localization in each compartment are also shown.

Although mutant constructs were not created for true negative controls, some of the non-interacting partners themselves act as negative controls. A better method would have been to create mutations in the binding interface of the two potential interacting accessory proteins and fuse these to the fluorescent gene segments to create chimeras that, once transfected into the cell, would not be able to form a complex. However, this would require knowledge about the interaction site of the two proteins, which in many cases is not known and, in addition, would have been extremely time consuming. In order to establish that the fluorescent signals observed were not due to nonspecific interactions, protein-protein pairs that were suspected not to interact were combined in transfection assays. Rio2p and Hrr25p are both protein kinases that phosphorylate components involved in maturation of the small ribosomal subunit. It would not be expected for these two enzymes to interact with each other, and reassuringly, they do not. Similarly, an experiment was designed with Nob1p, a protein found only in the cytoplasm, and Prp4 a protein required for mRNA processing in the nucleus. When Nob1-VC155 and Prp4-VN173 were co-transfected, a fluorescent signal was not observed. However, this does not prove that nonspecific signals are not occurring. In some experiments conducted in the lab, nonspecific interactions appeared to be occurring. This issue needs to be explored further due to the nature of rRNA processing in which pre-rRNA and accessory proteins are packed together within the nucleolus and protein-RNA binding interactions are taking place.

FURTHER STUDIES

The creation of mutant proteins that could not interact with each other for each pair of proteins studied would have been an extremely time-consuming task, and for detected protein-protein interactions in which the interaction interface has not been determined, this would not have been possible. However, if this were to be done for each interacting pair, the possibility of fluorescence observed from a nonspecific interaction or simply the reforming of the two fluorescent halves of the fluorescent protein in the absence of a true interaction could be addressed. Further studies would need to address the issue of possible nonspecific interactions.

The current study could be expanded to include additional proteins involved in ribosome biogenesis. Numerous accessory proteins have been identified through the use of tandem affinity purification assays coupled with mass spectroscopy in high-throughput studies. However, many of the proteins identified have not been fully characterized and their functions remain unknown. Since the Gateway cloning and BiFC method described and used in this study does not require specific knowledge regarding the structure or function of the proteins, these proteins of unknown function could be studied using this technique. It is possible that by identifying interacting partners of these proteins and observing the localization of the interaction within the cell that insight could be gained as to their functions. Many of the small ribosomal unit accessory proteins previously identified were not examined in this study. The current study could be expanded to

include all of the known small subunit accessory proteins that have been identified in other large-scale studies. In doing so, it would be possible to confirm interactions *in vivo* for proteins such as Mpp10p which has been shown to co-purify only with components involved in the earliest stages of ribosome synthesis and the 35S pre-rRNA and is thought to be a marker for this stage in ribosome biogenesis. By expanding the pool of proteins examined, a large, more comprehensive list of binary interactions among the *trans*-acting factors could be established and additional functions may be elucidated in the process.

In addition to expanding the pool of proteins being studied, it would also be of use to expand the number of fluorescent proteins utilized in the experiments and employ a multi-color BiFC technique. By using a second fluorescent protein, ternary interactions among these accessory proteins could be examined. Competition assays could be designed to test binding order of the various proteins involved in assembly of the ribosomal subunits. A very preliminary step toward multi-color BiFC was made in this study and seems to have promise for teasing out additional information regarding the protein-protein interactions that are taking place during this complex process.

Finally, fluorescent resonance transfer (FRET) could be employed to study protein-protein interactions. FRET involves transfer of energy between a donor and acceptor molecule. The energy transfer takes place only if the two molecules are in close proximity to one another and requires them to be within 10 nm of each other[156]. FRET can be detected by using flow cytometry or fluorescence microscopy. This technique could be combined with BiFC in an assay described as BiFC-FRET to prove ternary

reactions among proteins involved in ribosome biogenesis. Like BiFC, this assay would also permit visualization of the location in the cell where the interaction takes place.

REFERENCES CITED

1. Lafontaine, D.L. and D. Tollervey, *The function and synthesis of ribosomes.* Nat Rev Mol Cell Biol, 2001. **2**(7): p. 514-20.
2. Lecompte, O., et al., *Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale.* Nucleic Acids Res, 2002. **30**(24): p. 5382-90.
3. Omer, A.D., et al., *RNA-modifying machines in archaea.* Mol Microbiol, 2003. **48**(3): p. 617-29.
4. Potter, S., P. Durovic, and P.P. Dennis, *Ribosomal RNA precursor processing by a eukaryotic U3 small nucleolar RNA-like molecule in an archaeon.* Science, 1995. **268**(5213): p. 1056-60.
5. Potter, S., et al., *Preribosomal RNA processing in archaea: characterization of the RNP endonuclease mediated processing of precursor 16S rRNA in the thermoacidophile Sulfolobus acidocaldarius.* Biochem Cell Biol, 1995. **73**(11-12): p. 813-23.
6. Jeffares, D.C., A.M. Poole, and D. Penny, *Relics from the RNA world.* J Mol Evol, 1998. **46**(1): p. 18-36.
7. Morrissey, J.P. and D. Tollervey, *Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system.* Trends Biochem Sci, 1995. **20**(2): p. 78-82.
8. Young, R.A. and J.A. Steitz, *Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA.* Proc Natl Acad Sci U S A, 1978. **75**(8): p. 3593-7.
9. King, T.C., R. Sirdeskmukh, and D. Schlessinger, *Nucleolytic processing of ribonucleic acid transcripts in procaryotes.* Microbiol Rev, 1986. **50**(4): p. 428-51.
10. Brown, J.W. and P.J. Shaw, *The role of the plant nucleolus in pre-mRNA processing.* Curr Top Microbiol Immunol, 2008. **326**: p. 291-311.
11. Sharma, M.R., et al., *Cryo-EM study of the spinach chloroplast ribosome reveals the structural and functional roles of plastid-specific ribosomal proteins.* Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19315-20.
12. Barakat, A., K.F. Muller, and L.E. Saenz-de-Miera, *Molecular evolutionary analyses of the Arabidopsis L7 ribosomal protein gene family.* Gene, 2007. **403**(1-2): p. 143-50.
13. Hajarnavis, A. and R. Durbin, *A conserved sequence motif in 3' untranslated regions of ribosomal protein mRNAs in nematodes.* RNA, 2006. **12**(10): p. 1786-9.
14. Azuma, M., et al., *Perturbation of rRNA synthesis in the bap28 mutation leads to apoptosis mediated by p53 in the zebrafish central nervous system.* J Biol Chem, 2006. **281**(19): p. 13309-16.
15. Karsi, A., et al., *Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression.* Gene, 2002. **291**(1-2): p. 177-86.

16. McCallum, F.S. and B.E. Maden, *Human 18 S ribosomal RNA sequence inferred from DNA sequence. Variations in 18 S sequences and secondary modification patterns between vertebrates.* Biochem J, 1985. **232**(3): p. 725-33.
17. Newman, D.R., et al., *Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription.* RNA, 2000. **6**(6): p. 861-79.
18. Watkins, N.J., et al., *In vitro assembly of the mouse U14 snoRNP core complex and identification of a 65-kDa box C/D-binding protein.* RNA, 1998. **4**(5): p. 582-93.
19. Zahradka, P., D.E. Larson, and B.H. Sells, *Regulation of ribosome biogenesis in differentiated rat myotubes.* Mol Cell Biochem, 1991. **104**(1-2): p. 189-94.
20. Andersen, J.S., et al., *Nucleolar proteome dynamics.* Nature, 2005. **433**(7021): p. 77-83.
21. Wool, I.G., Y.L. Chan, and A. Gluck, *Structure and evolution of mammalian ribosomal proteins.* Biochem Cell Biol, 1995. **73**(11-12): p. 933-47.
22. Ansel, K.M., et al., *Mouse Eri1 interacts with the ribosome and catalyzes 5.8S rRNA processing.* Nat Struct Mol Biol, 2008. **15**(5): p. 523-30.
23. Goldman, W.E., et al., *Mouse rDNA: sequences and evolutionary analysis of spacer and mature RNA regions.* Mol Cell Biol, 1983. **3**(8): p. 1488-500.
24. Venema, J. and D. Tollervey, *Processing of pre-ribosomal RNA in Saccharomyces cerevisiae.* Yeast, 1995. **11**(16): p. 1629-50.
25. Kressler, D., P. Linder, and J. de La Cruz, *Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae.* Mol Cell Biol, 1999. **19**(12): p. 7897-912.
26. Venema, J. and D. Tollervey, *Ribosome synthesis in Saccharomyces cerevisiae.* Annu Rev Genet, 1999. **33**: p. 261-311.
27. Warner, J.R., *The economics of ribosome biosynthesis in yeast.* Trends Biochem Sci, 1999. **24**(11): p. 437-40.
28. Gallagher, J.E., et al., *RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components.* Genes Dev, 2004. **18**(20): p. 2506-17.
29. Bernstein, K.A. and S.J. Baserga, *The small subunit processome is required for cell cycle progression at G1.* Mol Biol Cell, 2004. **15**(11): p. 5038-46.
30. Gasch, A.P., et al., *Genomic expression programs in the response of yeast cells to environmental changes.* Mol Biol Cell, 2000. **11**(12): p. 4241-57.
31. Jorgensen, P., et al., *A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size.* Genes Dev, 2004. **18**(20): p. 2491-505.
32. Fatica, A., et al., *Nob1p is required for cleavage of the 3′ end of 18S rRNA.* Mol Cell Biol, 2003. **23**(5): p. 1798-807.
33. Fatica, A., D. Tollervey, and M. Dlakic, *PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA.* RNA, 2004. **10**(11): p. 1698-701.
34. Fatica, A. and D. Tollervey, *Making ribosomes.* Curr Opin Cell Biol, 2002. **14**(3): p. 313-8.

35. Scheer, U. and R. Hock, *Structure and function of the nucleolus.* Curr Opin Cell Biol, 1999. **11**(3): p. 385-90.
36. Pederson, T., *The plurifunctional nucleolus.* Nucleic Acids Res, 1998. **26**(17): p. 3871-6.
37. Visintin, R. and A. Amon, *The nucleolus: the magician's hat for cell cycle tricks.* Curr Opin Cell Biol, 2000. **12**(6): p. 752.
38. Lamond, A.I. and W.C. Earnshaw, *Structure and function in the nucleus.* Science, 1998. **280**(5363): p. 547-53.
39. Lamond, A.I. and J.E. Sleeman, *Nuclear substructure and dynamics.* Curr Biol, 2003. **13**(21): p. R825-8.
40. Henry, Y., et al., *The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site.* EMBO J, 1994. **13**(10): p. 2452-63.
41. Skryabin, K.G., et al., *Structure and function of the nontranscribed spacer regions of yeast rDNA.* Nucleic Acids Res, 1984. **12**(6): p. 2955-68.
42. Tschochner, H. and E. Hurt, *Pre-ribosomes on the road from the nucleolus to the cytoplasm.* Trends Cell Biol, 2003. **13**(5): p. 255-63.
43. Veldman, G.M., et al., *Some characteristics of processing sites in ribosomal precursor RNA of yeast.* Nucleic Acids Res, 1980. **8**(13): p. 2907-20.
44. Dunbar, D.A., et al., *Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast.* Mol Cell Biol, 1997. **17**(10): p. 5803-12.
45. Hughes, J.M. and M. Ares, Jr., *Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA.* EMBO J, 1991. **10**(13): p. 4231-9.
46. Udem, S.A. and J.R. Warner, *The cytoplasmic maturation of a ribosomal precursor ribonucleic acid in yeast.* J Biol Chem, 1973. **248**(4): p. 1412-6.
47. Fromont-Racine, M., et al., *Ribosome assembly in eukaryotes.* Gene, 2003. **313**: p. 17-42.
48. Warner, J.R., et al., *Ribosomal proteins and the assembly of ribosomes in eukaryotes.* Biochem J, 1972. **129**(3): p. 29P-30P.
49. Dlakic, M., *The ribosomal subunit assembly line.* Genome Biol, 2005. **6**(10): p. 234.
50. Granneman, S. and S.J. Baserga, *Ribosome biogenesis: of knobs and RNA processing.* Exp Cell Res, 2004. **296**(1): p. 43-50.
51. Ju, Q. and J.R. Warner, *Ribosome synthesis during the growth cycle of Saccharomyces cerevisiae.* Yeast, 1994. **10**(2): p. 151-7.
52. Tollervey, D., *Trans-acting factors in ribosome synthesis.* Exp Cell Res, 1996. **229**(2): p. 226-32.
53. Moy, T.I. and P.A. Silver, *Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins.* Genes Dev, 1999. **13**(16): p. 2118-33.
54. Kiss, T., *Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.* Cell, 2002. **109**(2): p. 145-8.

55. Tollervey, D. and E.C. Hurt, *The role of small nucleolar ribonucleoproteins in ribosome synthesis.* Mol Biol Rep, 1990. **14**(2-3): p. 103-6.
56. Krogan, N.J., et al., *High-definition macromolecular composition of yeast RNA-processing complexes.* Mol Cell, 2004. **13**(2): p. 225-39.
57. Osheim, Y.N., et al., *Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae.* Mol Cell, 2004. **16**(6): p. 943-54.
58. Bernstein, K.A., et al., *The small-subunit processome is a ribosome assembly intermediate.* Eukaryot Cell, 2004. **3**(6): p. 1619-26.
59. Kiss-Laszlo, Z., et al., *Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs.* Cell, 1996. **85**(7): p. 1077-88.
60. Kiss-Laszlo, Z., Y. Henry, and T. Kiss, *Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA.* EMBO J, 1998. **17**(3): p. 797-807.
61. Saffer, L.D. and O.L. Miller, Jr., *Electron microscopic study of Saccharomyces cerevisiae rDNA chromatin replication.* Mol Cell Biol, 1986. **6**(4): p. 1148-57.
62. Mougey, E.B., et al., *The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes.* Genes Dev, 1993. **7**(8): p. 1609-19.
63. Dragon, F., et al., *A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis.* Nature, 2002. **417**(6892): p. 967-70.
64. Sharma, K., J. Venema, and D. Tollervey, *The 5' end of the 18S rRNA can be positioned from within the mature rRNA.* RNA, 1999. **5**(5): p. 678-86.
65. Sharma, K. and D. Tollervey, *Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing.* Mol Cell Biol, 1999. **19**(9): p. 6012-9.
66. Grandi, P., et al., *90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors.* Mol Cell, 2002. **10**(1): p. 105-15.
67. Morrissey, J.P. and D. Tollervey, *Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis.* Mol Cell Biol, 1993. **13**(4): p. 2469-77.
68. Beltrame, M., Y. Henry, and D. Tollervey, *Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA.* Nucleic Acids Res, 1994. **22**(23): p. 5139-47.
69. Allmang, C., et al., *Processing of the yeast pre-rRNA at sites A(2) and A(3) is linked.* RNA, 1996. **2**(1): p. 63-73.
70. Hage, A.E. and D. Tollervey, *A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria?* RNA Biol, 2004. **1**(1): p. 10-5.
71. Beltrame, M. and D. Tollervey, *Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis.* EMBO J, 1995. **14**(17): p. 4350-6.
72. Venema, J. and D. Tollervey, *RRP5 is required for formation of both 18S and 5.8S rRNA in yeast.* EMBO J, 1996. **15**(20): p. 5701-14.

73. Beltrame, M. and D. Tollervey, *Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA.* EMBO J, 1992. **11**(4): p. 1531-42.
74. Venema, J., Y. Henry, and D. Tollervey, *Two distinct recognition signals define the site of endonucleolytic cleavage at the 5'-end of yeast 18S rRNA.* EMBO J, 1995. **14**(19): p. 4883-92.
75. van Nues, R.W., et al., *Separate structural elements within internal transcribed spacer 1 of Saccharomyces cerevisiae precursor ribosomal RNA direct the formation of 17S and 26S rRNA.* Nucleic Acids Res, 1994. **22**(6): p. 912-9.
76. Allmang, C., et al., *Recognition of cleavage site A(2) in the yeast pre-rRNA.* RNA, 1996. **2**(1): p. 51-62.
77. Lafontaine, D.L., T. Preiss, and D. Tollervey, *Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis?* Mol Cell Biol, 1998. **18**(4): p. 2360-70.
78. Lafontaine, D., et al., *The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18 S rRNA is essential in yeast.* J Mol Biol, 1994. **241**(3): p. 492-7.
79. Vanrobays, E., et al., *TOR regulates the subcellular distribution of DIM2, a KH domain protein required for cotranscriptional ribosome assembly and pre-40S ribosome export.* RNA, 2008. **14**(10): p. 2061-73.
80. Vanrobays, E., et al., *Dim2p, a KH-domain protein required for small ribosomal subunit synthesis.* RNA, 2004. **10**(4): p. 645-56.
81. Vanrobays, E., et al., *Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae.* Mol Cell Biol, 2003. **23**(6): p. 2083-95.
82. Vanrobays, E., et al., *Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein.* EMBO J, 2001. **20**(15): p. 4204-13.
83. Gelperin, D., et al., *Bms1p, a novel GTP-binding protein, and the related Tsr1p are required for distinct steps of 40S ribosome biogenesis in yeast.* RNA, 2001. **7**(9): p. 1268-83.
84. Tone, Y., et al., *Nob1p, a new essential protein, associates with the 26S proteasome of growing saccharomyces cerevisiae cells.* Gene, 2000. **243**(1-2): p. 37-45.
85. Tone, Y. and E.A. Toh, *Nob1p is required for biogenesis of the 26S proteasome and degraded upon its maturation in Saccharomyces cerevisiae.* Genes Dev, 2002. **16**(24): p. 3142-57.
86. Seiser, R.M., et al., *Ltv1 is required for efficient nuclear export of the ribosomal small subunit in Saccharomyces cerevisiae.* Genetics, 2006. **174**(2): p. 679-91.
87. Oeffinger, M., M. Dlakic, and D. Tollervey, *A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits.* Genes Dev, 2004. **18**(2): p. 196-209.
88. Liu, P.C. and D.J. Thiele, *Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis.* Mol Biol Cell, 2001. **12**(11): p. 3644-57.

89. Eschrich, D., et al., *Nep1p (Emg1p), a novel protein conserved in eukaryotes and archaea, is involved in ribosome biogenesis.* Curr Genet, 2002. **40**(5): p. 326-38.
90. Perez-Fernandez, J., et al., *The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism.* Mol Cell Biol, 2007. **27**(15): p. 5414-29.
91. Bleichert, F., et al., *The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation.* Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9464-9.
92. Rempola, B., et al., *Fcf1p and Fcf2p are novel nucleolar Saccharomyces cerevisiae proteins involved in pre-rRNA processing.* Biochem Biophys Res Commun, 2006. **346**(2): p. 546-54.
93. Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.
94. Ito, T., et al., *A comprehensive two-hybrid analysis to explore the yeast protein interactome.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4569-74.
95. Bousquet-Antonelli, C., et al., *A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs.* EMBO J, 1997. **16**(15): p. 4770-6.
96. Bousquet-Antonelli, C., et al., *Rrp8p is a yeast nucleolar protein functionally linked to Gar1p and involved in pre-rRNA cleavage at site A2.* RNA, 2000. **6**(6): p. 826-43.
97. Tremblay, A., et al., *A physical interaction between Gar1p and Rnt1pi is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins.* Mol Cell Biol, 2002. **22**(13): p. 4792-802.
98. Girard, J.P., et al., *GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast.* EMBO J, 1992. **11**(2): p. 673-82.
99. Ganot, P., M. Caizergues-Ferrer, and T. Kiss, *The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation.* Genes Dev, 1997. **11**(7): p. 941-56.
100. Henras, A., et al., *Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs.* EMBO J, 1998. **17**(23): p. 7078-90.
101. Gromadka, R. and J. Rytka, *The KRR1 gene encodes a protein required for 18S rRNA synthesis and 40S ribosomal subunit assembly in Saccharomyces cerevisiae.* Acta Biochim Pol, 2000. **47**(4): p. 993-1005.
102. Sasaki, T., E.A. Toh, and Y. Kikuchi, *Yeast Krr1p physically and functionally interacts with a novel essential Kri1p, and both proteins are required for 40S ribosome biogenesis in the nucleolus.* Mol Cell Biol, 2000. **20**(21): p. 7971-9.
103. Gromadka, R., et al., *A novel cross-phylum family of proteins comprises a KRR1 (YCL059c) gene which is essential for viability of Saccharomyces cerevisiae cells.* Gene, 1996. **171**(1): p. 27-32.
104. Karkusiewicz, I., et al., *Functional and physical interactions of Faf1p, a Saccharomyces cerevisiae nucleolar protein.* Biochem Biophys Res Commun, 2004. **319**(2): p. 349-57.

105. Roos, J., et al., *ENP1, an essential gene encoding a nuclear protein that is highly conserved from yeast to humans.* Gene, 1997. **185**(1): p. 137-46.
106. Chen, W., et al., *Enp1, a yeast protein associated with U3 and U14 snoRNAs, is required for pre-rRNA processing and 40S subunit synthesis.* Nucleic Acids Res, 2003. **31**(2): p. 690-9.
107. Schafer, T., et al., *The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes.* EMBO J, 2003. **22**(6): p. 1370-80.
108. Schafer, T., et al., *Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit.* Nature, 2006. **441**(7093): p. 651-5.
109. O'Farrell, H.C., J.N. Scarsdale, and J.P. Rife, *Crystal structure of KsgA, a universally conserved rRNA adenine dimethyltransferase in Escherichia coli.* J Mol Biol, 2004. **339**(2): p. 337-53.
110. Zhang, Y.Z., et al., *The evolutionarily conserved Dim1 protein defines a novel branch of the thioredoxin fold superfamily.* Physiol Genomics, 1999. **1**(3): p. 109-18.
111. O'Farrell, H.C., et al., *Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution.* RNA, 2006. **12**(5): p. 725-33.
112. Poldermans, B., C.P. Van Buul, and P.H. Van Knippenberg, *Studies on the function of two adjacent N6,N6-dimethyladenosines near the 3' end of 16 S ribosomal RNA of Escherichia coli. II. The effect of the absence of the methyl groups on initiation of protein biosynthesis.* J Biol Chem, 1979. **254**(18): p. 9090-3.
113. O'Farrell, H.C., et al., *Sequence and structural evolution of the KsgA/Dim1 methyltransferase family.* BMC Res Notes, 2008. **1**(1): p. 108.
114. Lafontaine, D., J. Vandenhaute, and D. Tollervey, *The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast.* Genes Dev, 1995. **9**(20): p. 2470-81.
115. Carnahan, R.H., et al., *Dim1p is required for efficient splicing and export of mRNA encoding lid1p, a component of the fission yeast anaphase-promoting complex.* Eukaryot Cell, 2005. **4**(3): p. 577-87.
116. Senapin, S., et al., *RRP20, a component of the 90S preribosome, is required for pre-18S rRNA processing in Saccharomyces cerevisiae.* Nucleic Acids Res, 2003. **31**(10): p. 2524-33.
117. Krecic, A.M. and M.S. Swanson, *hnRNP complexes: composition, structure, and function.* Curr Opin Cell Biol, 1999. **11**(3): p. 363-71.
118. Gavin, A.C., et al., *Functional organization of the yeast proteome by systematic analysis of protein complexes.* Nature, 2002. **415**(6868): p. 141-7.
119. Peng, W.T., et al., *A panoramic view of yeast noncoding RNA processing.* Cell, 2003. **113**(7): p. 919-33.
120. Loar, J.W., et al., *Genetic and biochemical interactions among Yar1, Ltv1 and Rps3 define novel links between environmental stress and ribosome biogenesis in Saccharomyces cerevisiae.* Genetics, 2004. **168**(4): p. 1877-89.

121. Geerlings, T.H., et al., *Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae.* J Biol Chem, 2003. **278**(25): p. 22537-45.
122. Piehler, J., *New methodologies for measuring protein interactions in vivo and in vitro.* Curr Opin Struct Biol, 2005. **15**(1): p. 4-14.
123. Wu, L.F., et al., *Large-scale prediction of Saccharomyces cerevisiae gene function using overlapping transcriptional clusters.* Nat Genet, 2002. **31**(3): p. 255-65.
124. Brown, M.P., et al., *Knowledge-based analysis of microarray gene expression data by using support vector machines.* Proc Natl Acad Sci U S A, 2000. **97**(1): p. 262-7.
125. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-6.
126. Ito, T., et al., *Roles for the two-hybrid system in exploration of the yeast protein interactome.* Mol Cell Proteomics, 2002. **1**(8): p. 561-6.
127. Phizicky, E.M. and S. Fields, *Protein-protein interactions: methods for detection and analysis.* Microbiol Rev, 1995. **59**(1): p. 94-123.
128. Uetz, P., et al., *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. **403**(6770): p. 623-7.
129. Puig, O., et al., *The tandem affinity purification (TAP) method: a general procedure of protein complex purification.* Methods, 2001. **24**(3): p. 218-29.
130. Ho, Y., et al., *Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry.* Nature, 2002. **415**(6868): p. 180-3.
131. Kumar, A., et al., *An integrated approach for finding overlooked genes in yeast.* Nat Biotechnol, 2002. **20**(1): p. 58-63.
132. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.
133. Combs, D.J., et al., *Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis.* Mol Cell Biol, 2006. **26**(2): p. 523-34.
134. Wade, C.H., M.A. Umbarger, and M.A. McAlear, *The budding yeast rRNA and ribosome biosynthesis (RRB) regulon contains over 200 genes.* Yeast, 2006. **23**(4): p. 293-306.
135. Grigull, J., et al., *Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors.* Mol Cell Biol, 2004. **24**(12): p. 5534-47.
136. Shimomura, O., F.H. Johnson, and Y. Saiga, *Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.* J Cell Comp Physiol, 1962. **59**: p. 223-39.
137. Prasher, D., R.O. McCann, and M.J. Cormier, *Cloning and expression of the cDNA coding for aequorin, a bioluminescent calcium-binding protein.* Biochem Biophys Res Commun, 1985. **126**(3): p. 1259-68.
138. Ormo, M., et al., *Crystal structure of the Aequorea victoria green fluorescent protein.* Science, 1996. **273**(5280): p. 1392-5.

139. Matz, M.V., et al., *Fluorescent proteins from nonbioluminescent Anthozoa species.* Nat Biotechnol, 1999. **17**(10): p. 969-73.
140. Shyu, Y.J., et al., *Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions.* Biotechniques, 2006. **40**(1): p. 61-6.
141. Hu, C.D., Y. Chinenov, and T.K. Kerppola, *Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation.* Mol Cell, 2002. **9**(4): p. 789-98.
142. Hu, C.D. and T.K. Kerppola, *Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis.* Nat Biotechnol, 2003. **21**(5): p. 539-45.
143. Kerppola, T.K., *Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells.* Nat Protoc, 2006. **1**(3): p. 1278-86.
144. Hu, C.D., A.V. Grinberg, and T.K. Kerppola, *Visualization of protein interactions in living cells using bimolecular fluorescence complementation (BiFC) analysis.* Curr Protoc Protein Sci, 2005. **Chapter 19**: p. Unit 19 10.
145. Rizzo, M.A., et al., *An improved cyan fluorescent protein variant useful for FRET.* Nat Biotechnol, 2004. **22**(4): p. 445-9.
146. Hartley, J.L., G.F. Temple, and M.A. Brasch, *DNA cloning using in vitro site-specific recombination.* Genome Res, 2000. **10**(11): p. 1788-95.
147. Landy, A., *Dynamic, structural, and regulatory aspects of lambda site-specific recombination.* Annu Rev Biochem, 1989. **58**: p. 913-49.
148. Ausubel, F.M., *Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology*. 5th ed. 2002, New York: Wiley. 2 v. (various pagings).
149. Bernard, P. and M. Couturier, *Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes.* J Mol Biol, 1992. **226**(3): p. 735-45.
150. *Cos-7*. [Electronic] 2008 [cited 2008; Cos-7 cell description, ATCC number CRL-1651]. Available from: http://www.atcc.org/ATCCAdvancedCatalogSearch/ProductDetails/tabid/452/Default.aspx?ATCCNum=CRL-1651&Template=cellBiology.
151. (2004) *Mirus TransIT-COS Transfection Kit*. 1-5.
152. Sprinzak, E., S. Sattath, and H. Margalit, *How reliable are experimental protein-protein interaction data?* J Mol Biol, 2003. **327**(5): p. 919-23.
153. Carpenter, A.E., et al., *CellProfiler: image analysis software for identifying and quantifying cell phenotypes.* Genome Biol, 2006. **7**(10): p. R100.
154. Lamprecht, M.R., D.M. Sabatini, and A.E. Carpenter, *CellProfiler: free, versatile software for automated biological image analysis.* Biotechniques, 2007. **42**(1): p. 71-5.
155. Vokes, M.S. and A.E. Carpenter, *Using CellProfiler for automatic identification and measurement of biological objects in images.* Curr Protoc Mol Biol, 2008. **Chapter 14**: p. Unit 14 17.

156. Stryer, L., *Fluorescence energy transfer as a spectroscopic ruler.* Annu Rev Biochem, 1978. **47**: p. 819-46.

APPENDIX A

PCR PRIMERS

Primer Extensions

Primer extensions added to gene-specific primers for each gene of interest in order to provide attachment site sequences for recombining into Gateway vectors are shown below.

5’ sequence prior to ATG of the gene of interest:
GGGGACAAGTTTGTACAAAAAAGCAGGCTCC

3’ sequence added to the end of the gene of interest
GGACCCAGCTTTCTTGTACAAAGTGGTCCCC
(Written from 5’ to 3’ this extension is as follows:
GGGGCCACTTTGTACAAGAAAGCTGGGTCC)

Primers for Each Gene of Interest with BiFC Nested Primer Extensions

Primers are written in the 5’ to 3’ direction and include the BiFC nested primer segments. Gene-specific sequences are in bold type.

Dim1 5’: GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGCCGAAGGTCAAGTC**
Dim1 3’: GGGGACCACTTTGTACAAGAAAGCTGGGTCC**TATCTGTTTCACAGGATACACAG**

Emg1 5’: GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGCCGCGCCCA**
Emg1 3’: GGGGACCACTTTGTACAAGAAAGCTGGGTCC**AATGACCCCCCATACTTCC**

Enp1 5’: GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGCCCAAATTCAAGGC**
Enp1 3’: GGGGACCACTTTGTACAAGAAAGCTGGGTCCCTCC**ACGGTGATGGGAAC**

Enp2 5’: GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGCAGGTCTCCAGCCTC**
Enp2 3’: GGGGACCACTTTGTACAAGAAAGCTGGGTCC**AAACGACCGTCCTCTTT**

Fcf1 5’: GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGGGAAGCAAAAGAAAAC**
Fcf1 3’: GGGGACCACTTTGTACAAGAAAGCTGGGTCC**GAATCGAGGGGCTCCAT**

Fcf2 5’: GTACAAAAAAGCAGGCTCC**ATGGTGGTTACCAGATCTGC**
Fcf2 3’: GTACAAGAAAGCTGGGTCC**ATTGCGAAATTTCTTCTTCTTT**

Gar1p 5’: GTACAAAAAAGCAGGCTCC**ATGTCTTTTCGAGGCGGA**
Gar1p 3’: GTACAAGAAAGCTGGGTCC**ATGTCCTCTCCCTCTGAAACC**

Hrr25p 5’: GTACAAAAAAGCAGGCTCC**ATGGAGCTACGTGTGGGG**
Hrr25p 3’: GTACAAGAAAGCTGGGTCC**CTTCCCGAGATGGTCAAATG**

Krr1p 5': GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGCGTCTCCCTCGC**
Krr1p 3': GGGACCACTTTGTACAAGAAAGCTGGGTCC**CTTTTTTTTCTTCTTTTTCTTTTCA**

Ltv1 5': GTACAAAAAAGCAGGCTCC**ATGCCTCACAGGAAGAAAA**
Ltv1 3': GTACAAGAAAGCTGGGTCC**TAGCTTTAGACCCTCAACATT**

Nob1 5': GTACAAAAAAGCAGGCTCC**ATGGCTCCAGTGGAGCA**
Nob1 3': GTACAAGAAAGCTGGGTCC**CCTTTTCTTCACAAACTTCTTTC**

Rio2 5': GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGGGAAAGTGAATGTGG**
Rio2 3': GGGGACCACTTTGTACAAGAAAGCTGGGTCC**TTCTCCCCAAAAGCTGG**

Rrp20 5': GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGAATCCGAAATGGAA**
Rrp20 3': GGGGACCACTTTGTACAAGAAAGCTGGGTCC**GAATCGATCTGCTGATCTGC**

Tsr1 5': GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGGCGGCCCACC**
Tsr1 3': GGGGACCACTTTGTACAAGAAAGCTGGGTCC**CTCCATGCCCCCTTGAG**

Utp4 5': GTACAAAAAAGCAGGCTCC**ATGGGTGAATTTAAGGTCCAT**
Utp4 3': GTACAAGAAAGCTGGGTCCGGTTC**CAAATTTCTTCTTTTTA**

Utp23 5': GGGGACAAGTTTGTACAAAAAAGCAGGCTCC**ATGAAGATCACAAGGCAGAA**
Utp23 3': GGGGACCACTTTGTACAAGAAAGCTGGG**TCCTTCTCCTTCTGCATTCTGCT**

Universal Screening and Sequencing Primers

Several primers were used for screening and sequencing constructs. These were designated 5' universal which is complementary to the 5' region of all BiFC constructs. Two additional primers, 3'-middle and 3'-end were also used. The 3'-end primer was designed to amplify the 3' end of the BiFC coding region. The 3'-middle primer was designed to amplify the early portion of the BiFC coding region and during sequencing reactions would permit a sequencing trace that would include both a portion of the gene encoding the fluorescent protein and the end of the coding region of the accessory protein cloned into the final product. Additional screening primers were developed to amplify the BiFC constructs and verify their sequences. The 5' universal Venus/Cerulean primer

amplifies the 5’ end of fluorescent constructs while the 3’VC155 amplifies the 3’ end of C-terminus constructs and 3’ VN173 amplifies the N-terminus constructs.

5’ universal primer:	CGGTGGGAGGTCTATATAAGC
3’-end primer:	AGGAGAGGCACTGGCG
3’-middle primer:	GGTGGCGATGGATCTTCT
3’ universal VN-173:	ACCAGGATGGGCACCAC
3’ universal VC-155:	CTCGATGTTGTGGCGAT

Primers were also used for screening donor constructs prior to recombining into destination vectors. These primers were designed specifically for pDONR201 and are listed below.

5’ pDONR:	TAACGCTAGCATGGATCTC
3’ pDONR:	GTAACATCAGAGATTTTGAGACAC