University of South Bohemia
Faculty of Science

Identifying Protein-Protein Interactions of *Trypanosoma brucei* F_oF_1-ATP Synthase Subunits Using the Yeast Two Hybrid System

Bachelor thesis

Michaela Kunzová

Supervisor: RNDr. Alena Panicucci Zíková, PhD.

České Budějovice, 2016

Kunzová M., 2016. Identifying Protein-Protein Interactions of Trypanosoma brucei FoF1-ATP Synthase Subunits Using the Yeast Two Hybrid System. (Bc.Thesis in English) - 56 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Annotation:
We are utilizing the yeast two-hybrid screen to identify the interacting partners of *the T. brucei* F_oF_1-ATP synthase subunits that have no known homologues outside of Trypanosomatids. A better understanding of the framework of these 14 unknown proteins may suggest their function within this essential enzyme.

I hereby declare that I worked on this bachelor thesis on my own and used only the resources mentioned in the bibliography.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages.

Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

Date.....

Signature.....

Acknowledgement

I would like to thank PhD. Alena Panicucci Zíková for letting me peek behind the curtain of how modern science is made when she let me work in her lab. Next, I would like to express my gratitude towards Brian Panicucci, who was always helpful throughout the whole project, never turned me down, when I was asking many questions, and who corrected my grammar with a will in this thesis. Other big thank you goes to my families; to the one, in which I was born, and which taught me how to walk, talk and how to be a good person, and to the other one, which I entered in first year of my studies here in České Budějovice, which showed me that there are still good people in the world, who do not hesitate to help others and all of whom patiently helped me to do my first steps in the world in science. Thank you.

Table of content

1 Introduction

1.1 Trypanosoma brucei

Trypanosoma brucei is a flagellated parasite that belongs to the order Trypanosomatida. This group of protozoans is notoriously comprised of pathogens that are responsible for some of the world's worst neglected tropical diseases. *T. brucei* is the causative agent of human African trypanosomiasis (HAT) and nagana, which afflicts domesticated animals. The life cycle of this parasite alternates between a mammalian host and its insect vector, the tsetse fly (Figure 1). Since the insect vector is restricted to Sub-Saharan Africa, most of the people affected by this disease live in remote rural areas with limited or no access to proper health service. While recent vector control programs have successfully limited new cases of HAT to under 10,000 patients in 2009, the available treatments for this lethal disease are often toxic and difficult to administer (http://www.who.int/en/). In addition, *T. brucei* is rapidly displaying signs of drug resistance. Therefore, a renewed emphasis on the discovery of druggable targets is challenging research scientists to identify unique and essential *T. brucei* biological processes.

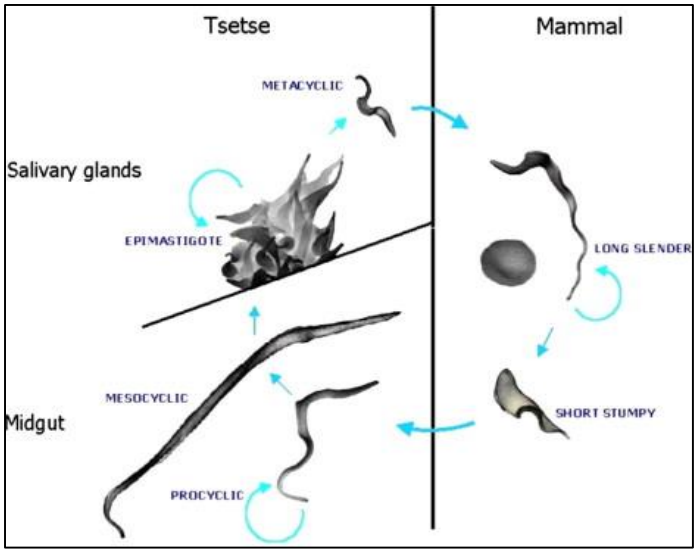

Figure 1: Life cycle of Trypanosoma brucei (Holmes, 2012).

1.2 Metabolic rewiring during *T. brucei* differentiation

As this extracellular parasite alternates between the blood and tissue fluid of the mammalian host and the midgut and salivary glands of the tsetse fly, it must adapt its metabolism as it encounters different nutrients in these various environments (Holmes, 2012). The procyclic insect form (PF) must utilize L-proline, and to a lesser extent L-threonine, as a carbon source to synthesize chemical energy in the form of ATP. During this life cycle, the mitochondrion is fully developed with a highly branched network and abundant cristae. Amino acid catabolism results in a citric acid (TCA) cycle intermediate that is further oxidized to contribute electrons to the electron transport chain (ETC). The proton-pumping ETC complexes subsequently contribute to the mitochondrial (mt) membrane potential (mΨΔ), which is coupled to ATP synthesis by the F_oF_1-ATP synthase (Figure 2). This oxidative phosphorylation (OXPHOS) pathway efficiently produces enough ATP to meet most of the energetic needs of the cell (Besteiro et al, 2005).

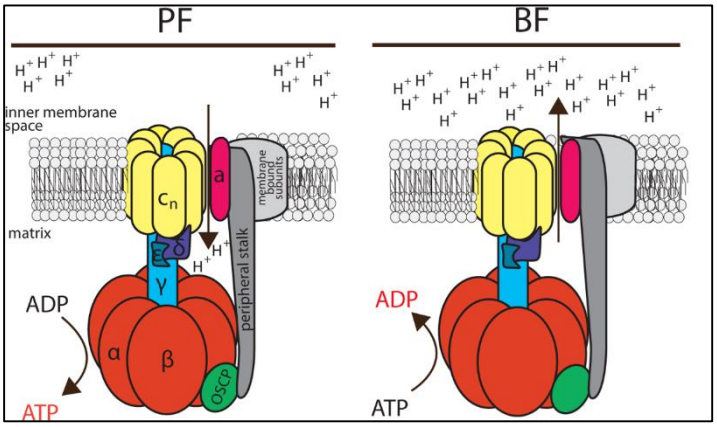

Figure 2: Schematic representing the function of F_oF_1-ATP synthase in the procyclic form (PF), where it synthesizes ATP, and the bloodstream form (BF), which pumps protons to maintain the mitochondrial membrane potential (mΨΔ) (Šubrtová et al, 2015).

However, when the parasite enters the glucose-rich bloodstream of the mammalian host, this bloodstream form (BF) is able to produce enough ATP through glycolysis alone. This results in a reduced tubular mitochondrion that lacks cristae, TCA enzymes and cyctochrome-mediated respiration (Hannaert et al, 2003). The BF mitochondrion is not dormant though, as it still depends on mt DNA expression and it retains several vital processes like RNA-editing, Fe-S cluster assembly and mt fatty acid synthesis. All of these functions involve nuclear encoded gene products, which require the mΨΔ for protein import into the organelle (Šubrtová et al, 2015). Therefore, without the ETC to transfer protons, the enzymatic activity of F_oF_1-ATP

synthase reverses, hydrolyzing ATP to generate mechanical energy to pump protons into the mt inner membrane space. While this continuous hydrolytic activity of the F_oF_1-ATPase maintains the essential BF *T. brucei* mΨΔ, this function is rare in mammals, occurring only when cells briefly experience hypoxic conditions. Therefore, ATP hydrolysis by F_oF_1-ATPase is a potential drug target in the infectious stage of *T. brucei* (Schnaufer et al, 2005).

1.3 The Structure and Function of the F_oF_1 ATP-synthase

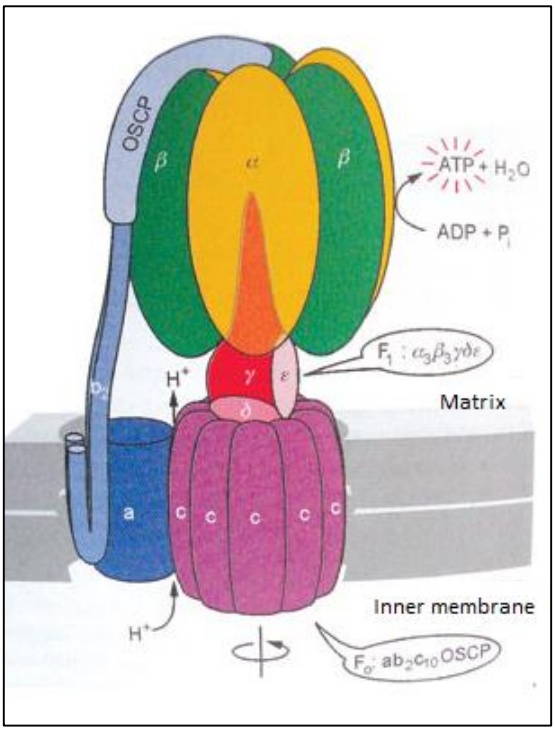

Figure 3: Eukaryotic F_oF_1-ATP synthase (Müller-Esterl, 2007).

The F_oF_1-ATP synthase is a multiprotein enzyme that consists of two functional domains: the catalytic F_1 entity that extends into the mt matrix and the F_o region that is embedded in the mt inner membrane (Figure 3). The F_o moiety translates electrochemical potential into mechanical energy by allowing the protons pumped into the mt inner membrane space by the ETC to flow down their concentration gradient and back into the mt matrix through the proton pore created between subunit a and the multimeric c-ring. This energetic flow of protons causes the c-ring to rotate within the plane of the membrane, acting as a rotary nanomachine whose kinetic energy is propelled through the F_1 coiled-coil subunit γ, the main component of the

central stalk. The asymmetric γ subunit, stabilized in the c-ring by subunits δ and ε, then causes conformational changes in the hexameric F_1 headpiece consisting of alternating α and β subunits. This creates different nucleotide binding affinities (open, loose and tight) at the β subunit catalytic site located at the interface with subunit α. In order for subunit γ to affect these structural changes, the catayltic headgroup must be held stationary by the peripheral stalk (subunits OSCP, b, d and F6) of the F_o domain. The OSCP subunit interacts with F_1 via subunit α, while subunit b anchors the peripheral stalk in the mt inner membrane by associating with subunit a of the proton pore. In addition, there are several supernumerary subunits that are membrane-bound F_o components that function in the dimerization of the complex or regulate its activity.

While the structure of this splendid enzymatic complex is highly conserved within opisthokants, the F_oF_1-ATPase of *Trypanosoma brucei* has several unique features. Until recently, every eukaryotic F_1 moiety described has always consisted of just α3, β3, δ, ε and γ. We have now demonstrated that *T. brucei* F_1 contains an additional subunit, named p18 after its relative mobility on an SDS-PAGE gel (Gahura, manuscript submitted). While the exact location of p18 is still unknown, its function is essential for the parasite as the F_1 subcomplex becomes destabilized when p18 is depleted. Furthermore, only subunits α, β, δ, ε, γ, a, c and OSCP have been identified from the 22 subunits composing the *T. brucei* F_oF_1-ATP synthase. Therefore, 14 subunits are either unique to trypanosomatids or they have diverged to the point where they are no longer recognized as homologues (Zíková et al, 2008). Since the *T. brucei* F_oF_1-ATPase is an attractive drug target, it is imperative that we elucidate the function of these hypothetical proteins. By mapping the interacting partners of known subunits with unknown subunits, we can begin to understand the structure of the complex and thus infer the function of these intriguing subunits.

1.4 Yeast Two-Hybrid Analysis

One of the classic methods used to determine protein-protein interactions is the yeast two-hybrid (Y2H) screen. This technique employs *Saccharomyces cerevisae* (baker's yeast), because with a doubling time of approximately 90min on rich medium, the growth and maintenance of these cells is simple, swift and economical. In addition, yeast are often used as a model organism for higher eukaryotes because they exhibit many of the same cellular processes that are observed in metozoans. Furthermore, the Y2H screen relies on the unique

attribute that yeast can live in either a haploid or diploid state. Diploids are formed when a haploid with the a mating type fuses with another haploid of the α mating type. These diploids then grow by mitosis until they encounter unfavourable growth conditions, often lacking nitrogen and carbon. At this point, the diploids undergo meiosis to form haploid spores that can survive the harsh environment until nutrients become abundant again (Lundblad and Struhl, 2008).

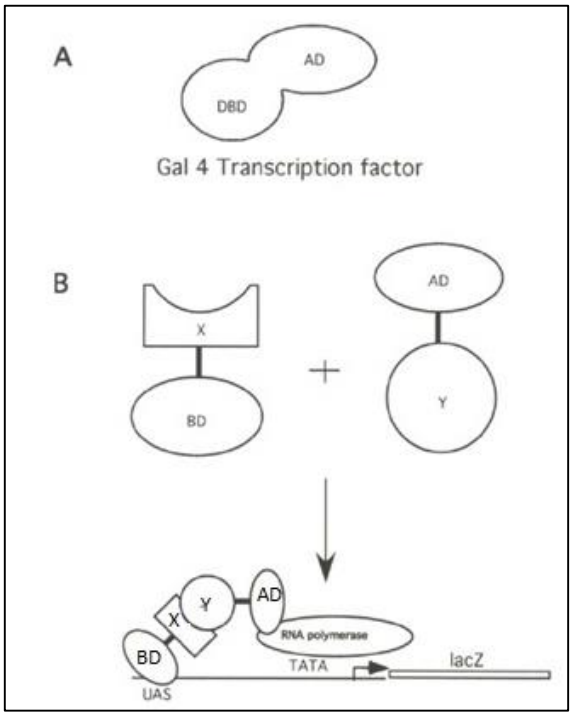

Figure 4: Transcriptional activation of the lacZ reporter in the Y2H assay. A) Domains of the wild type Gal4 transcription factor B) detection of protein-protein interactions. AD- transcriptional activation domain; BD- DNA binding domain; X, Y – proteins tested for interaction (Mordacq and Ellington, 2008).

The output of the Y2H screen is simply whether mutant yeast are able to grow on media lacking an essential amino acid or nucleobase. A positive result is achieved when the two proteins of interest, the bait and prey, interact and activate a reporter gene that complements the mutation and enables the synthesis of the missing nutrient (Figure 4). The basic biological process behind the Y2H screen involves transcription initiation, which requires a transcription factor to bind to a DNA enhancer element located upstream of the reporter gene. Since the GAL4 system has been well defined, this transcription factor is used in a Y2H screen to activate various reporter genes that have one of the GAL promoters (GAL1, GAL2 or GAL7). The GAL4 gene product binds the consensus sequence CGG-N11-CCG, so while each of the specific upstream activator sequences are different, they are all recognized by the GAL4

transcription factor. The GAL4 protein consists of two functional domains: the DNA binding motif that binds the promoter and the activation domain that interacts with the general transcription machinery. To determine if the bait and prey proteins interact, each has been fused to either the GAL4 binding domain or the activation domain. Therefore, this recombinant transcription factor is only able to recruit the RNA polymerase to the reporter gene when both domains are brought within close proximity via the bait/prey protein-protein interaction (Mordacq and Ellington, 2008).

There are a variety of genetically modified yeast available for Y2H screens. We have chosen the *S. cerevisiae* strain PJ69-4, created by James *et al.* in 1996. This host strain was established because it has a high transformation efficiency and contains three independent reporter genes that have been specifically modify to reduce the rate of false positives. First, a construct containing the HIS3 gene with the GAL1 promoter was introduced into the genome downstream of the LYS2 gene for tighter regulation. This decreases the need for the addition of the HIS3 enzymatic inhibitor, 3-aminotriazole (3-AT), which is employed during screens to eliminate false positives and identify only the strongest pairwise protein interactions. Another reporter includes the ectopic ADE2 gene, under the control of the GAL2 promoter, which replaces the endogeneous ADE2. When this mutant is grown on media with adenine, it accumulates an intermediate of adenine biosynthesis that turns red when oxidized. However, when the wild type (wt) ADE2 gene is activated, the colonies return to an opaque white color. Therefore, this reporter can act as an indicator of the strength of the bait/prey interaction because intermediate levels of the wt ADE2 gene will result in various shades of pink to white yeast. However, a more precise way to measure the binding strength of various bait/prey pairs utilizes a LAC2 reporter gene with the GAL7 promoter. LAC2 encodes the enzyme β-galactosidase, which is a glycoside hydrolase that can cleave the glycosidic bond of the X-gal substrate, producing a bright blue product that can be quantitated by a spectrophotometer (James et al, 1996).

Due to mutations introduced in key enzymes of several biochemical pathways, PJ69-4 is auxotrophic for tryptophan (Trp), leucine (Leu), histidine (His) and adenine (Ade), meaning they can't grow on media lacking any of these essential nutrients. Therefore, this strain can only grow on dropout media after it has acquired these wt genes from the inducible reporter constructs that also contain selectable markers (James et al, 1996).

As a positive control that demonstrates protein-protein interactions in the Y2H screen, we use two proteins known to interact very tightly. In order to commandeer a host cell for the purpose of replicating its viral genome, the Simian virus 40 (SV40) produces proteins that enable it to bypass cell cycle controls. One of these is the large T-antigen oncoprotein, which can transform a variety of cell lines and induce tumors in different animal models. This transforming activity of the large T-antigen is a consequence of its ability to bind and manipulate the functions of cell cycle regulatory proteins and certain tumor suppressors, like p53. Therefore, the amount of growth observed in the yeast diploid expressing the large T-antigen bait and the p53 prey constructs will demonstrate if the screen is working properly (Ali and DeCaprio, 2001) (Mordacq and Ellington, 2008).

2 Aims

- To clone the genes that encode selected subunits of the *T. brucei* F_oF_1-ATP synthase complex into the pGADT7 and pGBKT7 yeast two-hybrid (Y2H) plasmids
- To verify the protein expression from these plasmids in *Saccharomyces cerevisae*
- To perform a Y2H screen to determine the interacting partners of each protein, with the hope of deciphering the function of hypothetical proteins based on their location within the complex

3 Methods

3.1 Molecular cloning

Molecular cloning involves a series of techniques that insert a desired DNA fragment into a vehicle, usually referred to as a plasmid, which can self-replicate within a suitable host, often bacteria or yeast. These methods allow for the manipulation of genetic material to perform a desired task depending on the properties of the plasmid. The key features of this circular, non-chromosomal DNA molecule include an origin of replication, a selectable marker, suitable promoters and a multiple cloning site. The creation of this molecular tool involves two main processes that serve to amplify the recombinant DNA. The first involves an elegant enzymatic reaction called polymerase chain reaction (PCR), while the second step transforms the high-copy plasmids into an organism that can be quickly and cheaply grown into dense

populations. Once the recombinant plasmid is generated and transformed into the proper host organism, the resulting individual colonies must be screened to determine if they harbor the correct plasmid. This is achieved by positive antibiotic selection, plasmid DNA isolation, restriction enzyme digestions and plasmid sequencing. Specific details concerning each of these methods as they pertain to my thesis work will be discussed below.

3.1.1 Polymerase Chain Reaction

Almost all molecular cloning starts with PCR, which utilizes the ability of DNA polymerases to synthesize a new DNA strand complementary to the template provided. Since the polymerase can only elongate a DNA strand from a pre-exisiting nucleotide, short DNA oligonucleotides that specifically anneal to a chosen region of the template must be selected and synthesized. Multiple cycles of a three-step PCR, developed by Kary Mullis in 1985, then amplifies this targeted region exponentially by first denaturing the template at high temperatures, then lowering the temperature for primer annealing and finally allowing the DNA polymerase to elongate the replicated strand.. This method rapidly produces a specific DNA fragment that can then be inserted into a plasmid via various methods depending on the specific properties of the designed primers.

3.1.2 Primer design

Traditional methods of cloning often involve subcloning the PCR product into an intermediate vector that has unique properties to increase the efficiency of this step. After the plasmid is amplified in bacterial cultures and isolated, the DNA fragment of interest must be cut out from the intermediate vector using restriction digest enzymes and then ligated into the desired plasmid. This entire process requires at least 7 days of work. An alternative method takes advantage of homologous recombination to fuse together DNA fragments that share terminal end-homology in an *in vitro* cloning reaction. This greatly simplifies the cloning process and generates >90% positive clones (ref GENEART seamless cloning manual). We chose to utilize the GENEART® Seamless Cloning and Assembly Kit from Invitrogen. The included manual provides specific details and refers to a web-based tool that reduces the planning time involved in designing your PCR oligonucleotides. In short, the 5' end of the primers must contain the terminal 15 nucleotides of the linearized plasmid (in our case, we used

the Y2H plasmids pGADT7 or pGBKT7 from Clontech), which is followed by about 25 nucleotides that are unique for the open reading frame of the gene of interest. While designing these primers we attempted to balance the GC content around 50% as shown in Table I. Because of the restriction enzymes chosen to linearize the Y2H plasmids, it was necessary to insert a random nucleotide in the forward primer in order to keep the N-terminal tag in frame with our gene of interest. Therefore, adenine (highlighted A in Table I) was added immediately after the linearized plasmid sequence, which is underlined in Table I.

Table I: List of created primers.

Subunit	Oligo	Forward Primer	GC%	bp
α –mls cds pGADT7 AD	AZ0561	ACCAGATTACGCTCAAGCAGC AACGGCACCGGCAGGTG	61	38
α –mls cds pGBKT7 BD	AZ0563	AGAGGAGGACCTGCAAGCAGC AACGGCACCGGCAGGTG	66	38
α –mls N-term pGADT7 AD	AZ0561	ACCAGATTACGCTCAAGCAGC AACGGCACCGGCAGGTG	61	38
α –mls N-term pGBKT7 BD	AZ0563	AGAGGAGGACCTGCAAGCAGC AACGGCACCGGCAGGTG	66	38
α –mls C-term pGADT7 AD	AZ0567	ACCAGATTACGCTCAAGACAG CACGTTGGGTAAAG	49	35
α –mls C-term pGBKT7 BD	AZ0568	AGAGGAGGACCTGCAAGACAG CACGTTGGGTAAAG	54	35
δ –mls GADT7 AD	AZ0569	ACCAGATTACGCTCAACAATC GGCACCTCATGATC	49	35
δ –mls GBKT7 BD	AZ0571	AGAGGAGGACCTGCAACAATC GGCACCTCATGATC	54	35
ε –mls pGADT7 AD	AZ0573	ACCAGATTACGCTCAAAGCTC GTCATGGAGGGATCATG	50	38
ε –mls pGBKT7 BD	AZ0575	AGAGGAGGACCTGCAAAGCTC GTCATGGAGGGATCATG	55	38
p18 –mls pGADT7 AD	AZ0577	ACCAGATTACGCTCAAGCCGC CACCAGCGCCGCGAAGAA	62	39
p18 –mls pGBKT7 BD	AZ0579	AGAGGAGGACCTGCAAGCCGC CACCAGCGCCGCGAAGAA	67	39
OSCP cds pGADT7 AD	AZ0581	ACCAGATTACGCTCAAATGTTC CGCCGACTTTCTTC	47	36
OSCP cds pGBKT7 BD	AZ0583	AGAGGAGGACCTGCAAATGTT CCGCCGACTTTCTTC	53	36
Tb3320cds pGADT7 AD	AZ0599	ACCAGATTACGCTCAAATGTG GAGGTTTGGTG	47	32
Tb3320cds pGBKT7 BD	AZ0601	AGAGGAGGACCTGCAAATGTG GAGGTTTGGTG	53	32
Tb520 cds pGADT7 AD	AZ0603	ACCAGATTACGCTCAAATGCA GGGCAGTTGGTCGG	54	35
Tb520cds pGBKT7 BD	AZ0605	AGAGGAGGACCTGCAAATGCA GGGCAGTTGGTCGG	60	35
Tb2930cds pGADT7 AD	AZ0607	ACCAGATTACGCTCAAATGCG CCGTGTATCTTC	49	33
Tb2930cds pGBKT7 BD	AZ0609	AGAGGAGGACCTGCAAATGCG CCGTGTATCTTC	55	33
Tb6250 cds pGADT7 AD	AZ0611	ACCAGATTACGCTCAAATGTC GAAGCAACTAACG	44	34
Tb6250 cds pGBKT7 BD	AZ0613	AGAGGAGGACCTGCAAATGTC GAAGCAACTAACG	50	34
Tb1270 cds pGADT7 AD	AZ0615	ACCAGATTACGCTCAAATGAC AAAATATGAGCTG	38	34
Tb1270 cds pGBKT7 BD	AZ0617	AGAGGAGGACCTGCAAATGAC AAAATATGAGCTG	44	34
pGBKT7 for sequencing	AZ0558	GCCTCTAACATTGAGACAGC	50	20
pGADT7 for sequencing	AZ0559	GCGTTTGGAATCACTACAGG	50	20

Subunit	Oligo	Reverse Primer	GC%	bp
α –mls cds pGADT7 AD	AZ0562	CGATTCATCTGCAGCTCAC ACTGCCCGCTTACAAGAG	54	37
α –mls cds pGBKT7 BD	AZ0564	TATGCGGCCGCTGCATCAC ACTGCCCGCTTACAAGAG	60	37
α –mls N-term pGADT7 AD	AZ0565	CGATTCATCTGCAGCTTAG AGCCCCACCGGAACCTCGT	58	38
α –mls N-term pGBKT7 BD	AZ0566	TATGCGGCCGCTGCATTAG AGCCCCACCGGAACCTCGT	63	38
α –mls C-term pGADT7 AD	AZ0562	CGATTCATCTGCAGCTCAC ACTGCCCGCTTACAAGAG	54	37
α –mls C-term pGBKT7 BD	AZ0564	TATGCGGCCGCTGCATCAC ACTGCCCGCTTACAAGAG	60	37
δ –mls GADT7 AD	AZ0570	CGATTCATCTGCAGCTTAG TGGTGTTTTAGGGCAG	49	35
δ –mls GBKT7 BD	AZ0572	TATGCGGCCGCTGCATTAG TGGTGTTTTAGGGCAG	54	35
ε –mls pGADT7 AD	AZ0574	CGATTCATCTGCAGCTCAG TAATCCTTGGTGTGGATGG G	51	39
ε –mls pGBKT7 BD	AZ0576	TATGCGGCCGCTGCATCAG TAATCCTTGGTGTGGATGG G	56	39
p18 –mls pGADT7 AD	AZ0578	CGATTCATCTGCAGCTTAC TCTACCTTTACATC	42	33
p18 –mls pGBKT7 BD	AZ0580	TATGCGGCCGCTGCATTAC TCTACCTTTACATC	49	33
OSCP cds pGADT7 AD	AZ0582	CGATTCATCTGCAGCTCAA ACGCCATATTTCGCCTCC	51	37
OSCP cds pGBKT7 BD	AZ0584	TATGCGGCCGCTGCATCAA ACGCCATATTTCGCCTCC	57	37
Tb3320cds pGADT7 AD	AZ0600	CGATTCATCTGCAGCTCAC TCTTCAATTTCATTCATAC GC	43	40
Tb3320cds pGBKT7 BD	AZ0602	TATGCGGCCGCTGCATCAC TCTTCAATTTCATTCATAC GC	48	40
Tb520 cds pGADT7 AD	AZ0604	CGATTCATCTGCAGCTTAA GCTGTGTGTCGG	52	31
Tb520cds pGBKT7 BD	AZ0606	TATGCGGCCGCTGCATTAA GCTGTGTGTCGG	58	31
Tb2930cds pGADT7 AD	AZ0608	CGATTCATCTGCAGCTTAG TGATGGGCCTTTTTC	47	34
Tb2930cds pGBKT7 BD	AZ0610	TATGCGGCCGCTGCATTAG TGATGGGCCTTTTTC	53	34
Tb6250 cds pGADT7 AD	AZ0612	CGATTCATCTGCAGCTCAA AGTGCAGAAACCAC	49	33
Tb6250 cds pGBKT7 BD	AZ0614	TATGCGGCCGCTGCATCAA AGTGCAGAAACCAC	55	33
Tb1270 cds pGADT7 AD	AZ0616	CGATTCATCTGCAGCTTAA TAATTGTAGATGGCC	41	34
Tb1270 cds pGBKT7 BD	AZ0618	TATGCGGCCGCTGCATTAA TAATTGTAGATGGCC	51	39
pGBKT7 for sequencing	AZ0235	TTTTCGTTTTAAAACCTAA GAGTC	30	24
pGADT7 for sequencing	AZ0560	AGATCAGAGGTTACATGGC C	50	20

3.1.3 PCR

Since it is important that the amplified open reading frame of each gene of interest doesn't contain any errors incorporated by the DNA polymerase that would result in a nonsense mutation, we used KOD HOT START DNA polymerase (Novagen). This proofreading enzyme is mixed with two specific monoclonal antibodies that inhibit the activities of both the polymerase and the 3' → 5' exonuclease at ambient temperature. This decreases mispriming and primer degradation that can occur during setup and initial temperature increase. This results in high yields of very accurate blunt end PCR products.

A standard reaction setup consists of the following reagents provided by the polymerase kit: 10X Buffer for KOD Hot Start DNA Polymerase, MgSO4 (1.5mM), 100µM dNTPs (5µl) and KOD Hot Start DNA Polymerase (1µl). To this mixture was added 200ng of template genomic *T. brucei* DNA, 1.5µl of both the forward and reverse primers (final concentration is 300nM) and finally the appropriate volume of PCR grade water to reach a final reaction volume of 50µl. A typical PCR program consisted of a one-time polymerase activation step (95°C for 2min), then 30 cycles of a three step procedure that included a denaturation step (95°C for 20s) an annealing step (55°C for 10s) and the extension step (70°C for 20s). This extension time was dependent on the length of the amplified product, where amplicons smaller than 500bp were incubated for 10s, amplicons within the range of 500-1000bp were held at 70°C for 15s, while PCR products between 1000 to 3000bp were extended for 20s for every kb. The annealing temperature for each primer pair needed to be optimized based on the predicted melting temperature, but usually this value fell between 50°C and 55°C.

3.1.4 Gel electrophoresis

To determine the yield and specificity of each PCR, the amplicons were resolved on an agarose gel (Figure 5). Basically, gel electrophoresis is a technique that separates charged molecules based on their size as they move through a matrix consisting of roughly equivalent sized pores. Agarose and acrylamide are routinely used to resolve nucleotides based on their small pore sizes, with polyacrylamide gels usually being reserved for molecules typically less than 1000bp. The percentage of these reagents also proportionally affects the pore size, where more concentrated gels will have smaller pore sizes and resolve smaller fragments more

efficiently. Before DNA samples can be analyzed, they need to be mixed with a 10x loading dye that consists of glycerol to help the molecules sink to the bottom of a well in a horizontal gel and one or two dyes that can track the progress of the migrating DNA. Once the samples are loaded into the gel, it is important to run them in the direction of the cathode since the DNA is negatively charged due to the phosphate backbone.

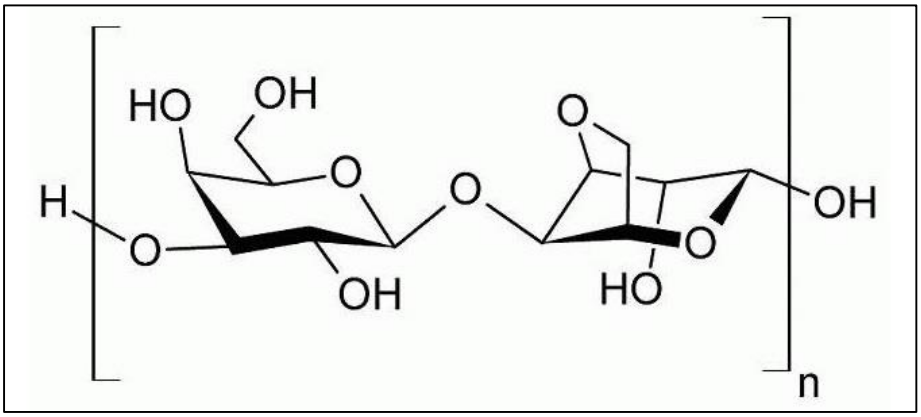

Figure 5: Agarose chemical structure (www.thermofisher.com).

Since most of our PCR amplicons ranged between 230 and 1714bp, we used 1% agarose gels (1g agarose in 100ml of TAE buffer). To dissolve the powdered agarose, the mixture was heated in a microwave until it began to gently boil and then liquid solution was cooled under running tap water to about 50°C (measured as when you can comfortably touch the glass container with your bare hand). At this cooler temperature, 1.5µl of heat-labile ethidium bromide could be added. Ethidium bromide is a fluorescent dye that allows the DNA to be visualized when placed on UV light source because it is an intercalating agent that interacts between the two strands of DNA. The liquid agarose was gently swirled to evenly distribute the ethidium bromide before it was carefully poured into a gel tray containing an appropriately sized comb that creates wells to load the DNA samples into. The gel was allowed to solidify (40min) before it was placed in an electrophoresis tank, covered in TAE running buffer consisting of the necessary electrolytes and the comb was removed.

To accurately determine the size of the amplicons created by PCR, a DNA marker (1kb DNA Ladder Plus, Invitrogen) was loaded into a well. Since the yield & specificity of each PCR is unknown, all of the reaction volume for each sample was loaded into the remaining wells and the gel was run for approximately 45min at 90V. Afterwards, the DNA products were visualized under UV light and a picture was taken to document the results of the PCR. Resolving the PCR products on an agarose gel have the added benefit of separating the desired

amplicon from the DNA template and any unincorporated dNTP's, but now the correct sized DNA band needs to be extracted from the agarose before it can be inserted into the vector. For this purpose, we used the GenElute™ Gel Extraction Kit (Sigma-Aldrich), which contain a reagent with a pH indicator to melt the agarose, spin columns with a silica membrane to bind the DNA and an elution buffer. While this process is very quick, it also usually results in a 50% loss of linear DNA molecules between 100bp – 10kb. Due to these significant losses, it is important to determine the DNA concentration and purity for downstream applications. Therefore, we used a NanoDrop™ (Thermo Scientific), which is a spectrophotometer that can calculate the concentration of very small samples (0.5µl-2µl) based on their UV absorption at 260nm. Furthermore, you can determine the purity because DNA has a maximum absorption at 260nm whereas protein has a maximum at 280, thus a A_{260}/A_{280} ratio below 1.8-2.0 indicates that the nucleic acid preparation has protein contamination.

3.1.5 Engineering DNA molecules by homologous recombination

We engineered specialized plasmids (pGADT7 & pGBKT7) that could be replicated in both bacteria and yeast by utilizing linearized DNA fragments with terminal end-homology. Both vectors are generally very similar but they have some important distinctions. Specifically, the pGBKT7 plasmid contains the DNA binding domain (BD) of the *S. cerevisiae* GAL4 transcription factor, while the pGADT7 vector possesses the activation domain (AD) from the same protein. By introducing both vectors into a single yeast, we can screen for protein interactions by determining if a prey protein of interest fused to the GAL4 AD can bind a particular bait protein fused to the DNA BD and thus bridge the two halves of the transcription factor and induce transcription of a reporter gene with the GAL4 promoter. Other differences between these two plasmids include their bacterial antibiotic resistance marker (pGADT7 – ampicillin; pGBKT7 - kanamycin), an auxotrophic amino acid marker (pGADT7 – leucine, pGBKT7 – tryptophan) and the SV40 nuclear localization targeting sequence fused to the GAL4 AD in pGADT7. Finally, a shared unique attribute of these plasmids is that they incorporate both a bacterial and yeast origin of replication so they can be propagated in both organisms.

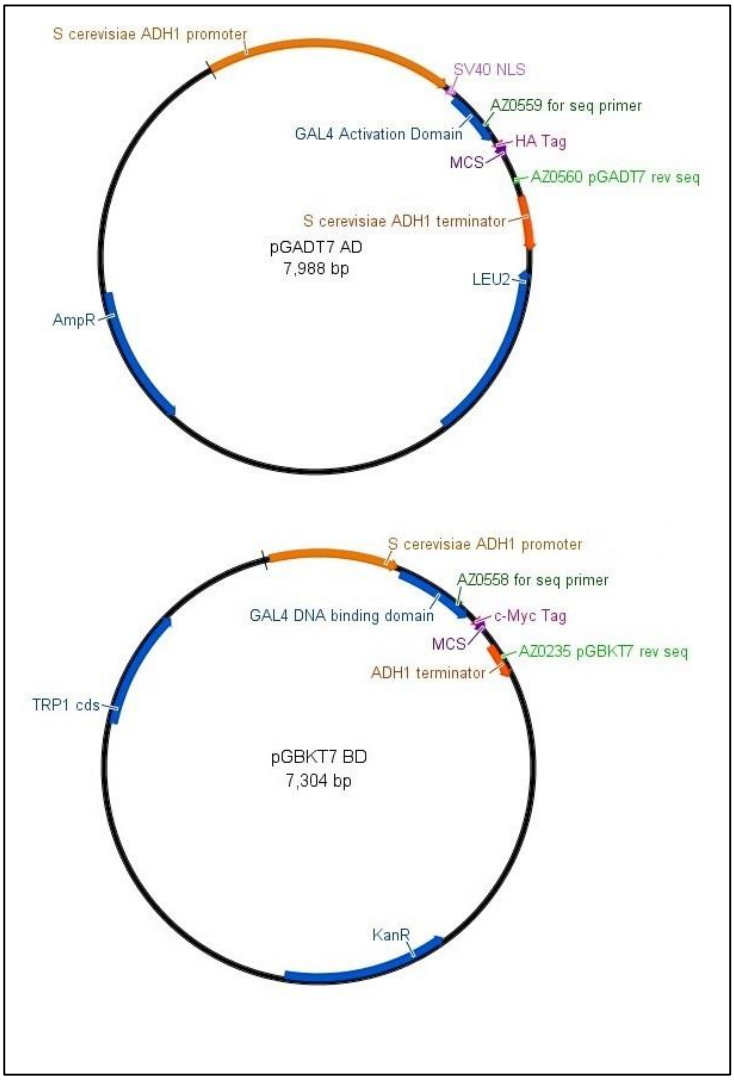

Figure 6: pGADT7 and pGBKT7 plasmid maps.

The GeneART® Seamless Cloning and Assembly kit (Invitrogen) is based on homologous recombination and can be used to assemble up to four DNA fragments in one reaction. The user's manual contained the following equation that was used to calculate an insert to vector ratio of 2:1 for the highest assembly efficiency.

$$x\ ng\ insert = \frac{(2)(bp\ insert)(100\ ng\ linearized\ vector)}{(bp\ vector)}$$

The insert was then diluted into a working stock that contained 10x the calculated final concentration required for the enzymatic reaction. Since the total volume of this reaction was

brought up to 20µl with deionized water, 2µl of 10x insert was then added together with 2µl of a 10x working stock of the linearized plasmid, 4µl of the 5x reaction buffer and 2µl of the 10x enzyme mix. These reagents were then gently mixed by flicking the tube before it was incubated at room temperature for 30min.

3.1.6 Bacterial Transformation

Transformation is a commonly used method to introduce recombinant DNA into bacterial cells. Since the most popular prokaryote used in the lab is gram-negative *Escherichia coli*, which are not capable of transformation under natural conditions, they must be made competent either chemically or by electric shock. We used One Shot® TOP10 Chemically Competent *E. coli* that were made competent by suspending them in a buffer with divalent cations, usually calcium or manganese chloride. These positively charged ions help enable the DNA to navigate through the connections that form between the outer and inner membrane by neutralizing the negatively charged phosphate backbone of the recombinant DNA and coating the outer membrane of the bacteria to mask the negative charges of the phospholipids. Since these cations increase the fluidity of the normally rigid cell wall, it is important to gently thaw them on ice for 10-20min and avoid unnecessary pipetting or vortexing. Then 8µl of the seamless cloning assembly reaction are added and mixed with the bacteria by inverting the tube several times. The DNA was incubated with the cells for 20-30min on ice to stabilize the bacterial membranes and provide ample opportunity for the cations to interact with all the negatively charged molecules, allowing the plasmid DNA to come in close proximity to the prokaryote. The competent bacteria were then shocked with heat by placing them at 42°C for 30s. This further alters the fluidity of the membranes and allows the exogenous DNA to enter the cell at a more efficient rate, possibly due to surface invagination. Afterwards, the cells were incubated on ice for 2min, which is speculated to reduce the thermal motion of the recombinant DNA and promote it to enter the cell. Since actively growing cells in early log phase are more susceptible to transformation, 250µl of nutritionally rich SOC medium was added to the cells before placing them at 37°C for 1hr in a shaking incubator (200rpm). The cells were then diluted with another 700µl of SOC media before 100µl of cells were spread onto prepared LB-plates containing the appropriate antibiotics. Plates with bacterial cells were finally incubated overnight at 37°C to generate individual colonies.

3.1.7 Isolation of plasmid DNA

Since each bacterial colony is derived from an individual cell, each colony should contain the same plasmid DNA. Depending on the success of the transformation process, 3-5 colonies are often screened to determine if they contain the expected recombinant DNA molecule assembled with our gene of interest. Therefore, the plasmid DNA first needs to be isolated from these colonies. First the plasmid DNA is amplified by growing a single colony in 3ml LB medium containing the appropriate selectable antibiotic at 37°C in a shaking incubator overnight. It is important to harvest the cells the next morning when the bacteria are still in log phase growth, otherwise a very dense culture will start to lose their plasmid as the excreted antibiotic resistance protein depletes all of the selectable drug. In successive spins, all 3 ml of the culture were harvested at 12,000 x g for 1min. An aspirator was used to remove any traces of liquid media from the cell pellets, which were then thoroughly resuspended in Resuspension Solution provided by the GenElute™ HP Plasmid Miniprep Kit. The bacteria were then subjected to an alkaline-SDS lysis step, while care was taken to avoid harsh mixing that could shear the chromosomal DNA and result in contamination of the purified plasmid. This lysis step was neutralized within 5min to avoid permanently denaturing the supercoiled DNA plasmid. In the presence of high salts, this DNA could then be specifically bound to a silica spin column, which provides a convenient platform to quickly remove trace contaminants such as endotoxins and residual salts. Finally the plasmid DNA was eluted with 100µl of the Elution solution (10mM Tris-HCl, pH 8.5) and the DNA concentration was measured via NanoDrop™.

3.1.8 Plasmid screening by restriction digest analysis

The purified plasmid DNA can now be screened by restriction digest to determine if it contains our insert and produces the expected banding pattern. Restriction enzymes are often referred as “molecular scissors”. They are naturally found in bacteria or archaea cells as a defence mechanism against viruses. They recognize specific palindromic DNA sequences that are usually between 4 to 8bp long. The restriction enzyme then cuts both DNA strands between the sugar-phosphate backbones. Depending on the location of this incision, it can produce either blunt end DNA fragments (same site on both strands) or ones with overhangs (two different sites) that are often referred to as sticky ends, which are useful in directional molecular cloning.

The recombinant plasmids were also generated *in silico* using the Geneious software, which has numerous molecular biology features that can quickly identify unique restriction sites that can be used to verify the plasmids. Ideally, two compatible restrictions enzymes present just once in the sequence are chosen so that one cuts within the backbone vector and the other within the insert, producing two bands that are easy to resolve and accurately identify on an agarose gel.

To easily visualize the banding pattern on an agarose gel, about 500ng of plasmid DNA is required for the digest (Table II). Since the final volume of the reaction is 20µl, 2µl of 10x buffer was added. Also, 0.5µl of each restriction enzyme was included so that the glycerol content from the enzyme storage buffer doesn't cause the total glycerol in the reaction to exceed 10%, which prohibits enzyme activity. Once the volume of 500ng DNA was calculated for each plasmid, the remaining volume of the digestion reaction was supplied with MilliQ water. The restriction digest was then placed at 37°C for 3hrs. To ensure that the DNA was completely cut, another 0.5µl of each restriction enzyme was added and the whole digest was incubated for another hour. After the digest was completed, 5x loading dye was added to the samples before they were resolved on a 1% agarose gel and visualized under a UV light

Table II: Example of a restriction digest reaction.

Plasmid	500ng	Enzyme	Buffer	Water	Predicted bands with inserts (bp)	Predicted band empty (bp)
OSCP AD	4 µl	BglII, SalI	0	13µl	8089, 612	7992
α N AD	4.2 µl	HindIII	R	13.3µl	7192, 1125	uncut

3.1.9 Sequencing

While the restriction digests demonstrate if the insert is present in the plasmid, sequencing is required to confirm that our gene of interest doesn't contain any missense or nonsense mutations. Therefore, one clone for each plasmid was sequenced by the SEQme sequencing and qPCR company (website) using primers that enabled us to determine if the N-terminal Gal4 activation or binding domain, the epitope tag and the gene of interest are all in frame and without any deleterious mutations.

SEQme requires the following sequencing reaction components to be supplied in a thin-walled pcr tube: 2.5µl primer (10µM), up to 7.5µl of DNA (500ng), and

the remaining volume up to 10 µl provided by sterile water. The abi sequencing chromatogram files were usually accessible online in 3-4 days. Once imported into the Geneious program, any problematic sequencing ends could be trimmed and then aligned with the *in silico* files created for each plasmid. Every single nucleotide polymorphism (SNP) identified in the coding sequence was recorded for each plasmid and determined if it would be detrimental to the synthesized protein.

3.2 Yeast

With the plasmids assembled and verified, we will now switch organisms and discuss the yeast methods employed to perform the actual Y2H experiments. First, we need to transform *Saccharomyces cerevisae* with the Y2H plasmids and then verify the protein expression from these transformants using a western blot analysis (Chapter 3.3).

3.2.1 Media

Compared to using *E. coli* for the previous molecular cloning steps, *Saccharomyces cerevisae* prefers to grow at 30°C, but even at this temperature they grow more slowly, typically requiring three days of growth to generate a 2-3mm sized colony on an agar plate. Because of this slower growth, 0.1% ampicillin was added to all media to limit fast growing bacterial contamination. Enriched yeast media (1% yeast extract, 2% tryptone – source of amino acids from the tryptic digest of casein, 2% glucose), referred to as YPD (yeast, peptone, dextrose), allows for the growth of *S. cerevisae* without any nutritional selection. Nutritional selection refers to the selection of yeast containing a plasmid that is capable of synthesizing an enzyme of a mutated biochemical pathway that allows the yeast to survive on media lacking that essential nutrient. However, the yeast strains used in our study are adenine auxotrophs and also require the addition of 30mg/L adenine hemisulfate, resulting in media called YPDA. This liquid media is used to grow dense yeast cultures required for a variety of experiments, but it is best to seed these cultures with cells from a single colony that arise from yeast agar plates. These plates are created by simply adding 2% bacteriological agar and 0.01% NaOH to YPDA media. The NaOH is added to maintain the integrity of the agar when it is autoclaved, otherwise it can be degraded, leading to soft plates. All liquid media were filter sterilized, while the media for agar plates were autoclaved before being poured into Petri dishes under a flame.

To test for protein-protein interactions, the Y2H analysis determines if yeast are able to activate the transcription of a gene that can help synthesize an essential nutrient that is absent from the agar plates. Therefore, synthetic defined (SD) media, which includes a carbon source like glucose but lacks most of the essential amino acids, is used when nutritional selection is required. With this very basic media composition, you can adapt the media for each yeast strain by adding back the remaining essential components except for the ones that are used as the selectable markers for the Y2H ectopic reporters. Since our yeast utilize adenine and histidine as Y2H reporters, these auxotrophic yeast must have a final concentration of 30mg/L adenine hemisulfate and 20mg/L L-histidine in their agar plates when they are not being used to screen for protein-protein interactions. In addition, the yeast are also auxotrophic for either leucine (a mating type) or tryptophan (α mating type) and require a final concentration of either 100mg/L L-leucine or 40mg/L L-tryptophan in their media when they don't contain the pGADT7 or pGBKT7 vector, respectively. Therefore, agar plates to select for yeast containing the pGADT7 vector would consist of ingredients with the following final concentrations: 6.7g/L yeast nitrogen base without amino acids, 1x dropout media (10x composition listed in Table III), 30mg/L adenine hemisulfate, 20mg/L L-histidine and 40mg/L L-tryptophan. Once all of these components are dissolved, the solution is filter sterilized and stored at 4°C up to 1 year (Clontech, 2009).

Table III: 10x Dropout media.

Reagent	500ml
Arginine HCl	100mg
Isoleucine	150mg
Lysine HCl	150mg
Methionine	100mg
Phenylalanine	250mg
Threonine	1g
Tyrosine	150mg
Uracil	100mg
Valine	750mg
Q.S. dH2O	500ml

3.2.2 Transformations

Yeast transformation is a process that allows for the genetic modification of yeast cells by encouraging them to take up plasmid DNA, usually isolated from bacteria. We employed the classical lithium acetate approach that was adapted from Clontech's Yeast Protocols Handbook (Clontech, 2009). At first, the two yeast strains were streaked out on a plate without any nutritional selection and allowed to grow for 3 days. Several colonies of the yeast strains PJ69-4a or PJ69-4α were inocculated in 50ml of the rich YPDA (Yeast Peptone Dextrose + Adenine) media and incubated at 30°C for 16-18hrs at 250rpm. 30ml of the overnight culture was then transferred into a 2L Erlenmeyer flask containing 300ml of YPDA. The OD_{600} of this diluted culture was measured on the Eppendorf BioPhotomere. To reach the optimal density for a successful yeast transformation in a defined time period, the OD_{600} had to be around 0.2-03. If it was less, then a little more of the overnight culture was added. This 2L Erlenmeyer flask was incubated at 30°C for 3hrs on a shaking (230rpm) platform. When the OD_{600} reached 0.4-0.6, the cells were centrifuged in conical falcons at 1,000xg for 5min at room temperature. Using aseptic technique under a flame, the supernatant was discarded and the pellet was resuspended in sterile distilled water. The cells were centrifuged as before and the supernatant was discarded. Finally, the yeast cell pellet was resuspended in 1.5ml of freshly prepared and sterile 100mM lithium acetate/TE, pH 7.5 (Table IV). The final pH of this solution was adjusted with 1% acetic acid before it was autoclaved. It is also highly recommended to use fresh lithium acetate that is not more than 1 month old.

Table IV: 100mM Lithium Acetate/TE, pH 7.5.

Reagent	[Final]	50ml
dH2O		45ml
Lithium Acetate	100mM	0.510g
Tris-HCl, pH 8.0	10mM	500µl
EDTA, pH 8.8	1mM	100µl

Transformations of each selected gene consisted of 100ng of plasmid DNA and 100µg of carrier DNA in a sterile Eppendorf tube. While the exact mechanism of the carrier DNA is still poorly understood, it is thought to either protect the plasmid DNA from nucleases or the yeast cell wall, which due to its chemical structure tends to bind DNA. After the DNA was mixed by pipetting up and down, 100µl of competent yeast cells were added and vortexed for 10s. Then 600µl of 40% PEG 3350/100mM lithium acetate/TE, pH 7.5 (Table V) was added to the solution, which was vortexed again for 10s. The PEG solution, which helps to hold the cell pores open and facilitate more efficient DNA uptake, was adjusted to pH 7.5 with 1% acetic acid. This was done before the addition of viscous PEG, which requires 50°C to dissolve into solution. Finally the solution was made sterile by autoclaving.

Table V: 40% PEG 3350/100mM Lithium Acetate/TE, pH 7.5.

Reagent	[Final]	50ml
dH2O		30ml
Lithium Acetate	100mM	0.510g
Tris-HCl, pH 7.5	10mM	500µl
EDTA, pH8	1mM	100µl
PEG 3350	40%	20g

To allow time for the yeast to become coated in DNA, the Eppendorf tubes were incubated at 30°C for 3min at 200rpm. Afterwards, 70µl of dimethyl sulfoxide was added and mixed by gentle inversion before the cells were heat shocked for 15min at 42°C in a thermoblock. The cells were immediately chilled on ice for 12min and then centrifuged for 5s at 14,000rpm at room temperature. The supernatant was removed and the cells were resuspended in 0.5ml of TE buffer, pH 7.5 (Table VI). The pH of the TE was adjusted with HCl before it was autoclaved. 50µl of resuspended cells were diluted in 150µl of sterile TE buffer and then spread onto prepared selection plates lacking either leucine for PJ69-4a strain

or tryptophan for the PJ69-4α strain. The plates are incubated at 30°C for 3 days. The first colonies were visible after the second day of incubation.

Table VI: TE buffer, pH 7.5.

Reagent	[Final]	50ml
dH2O		49.4ml
Tris-HCl, pH 8.0	10mM	500µl
EDTA, pH 8.0	1mM	100µl

3.3 Protein expression

We performed a western blot analysis to determine the protein expression levels of the tagged protein in yeast transformed with the Y2H vectors. Since yeast possess a rigid cell wall compared to bacteria, a mild alkaline method was employed to prepare protein samples for SDS PAGE analysis.

3.3.1 Protein isolation

A single yeast colony was picked from an SD-plate with a biological loop and grown in 5 ml YPDA media at 30°C for 22hrs at 200rpm. In order to accurately measure the OD_{600} with a biophotometer, 40µl of the yeast culture was diluted in 160µl YPDA. This diluted solution was placed into a microcuvette and the OD_{600} recorded. It was determined that the ideal number of cells to harvest for this protein isolation procedure was 1ml of a culture with an OD_{600} of 0.5. These cells were spun down at 6000rpm for 5min at room temperature. If the density of the culture varied between an OD_{600} of 0.2 – 1.0, the amount of harvested cells was calculated based on the previous reference. For instance, if the OD_{600} was 0.2, then 2.5 ml culture was spun down. The media was then decanted and the yeast resuspended in 100µl distilled H_2O. Next, 100µl 0.2 M NaOH was added and the reagents were mixed by flicking the tube well. was After the cells were incubated for 5min at room temperature, they were pelleted by spinning them for 5min at 6000rpm at room temperature. All of the liquid was carefully decanted, leaving behind no traces of it behind. The cell pellet was then resuspended in 50µl 1xSDS sample buffer (Table VII) by pipetting up and down. The sample buffer contains SDS, an anionic detergent with a negatively charged head group that helps denature proteins. The buffer also contains β-mercaptoethanol to further reduce any disulfide bonds

between the proteins, while the dense glycerol helps with sample loading. Finally, the bromophenol blue helps track the migration of the proteins on the gel (OpenWetWare contributors, 2013). These protein samples were then boiled at 97°C and pelleted again at 6000 rpm for 5min at room temperature. The supernatant was extracted to a new Eppendorf tube and stored at -20°C.

Table VII: 1x SDS sample buffer.

Reagent	[Final]	14ml
Tris pH 6.8	60mM	840µl
Glycerol	5%	875µl
SDS	2%	1.4ml
β-mercaptoethanol	4%	560µl
Bromophenol Blue	0.0025%	17.5µl
Qs with MilliQ		14ml

3.3.2 Western blotting

Western blot analysis is a method that utilizes antibodies to detect specific proteins that are resolved on a SDS-PAGE gel and transferred to a PVDF (polyvinylidene difluoride) membrane for stability. (YANG and Mahmood, 2012). The primary antibodies specifically recognize the N-terminal epitope tags that are fused to the bait or prey. Each Y2H plasmid has a different epitope tag: pGADT7 contains an HA tag that is recognized by a mouse monoclonal antibody and pGBKT7 contains a c-Myc tag that reacts with a rabbit polyclonal antibody. The secondary antibody is typically produced in goats and recognizes either mouse or rabbit IgG. This antibody is conjugated with horseradish peroxidase, an enzyme that produces chemiluminescence when the ECL substrate is applied, serving to amplify the protein signal (Thermo Fisher Scientific, 2013).

3.3.2.1 SDS – PAGE

SDS-PAGE (sodium dodecyl sulphate-polyacrylamide gel electrophoresis) is a method that resolves charged proteins based on their molecular weight as they migrate through a gel in an electric field. The secondary and tertiary structures of the proteins are abolished as the proteins are coated in the negatively charged detergent, SDS, which also causes the proteins to migrate towards the cathode (Oswald, 2008). The gel itself is composed of the stacking

(Table VIII) and resolving (or running) gels (Table IX), which are defined by the percentage of acrylamide and the pH. We used a 12% resolving gel because it best resolves proteins with molecular weights between 10-200 kDa (Sino Biological, 2013).

Table VIII: 5% Stacking gel recipe.

Reagents	5 ml (2 gels)
H_2O	3.4 ml
30% acrylamide mix	840 µl
1M Tris 6.8	630 µl
10% SDS	50 µl
10% APS	50 µl
TEMED	10 µl

Table IX: 12% Resolving gel recipe.

Reagents	12 ml (2 gels)
H_2O	4 ml
30% acrylamide mix	4.8 ml
1.5M Tris 8.8	3 ml
10% SDS	120µl
10% APS	120µl
TEMED	12µl

Figure 7 (Oswald, 2008) is a schematic to illustrate the pH differences in the gels used to resolve proteins. The stacking gel (the upper one) has a pH of 6.8, while the resolving gel has a pH 8.8. The different pH values cause the negatively charged glycine ions present in the buffer (1xSDS running buffer) to travel from the higher pH in the resolving gel towards the stacking gel with the lower pH. As glycine transitions into the stacking gel, it loses its charge and starts to move slower to the anode. Meanwhile, Cl^- ions from Tris-HCl, travel very quickly in the electric field and create a steep voltage gradient, which drags glycine with them. The protein sample, possessing its own electric mobility, is now pulled along between these two lines (Oswald, 2008). In order to determine the mobility of proteins migrating through the gel, 4µl prestained protein ladder was used as a marker (PageRuler™ Prestained Protein Ladder, Fermentas – Life Sciences). The other wells were loaded with 20µl of protein samples. The gels were run at 90V until the proteins have left the stacking gel, at which point the voltage was increased to 110V for about 90min.

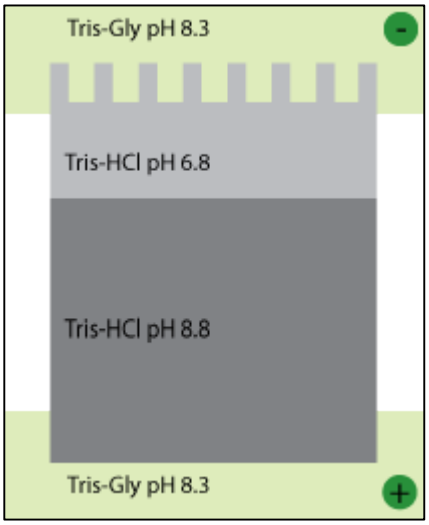

Figure 7: SDS-PAGE gel scheme.

3.3.2.2 Western Blot Analysis

Figure 8 (Amino acid analysis, 2003) demonstrates the arrangement for transferring the proteins onto a durable membrane that can then be analyzed with antibodies. The gel with the resolved proteins is inserted between the activated PVDF membrane and the filter paper. In order to activate the membrane, it was sequentially soaked for 30s in methanol, then in distilled H_2O for 2min and finally in 1 x Transfer buffer for 5min. The most important thing is to ensure that there are no air bubbles between any of the layers, especially between the gel and membrane. Since the proteins are still negatively charged, they will continue to migrate towards the anode. This transfer procedure is performed for 90min at 90V in 1x Tris-glycine transfer buffer containing 20% methanol.

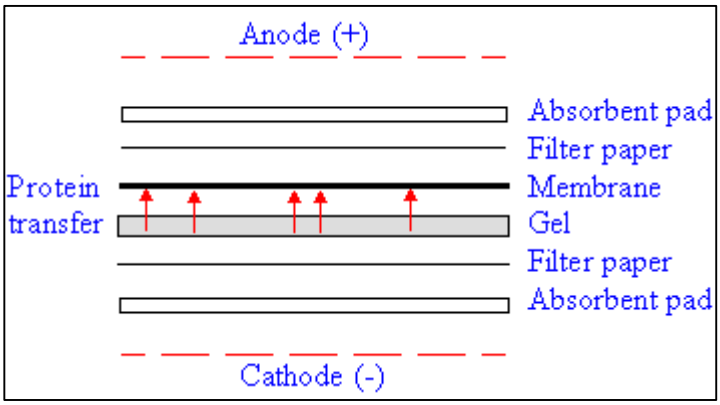

Figure 8: Schematic of Electroblotting.

The membrane coated with proteins was inserted into a sterile falcon tube with 5% milk dissolved in PBS-T (phosphate buffered saline with Tween). It was kept rolling for about 10min, before it was stored at 4°C overnight. Proteins from the milk, primarily casein, act as a blocking reagent that limits nonspecific interactions between random proteins and the antibodies raised against a specific antigen. The falcon tubes were rolled again for about 30min the next morning before the blocking agent was discarded and 5ml of fresh 5% milk was added to the falcon. The primary antibody was then diluted 1:1000 directly in the milk and incubated on a roller for 1hr. Afterwards, it was quickly rinsed with 25ml of PBS-T to remove the primary antibody. Then it was further washed with 25ml of PBS-T, first for 15min on the roller and then three times for 5min. Next, the secondary antibody was diluted 1:2000 in 5ml of 5% milk and incubated with the membrane in the falcon tube on the roller for 1hr. Afterwards, the same washes were performed as before. The last step was to mix the Clarity ECL reagents (Pierce™ ECL Western Blotting Substrate) in a 1:1 ratio and incubate them with the membrane for 1 minute at room temperature. The proteins were then visualized on the Chemidoc MP Imaging System (BioRad).

3.4 Yeast two hybrid analysis

Now that the protein expression has been determined, it was possible to finally perform the Y2H screen. Since we had created haploid yeast with either the pGADT7 plasmid containing the Gal4 activation domain or the pGBKT7 vector with the DNA binding domain, we first needed to create diploids that contained both plasmids. The interactions between bait and prey proteins could then be analyzed in these diploids when they were grown on media lacking selectable nutrients.

3.4.1 Mating of yeast

The two yeast strains (PJ69-4a and PJ69-4α) were mated in 96 well microtiter plates that contained 160µl YPDA. A single colony was picked with a biological loop from SD-Leucine (SD-Leu) or SD-Tryptophan (SD-Trp) plates for every gene and resuspended in an Eppendorf tube with 1 ml of YPDA by vortexing for 45s. 20µl of this cell suspension was then immediately pipetted into one row or column designated for the gene, as depicted in Table X. This procedure was repeated for every gene in both plasmids.

Table VIII: Example of matting haploids on a 96 well microtiter plate.

	pGBKT7	α	β	δ
pGADT7				
α				
β				
δ				
ε				

The microtiter plate was placed on rotating platform shaker and incubated at 30°C for 20hrs at 200rpm. To screen for the formation of zygotes or "Mickey Mouse" cells, 10µl of a culture from a sampling of wells were mixed with 100µl of 1x PBS and observed under the microscope (40x objective). Zygotes are two haploids that have created a single diploid cell, as shown in Figure 9 (Kimball, 2003). If few zygotes were identified, the culture was allowed to incubate for an additional 4hrs.

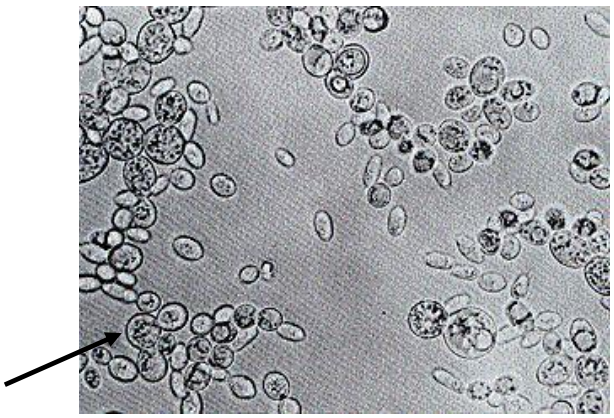

Figure 9: Yeast "Mickey Mouse" cells.

3.4.2 Nutritional selection

Droplets of liquid containing diploid cells were transported with a metallic pin stamp onto SD agar plates lacking leucine and tryptophan. Droplets of liquid are immediately visible on the agar plates and help to verify even plating of all diploid wells. These diploids were allowed to absorb into the agar plates before they were inverted and incubated at 30°C for 3 days. On the third day, pictures were taken to record the amount of diploid growth. Then using a biological loop, approximately the same amount of colonies were streaked out on agar plates lacking leucine, tryptophan and histidine. These plates were incubated at 30°C for

3 days, at which time more pictures were taken to record the growth induced from protein-protein interactions.

4 Results

4.1 F_oF_1-ATP synthase subunits selected for Y2H analysis

The goal of the Y2H analysis is to determine the interacting protein partners of the *T. brucei* F_oF_1-ATP synthase. This is a valuable exercise because 14 of the 22 subunits have no known homology outside of trypanosomatids and this structural information could potentially suggest a function for these hypothetical subunits. Since the Y2H will often give false positive results with very hydrophobic proteins, the best baits include the components of the hydrophilic F_1-ATPase. While these proteins are extremely conserved throughout most eukaryotes studied so far, we have already identified an additional subunit, a hypothetical protein named p18 that is required for F_1-ATPase activity (Gahura, in preparation). Therefore, we can use the other F_1-ATPase subunits as positive controls while searching for the binding partner of p18. All of the remaining *T. brucei* hypothetical proteins are most likely either part of the peripheral stalk, possibly containing a transmembrane domain, or they comprise the supernumerary subunits, which are often membrane embedded. Therefore, we performed some hydrophilicity plots of the each of the subunits to estimate how much of the protein is hydrophobic. Due to the arrangement of the 96 well plates used to perform a high-throughput screen, we chose the 14 most hydrophilic subunits that we were most interested in based on other data previously collected in the lab.

The next factor to consider for each of the subunits is the mitochondrial targeting sequence, typically a short peptide at the N-terminus that directs a nascent protein to the mitochondria. This sequence usually consists of alternating hydrophobic and charged amino acids, forming what is called an amphipathic helix that is cleaved by a signal peptidase once it enters the mt. The properties of the MLS lend itself to create false positives in the Y2H screen, plus we are not sure if the inclusion of the targeting signal would affect the normal folding of the mature protein. Therefore, we decided to exclude the MLS whenever it had been experimentally determined (Table XI). A colleague in the lab used Edmann degradation to sequence the N-terminus of several subunits. When that chemistry failed, the mass spectrometry data for the full mature protein allowed us to identify the targeting sequence for δ and ε. Of the six

subunits with experimentally proven MLS, only 2 were accurately predicted by the online tool MitoProt. Therefore, we decided to use the entire coding sequence for all the remaining subunits used in the screen. In an effort to save time, we reused the β and γ constructs created by an earlier bachelor student in the lab, Matthias Guggenberg. Since these constructs were made before we had any of the experimental evidence for the MLS of the F_oF_1-ATP synthase subunits, the β constructs include the determined MLS. This was not an issue for the γ constructs because they apparently have an unidentified internal MLS.

Table IXI: Experimentally determined MLS for the F_oF_1-ATP synthase candidates used in the Y2H screen. (-MLS – has no N-terminal MLS, ND - not determined, NA – not applicable,)

Protein	Moiety	Gene ID	MLS
α cds	F_1	Tb927.7.7420	MRRFGSKFASGLASRCALACPLAS
α N	F_1	Tb927.7.7420	MRRFGSKFASGLASRCALACPLAS
α C	F_1	Tb927.7.7420	NA
β	F_1	Tb927.3.1380	MLTRFRSAVLRGAVSITGARA
δ	F_1	Tb927.6.4990	MFRTFGRRLVSCTLPLL
ε	F_1	Tb927.10.5050	MIRRSCALL
γ	F_1	Tb927.10.180	-MLS
p18	F_1	Tb927.5.1710	MMRRVYAPVFCSVAAARF
OSCP	F_o	Tb927.10.8030	MFRRLSSSARAVVAARF
520	F_o	Tb927.10.520	ND
1270	F_o	Tb927.11.1270	ND
2930	F_o	Tb927.5.2930	ND
3320	F_o	Tb927.8.3320	ND
6250	F_o	Tb927.11.6250	ND

We have also determined the exact cleavage sites within subunit α that produce an N-terminal fragment and a longer C-terminal fragment, both of which remain associated with F_1-ATPase. Since the N-terminal peptide has a domain known to interact with OSCP, we also included these individual peptides in the Y2H screen.

4.2 Bacterial cloning

We employed KOD Hot Start DNA Polymerase to amplify gene inserts for each plasmid. The outcome was tested via electrophoresis, which allows us to observe if the amplification was specifc (only one band detected) and accurate (band is expected size). Only the amplification of 520 proved troublesome, but the addition of a 1µl of DMSO, which aids

in the denaturing of templates with high GC content, produced enough of the correct amplification product. A representative 1% agarose gel demonstrates the specificity and accuracy of the PCR amplifications (Figure 10).

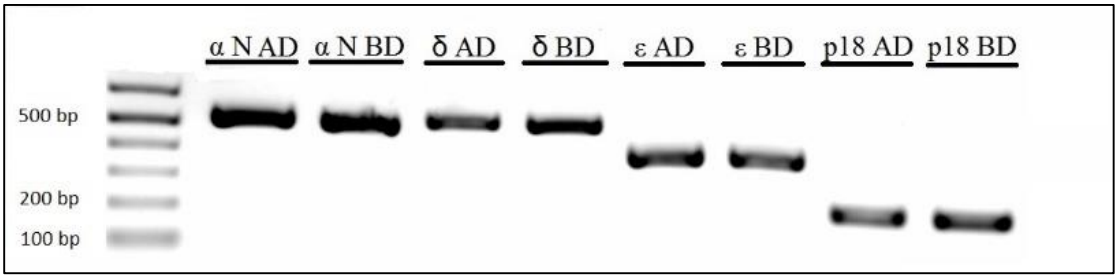

Figure 10: Representative PCR fragments visualized on a 1% agarose gel stained with ethidium bromide.

Once the vector was assembled by end homology with the GENEART Seamless Cloning Kit, it was transformed into chemically competent *E. coli* that were then plated on agar. The antibiotics kanamycin and ampicillin were used to select for positive bacterial colonies containing the pGBKT7 or pGADT7 plasmids, respectively. The transformation process was quite successful as we routinely observed between 50-100 colonies on every plate. In order to verify that the bacteria contained vectors with the expected pcr inserts, plasmid DNA isolated from two individual bacterial colonies was cleaved with analytical restriction enzymes. 500ng of plasmid was digested because it produces enough of the expected smaller DNA bands to be visualized by ethidium bromide while not requiring a significant amount of enzyme and incubation time. We carefully chose the combination of restriction enzymes so that we would get a clear distinction between empty vectors and vectors with correct inserts. Several factors that affected this decision involved the restriction sites present in the distinct multiple cloning sites of the AD and BD vectors, if any of these sites occurred in our pcr insert and if the interval of these sites produced bands with varying sizes that were easy to distinguish.

Table XII: Expected banding pattern for a selection of several analytical restriction digests.

Plasmids	Restriction enzyme	Expected sizes (bp)	Parental sizes (bp)
1270 AD	Hind III	7192, 1251	7192, 804
6250 AD	Hind III	7192, 1551	7192, 804
6250 BD	StuI	8076	-
1270 BD	XhoI, NotI	6960, 824	6960, 352

Table XII indicates the restriction enzymes used for several Y2H plasmids and the expected sizes of the resulting DNA fragments if the plasmid contains the insert (expected) or not (empty vector). The empty 6250 BD doesn't contain a StuI restriction site, so it would remain an uncut plasmid, which usually produces two bands that represent supercoiled and nicked plasmids. Figure 11 contains the resulting digests, where the control lanes are the respective empty parental vectors (pGADT7 – lane 3, pGBKT7 – lane 7 and 10) digested with the same restriction enzymes. This makes it easy to distinguish between vectors with inserts and empty vectors. There is no sample loaded in lane 4 and the arrows mark very faint bands observed in lane 3 and 10. Typically, we observed that both clones of each plasmid were correct. For every plasmid we obtained at least one positive clone.

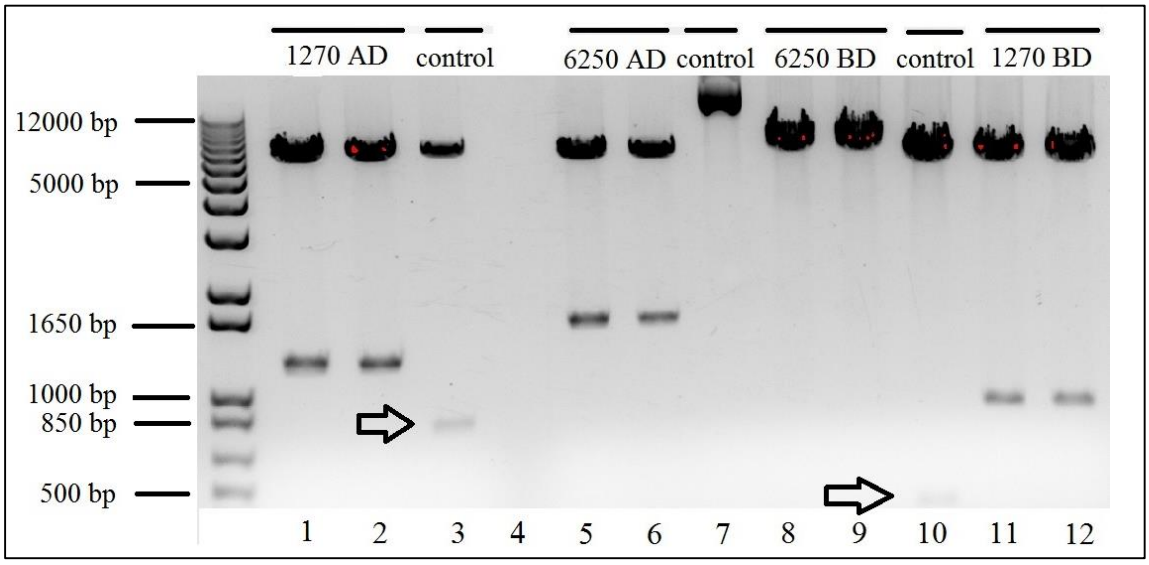

Figure 11: Representative restriction digest analysis of various Y2H vectors.

4.3 Sequencing

While the restriction digest analysis identified positive clones containing the pcr insert, the plasmids need to be sequenced to make sure that the coding sequence doesn't contain any nonsense or missense mutations introduced by the DNA polymerase during PCR. Each positive clone was sequenced by Seqme s.r.o. with two primers that we designed specifically for the plasmids.

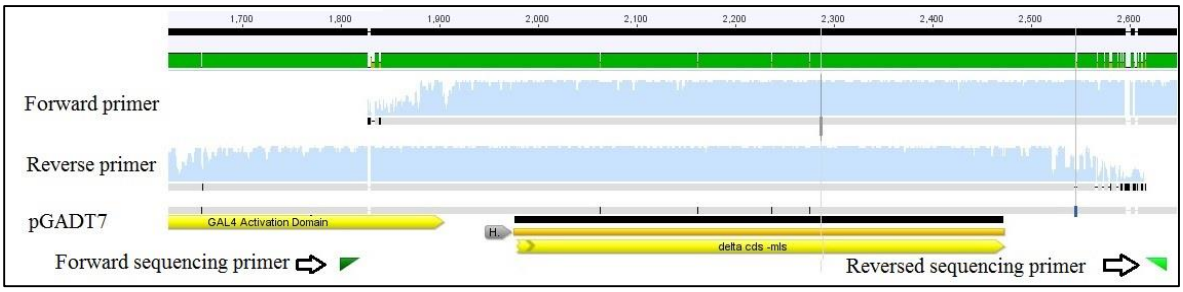

Figure 12: Geneious schematic of the sequencing data obtained for δ pGADT7, the forward and reverse sequencing primers are depicted.

Figure 12 illustrates where the sequencing primers (AZ0559 forward, AZ0560 reverse) anneal to the pGADT7 plasmid. Similar primers (AZ0561 forward and AZ0562 reverse) were also created for the pGBKT7 vector. These primers were created to generate sequences that could determine if the fused GAL4 domain was in frame with the epitope tag and our gene of interest. Once the resulting sequencing files were downloaded online, the Geneious software package implemented the MUSCLE algorithm to align the sequencing chromatograms with their respective plasmid created *in silico*. A typical Seqme sequencing read with high quality nucleotide assignment is about 900bp, which allowed us to obtain 2x coverage for most of the gene insert and epitope tag. Only the α plasmids (containing the longest pcr insert, 1714bp) didn't afford such excellent coverage. The Geneious alignments enable the chromatogram sequences to be translated in the correct reading frame, which led to the quick identification of any problematic regions. A summary of all the mutations identified in each of the plasmids used in the Y2H screen are listed in Table XIII. Only mutations that were observed in both the forward and reverse sequencing reads were recorded, unless the mutation was observed within a region of high quality and there was no data for the opposite strand.

There are several different types of mutations that can arise. Silent mutations generally occur in the third wobble position of the codon and encode the same amino acid. Missense mutations result in an altered codon that encodes a different amino acid in the protein and caution should be paid when this mutation results in an amino acid with different chemical properties. When nonsense mutations arise, the gene product is prematurely truncated because a new stop codon has been inserted.

Table XIII: Mutation identified by sequencing, X – mutation, (–) no mutation, ND – no data, X – probable PCR error, S114T – mutation results in aa with conserved chemical properties, N282Y – mutation results in aa with different chemical properties.

Gene	Protein	Codon	pGADT7 forward	pGADT7 reverse	pGBKT7 forward	pGBKT7 reverse
α	I56I	ATT->ATC	-	ND	X	ND
	I149I	ATC->ATT	X	ND	X	ND
	Q442Q	CAA->CAG	ND	-	ND	X
α - N	I56I	ATT->ATC	X	X	-	-
α - C	I14I	ATC->ATT	X	ND	X	ND
	Q310Q	CAA->CAG	ND	X	ND	X
δ	N29N	AAT->AAC	X	X	X	X
	Y62Y	TAC->TAT	X	X	X	X
	G87G	GGG->GGA	X	X	X	X
	H100H	CAC->CAT	X	X	X	X
ε	T19T	ACC->ACG	X	X	X	X
p18	G118G	GGC->GGT	X	X	X	X
OSCP	H162H	CAC->CAT	X	X	X	X
	G201G	GGT->GGC	X	X	X	X
520	N282Y	AAC->TAC	ND	X	ND	-
	R286R	CGA->CGG	ND	X	ND	X
	E330E	GAG->GAA	ND	X	ND	X
	G348G	GGG->GGA	ND	X	ND	X
	G361G	GGG->GGA	ND	X	ND	X
1270	G99G	GGC->GGT	X	X	ND	X
	S114T	TCT->ACT	X	X	X	X
	S143S	AGT->AGC	X	X	X	X
	T55T	ACA->ACG	-	-	X	X
3320	T28T	ACA->ACG	X	ND	X	ND
	G37G	GGG->GGA	X	ND	X	ND
	Q76Q	CAG->CAA	X	ND	X	ND
	L120L	CTT->CTA	X	ND	X	ND
	A126A	GCG->GCC	X	ND	X	ND
	R256R	CGA->CGG	X	X	X	X
	D308E	GAT->GAG	ND	X	ND	X
	G316E	GGG->GAG	ND	X	ND	X
6250	A62V	GCC->GTC	X	X	X	X
	A125A	GCT->GCC	X	X	X	X
	E163G	GAG->GGG	X	X	-	-
	S176V	TCC->GTC	X	X	-	-
	A197G	GCA->GGA	-	-	ND	X
	G208G	GGA->GGG	-	-	X	X

Since the β and γ plasmids were created and sequenced previously by Matthias, the data is not included here. Also, both the AD and BD plasmids of 2930 didn't have any identified mutations. Upon closer inspection, most of the recorded mutations occur at the wobble position and did not have any effect on translation since the altered codon still encoded for the same amino acid. These silent mutations are often assumed to be differences between *T. brucei* lab strains and don't pose a threat to the proper folding of the gene product. Other mutations resulted in the change of the amino acid (1270 S114T or 3320 D308E), but the chemical properties of the side chain remained the same. Since these mutations were seen in both the pGADT7 and pGBKT7 vectors, these were also interpreted as minor differences between lab strains. However, the 520, 3320 and 6250 plasmids contain a few mutations that resulted in amino acids with quite different properties (520 N282Y, 3320 G316E, 6250 E163G, 6250 S176V, and 6250 A197G). Only the G316E mutation in the 3320 plasmid was found in both vectors and might be considered a strain difference. On the other hand, the other mutations were only present in one of the two plasmids, which strongly suggests that there was an error in that specific pcr product. All similar mutations are marked as probable PCR errors, which means that the sequence should be verified for the following 6 plasmids (α pGBKT7, α-N pGADT7, 520 pGADT7, 1270 pGBKT7, 6250 pGADT7 and 6250 pGBKT7). Since both α plasmids and the 1270 vector contain silent mutations, these will be fine to include in the Y2H analysis. However, if after resequencing, both the 520 and 6250 mutations persist, these three plasmids will most likely need to be created again using new pcr products. Fortunately, none of the plasmids contained any nonsense mutations that resulted in a premature stop codon.

We also analyzed the sequence for the empty parental plasmids (pGADT7 and pGBKT7). Using the forward primers, we were able to confirm that these parental vectors have no inserted bait or prey gene fragment because we observe the epitope tag, the MCS and then the stop codon. The reverse primer again confirms there is no insert, while also providing sequence for the Gal4 domain through the start methionine. Neither plasmid appears to have any mutations, which will become important for later analyses using these constructs in the Y2H screen.

4.4 Yeast transformation

Since there are 226 pairwise combinations we would like to screen by Y2H, we decided it would be more economical to transform all of the AD vectors into the haploid mating strain a and all of the BD vectors into the haploid mating strain α. Then we could mate haploids containing each individual bait and prey plasmids to create diploids that could be tested for protein-protein interactions. The transformation of these Y2H plasmids into *S. cerevisiae* was quite successful as we obtained about 50-100 round white colonies on every plate. The transformation plates lacked an essential amino acid so that only yeast containing the desired plasmid would be able to grow. Contamination occurred only in a few cases and usually manifested in yellow colonies with a different morphology. Still, this contamination occurred on just a small portion of the plate and the correct yeast colonies were still able to be selected.

4.5 Western blots

While the yeast must contain the Y2H plasmid to grow on the dropout media, we need to verify that they are synthesizing the Gal4 fusion protein of interest. Incorporated into the plasmid is an epitope marker that is expressed as part of the protein. A commercial antibody specific for this tag can then be used for a western blot analysis to determine the size of the ectopic protein and its expression levels. Preparing whole cell lysate samples for Western blot analysis is complicated by the rigid yeast cell wall. Therefore, we decided to use a mild alkali treatment that is followed by boiling in laemmli buffer. This very quick approach requires a small amount of cells and is fairly reproducible. Roughly equal amounts of cells were harvested based on their optical density and then we loaded 20µl of the prepared samples per well. The mouse anti-HA primary antibody was used to detect proteins expressed from the pGADT7 vector, while a rabbit c-Myc antibody recognized pGBKT7 proteins. Table XIV contains the proteins used in this study, their expected molecular weight, their fused epitope tag and the expression plasmid.

Table XIV: Expected molecular weights of tagged proteins.

Protein	kDa	Tag	Plasmid
empty	20	HA	AD
SV40	91	HA	AD
α	79	HA	AD
β	71	HA	AD
α C	65	HA	AD
520	64	HA	AD
3320	64	HA	AD
2930	61	HA	AD
γ	52	HA	AD
OSCP	47	HA	AD
6250	45	HA	AD
1270	38	HA	AD
p18	37	HA	AD
δ	36	HA	AD
α N	31	HA	AD
ε	26	HA	AD
empty	22	c-Myc	BD
α	81	c-Myc	BD
β	74	c-Myc	BD
α C	67	c-Myc	BD
520	67	c-Myc	BD
3320	65	c-Myc	BD
2930	63	c-Myc	BD
p53	56	c-Myc	BD
γ	54	c-Myc	BD
OSCP	49	c-Myc	BD
6250	48	c-Myc	BD
1270	40	c-Myc	BD
p18	39	c-Myc	BD
δ	38	c-Myc	BD
α N	33	c-Myc	BD
ε	28	c-Myc	BD

First we wanted to determine if protein expression levels varied between individual yeast colonies on a transformation plate. Also, due to the limited availability to perform lengthy experiments, we tested if storing transformed yeast at 4°C for a period of time had a negative effect on protein expression. Therefore, we analyzed three colonies taken from a fresh transformation plate and three colonies from a 4°C plate that were re-streaked and allowed to grow for three days on an agar plate (Figure 13). The 1270 AD proteins had similar levels of expression from the individual colonies on a single plate, but the overall expression

seemed diminished by half in the older yeast. The tagged proteins from the 1270 BD vectors had more variability between yeast colonies on a single plate, where expression levels were either robust or not detectable. Again, the lone positive colony from the fresh transformation plates seemed to have 2x the amount of protein compared to the older yeast stocks. It is impossible at this point to know if the fluctuation in the western signals are due to reproducibility issues with preparing the yeast whole cell lysates or if there are generally large protein variances between yeast colonies. Since we detect a high molecular weight non-specific band in 5 out of the 6 colonies tested for 1270 BD, it would suggest that cell lysis was at least partially successful. However, this will need to be examined more closely to ensure that negative interactions in the large Y2H screen are not due to the protein expression of individual yeast colonies.

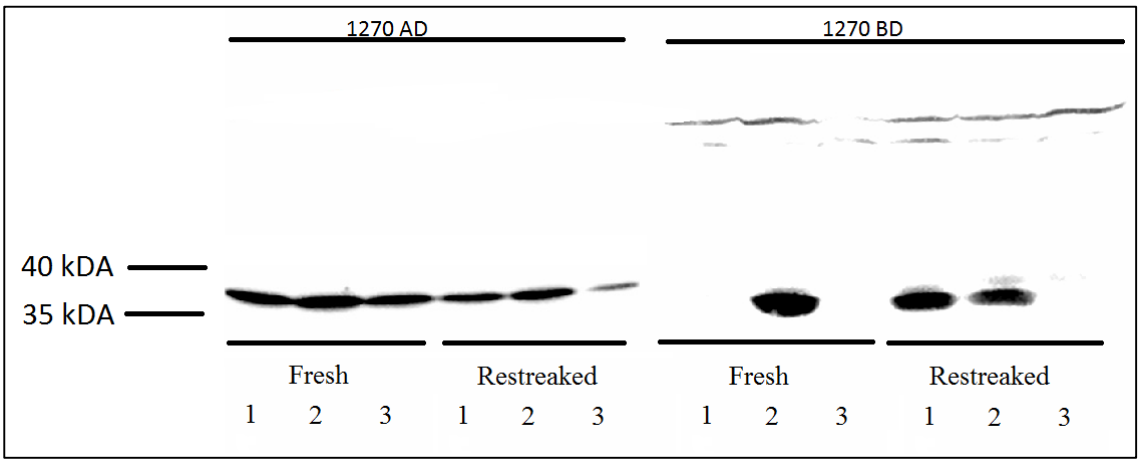

Figure 13: Western blot illustrating the varying levels of protein expression between freshly transformed yeast and those stored at 4°C.

Since these results were rather inconclusive, we continued to analyze three colonies from most transformations to verify protein size and expression (Figure 14). We observed that most proteins had good expression levels and ran at their expected mobility, without any signs of proteolytic damage. Furthermore, in support of the 1270 AD results from Figure 13, we usually didn't observe large differences in protein expression between selected yeast colonies. This outcome will simplify our screen and it is a reasonable expectation as most published Y2H screens don't thoroughly test for protein expression for logistical reasons. It is important to note that in most instances when we observe very low protein expression levels, it is usually from yeast colonies that grew poorly after having been stored at 4°C for 6 weeks or longer. That being said, we obtained only weak expression from the α, α N and OSCP BD vectors, but they were a single band of the expected size. Unfortunately, yeast containing these

vectors were only tested once. It is probable that the overnight growth of these yeast was less than optimal and less material from the whole cell lysate was actually loaded onto this gel. New samples for these yeast will need to be prepared and analyzed again. Finally, the empty pGADT7 and pGBKT7 plasmids produced the expected small protein consisting of the corresponding Gal4 domain fused with the epitope tag, providing additional evidence that these parental vectors express the correct proteins.

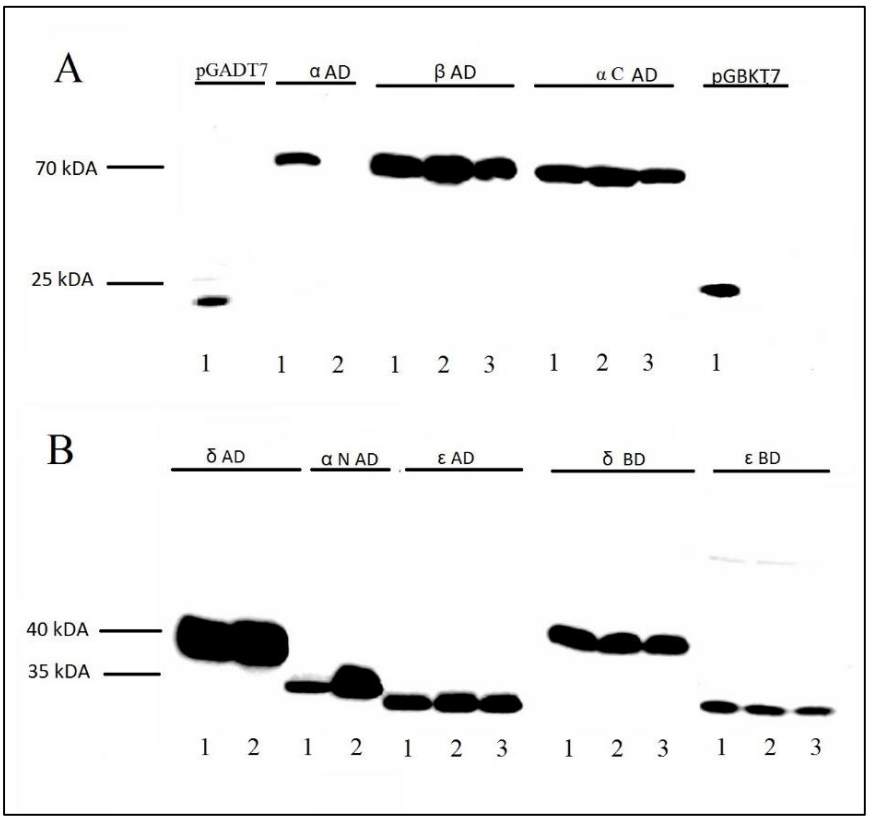

Figure 14: Western blot depicting proteins running at their expected mobility on a A) 10% or B) 12% SDS-PAGE gel.

A few of the proteins we analyzed demonstrated either double bands or mobilities that differed from their expected sizes (Figure 15). OSCP fused with the Gal4 activation domain produced two bands that were detected by the epitope tag antibody. Since the lower band runs at the predicted mobility, it would indicate that the higher band might contain some post-translation modification. Interestingly, the ratio between the upper and lower band is consistent between colonies (~1:1). These samples will need to be prepared again though because the OSCP fusion protein expressed in the BD vector produced just a single band of the correct size, suggesting that either the AD protein isolation was defective or somehow the Gal4 activation domain in this context gets modified in yeast. While we also saw a double band appear in one of the p18 AD samples, we also observed a much greater range of protein

expression in these yeast colonies. The lower band again runs at the predicted size, but this time only a very small portion of the total protein detected by the HA antibody has a higher mobility. This low level of modified protein will probably be insignificant in the Y2H screen, but since no data was collected from the p18 fusion protein expressed in the BD vector, these samples will have to be investigated again. Finally, the AD γ subunit also produced unexpected results as the signal was detected around 23kDa, which is significantly less than the predicted 52kDa. This difference is most likely the result of proteolytic cleavages that occurred during a flawed sample preparation, since the ectopic γ fusion protein produced by the BD vector was the correct size. Furthermore, the isolation of the *T. brucei* F_1-ATPase complex contains the γ subunit resolved at its expected size of 34kDa (Gahura manuscript, submitted). Therefore, we will need to harvest and analyze new samples for AD γ, as well as for AD 3320 and BD p18, which we have not been able to detect any protein signal for yet.

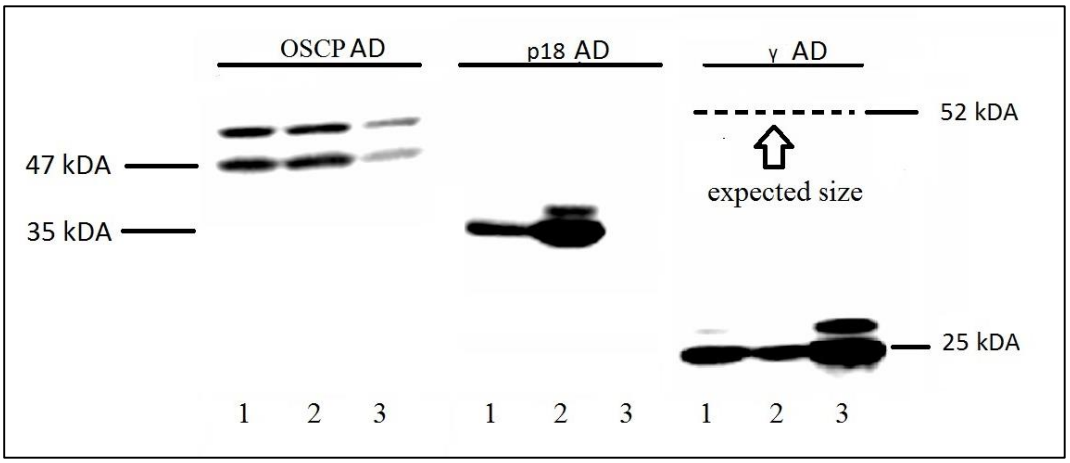

Figure 15: Western blot results of OSCP, p18 and γ AD proteins resolved on a 10% SDS PAGE gel.

4.6 Y2H

The creation of diploids, containing both the bait and prey plasmids, involved mixing haploids of the α and a mating types into a 96 well plate containing nutrient rich media and allowing them to incubate for 20-24 hrs at 30°C. Typically, around 20 hrs the culture should contain zygotes that appear as 3-lobe structures, resembling a clover leaf or a "Mickey Mouse" face. However, when we observed the yeast cultures at this time under an inverted light microscope, we didn't visualize many of these structures and thus let them incubate for an additional 4hrs, as recommended by the protocol. It is possible that the vast majority of haploids already mated and were now dividing as budding diploids, which are very similar

to dividing haploids, except with a slightly larger almond shape. These cultures were then transferred by a metal pin stamp onto SD-Leu-Trp agar plates that were the same size as the 96 well plate (Figure 16). This selection for diploids was done a total of 3 times. Unfortunately, while each experiment contained robust growth for most of the colonies, there were always some that failed to grow as well. For example, most spots of yeast formed a circular lawn after three days of growth at 30°C, but some seemingly random spots would only produce a few individual colonies. Since this diminished growth was observed at various locations on the plate, it was probably not due to the metal stamp or the location on the agar plate. The most likely explanation is that the amount of yeast cells used to seed each of the overnight wells was inconsistent. This irregular growth creates problems when the pin stamp is used to replicate these diploids onto agar plates used to screen for Y2H interactions based on the amount of growth.

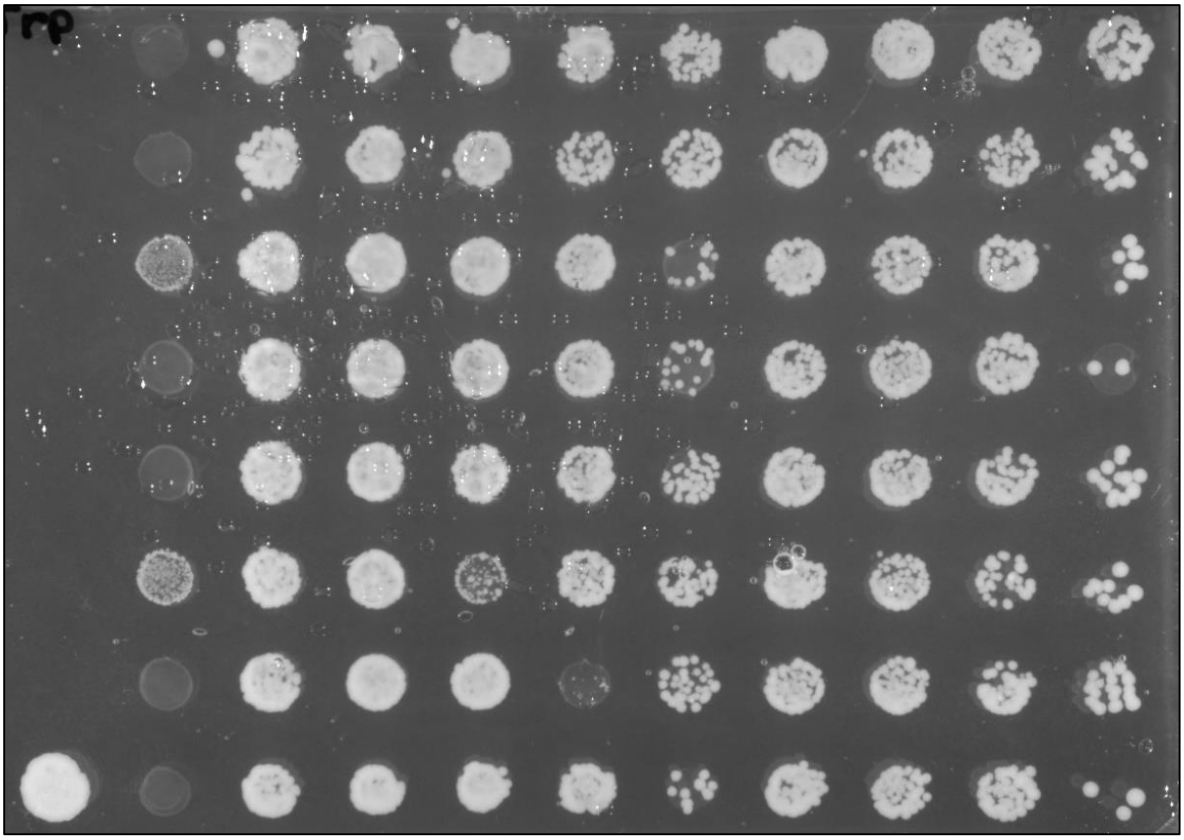

Figure 16: Diploid selection with inconsistent growth on SD-Leu-Trp plate.

While we performed the Y2H screen 3 times, our results were often inconclusive for several reasons. The first problem we encountered was that when either the empty AD or BD vector was paired with its counterpart vector containing an ATP synthase subunit, it seemed

to auto-activate, resulting in vigorous growth on SD-Leu-Trp-His plates. While this unfortunate outcome can be expected to occur on occasion, it happened far too frequently in our screens. These false positives makes it almost impossible to make any strong conclusions about the remaining data, but for the sake of presenting more of the hard work that was achieved, we will attempt to analyze the rest of our preliminary results.

One optimistic result involved the robust growth we observed with a diploid harboring both the p53 BD and SV 40 AD vectors, a positive control that was included in almost all experiments (Figure 17). In fact this interaction was so strong, that significant growth was observed on SD-Leu-Trp-His plates containing 3-AT, a specific inhibitor of the histidine synthesis pathway. While there was an inverse linear relationship between the amount of growth and the 3-AT concentration added to the agar plate, growth was still observed at 30mM 3-AT. This indicates that the manner of applying 3-AT to the plates rendered it an effective tool to objectively gauge the strength of any interaction pairs and could possibly be used to tease apart growth due to auto-activation versus protein-protein interaction.

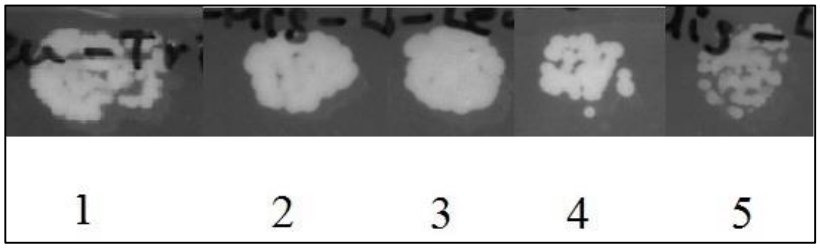

Figure 17: Y2H positive control for protein-protein interactions (1) SD-Leu-Trp, (2) SD-Leu-Trp-His, (3) SD-Leu-Trp-His 2.5mM 3-AT, (4) SD-Leu-Trp-His 10mM 3-AT, (5) SD-Leu-Trp-His 30mM 3-AT).

Unfortunately, we rarely detected any other pairings that resulted in such significant growth, even on the SD-Leu-Trp-His plates with just 2.5mM 3-AT. While the amounts of 3-AT applied will always need to be adjusted based on the strength of interactions witnessed in each Y2H screen, 2.5mM 3-AT is considered a fairly low value that should not thoroughly disrupt the growth as we often saw. Therefore, from this point forward, any interactions we mention will be from significant growth (scored at least a 3 out of 5) observed on just the SD-Leu-Trp-His plates with 0mM 3-AT.

The most obvious protein-protein interactions we observed occurred when diploids were grown on SD-Leu-Trp-His plates for three days before they were streaked out onto SD-Leu-Trp-His plates with 0mM 3-AT. It is unclear to us if these yeast colonies led to better growth because the diploids were able to acclimate to SD-Leu-Trp media before being screened for interactions or whether the physical method of streaking the cells on the plates

with a wire loop vs a metal stamp resulted in better overall fitness. The interactions observed on these plates were generally much more clear to interpret as either yeast were able to grow robustly or not at all (Figure 18). However, we still observe strong auto-activation with both the empty AD and BD vectors (Table XV). Furthermore, the interactions recorded between components of the known F_1-ATPase in this study don't seem self-evident, at least based on the well-established bovine model. Unfortunately, there was miscommunication between myself and the supervisor, who streaked out the diploids onto the SD-Leu-Trp-His, and it is possible that the interactions were mislabeled. Nevertheless, this technique seems promising and will be further optimized and used in combination with the more stringent adenine reporter.

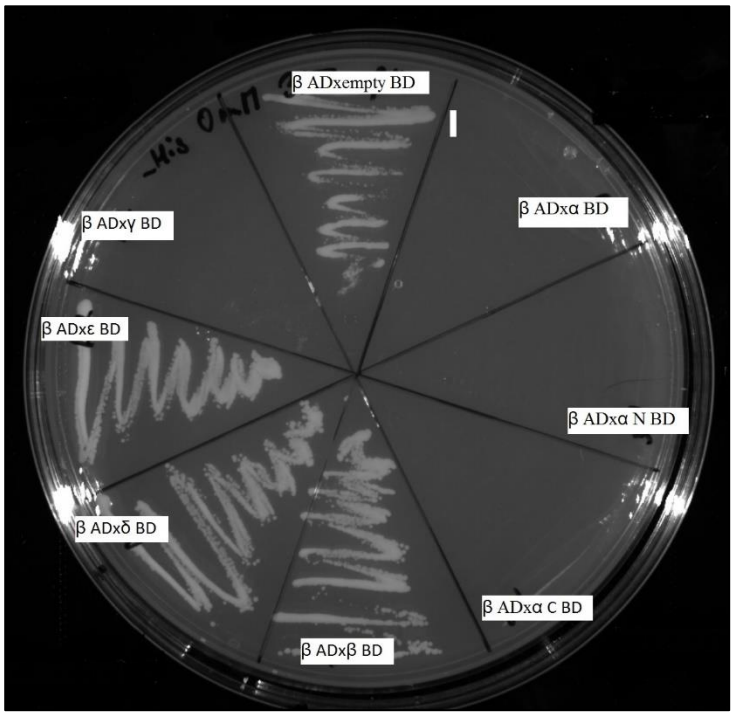

Figure 18: Y2H results with using the streaking out method on SD-Leu-Trp-His.

TableXV: Y2H results with restreaking method (growth: + small, ++ intermediate , +++ big).

	empty AD	α –mls cds	α –mls N-term	α –mls C-term	β –mls
empty BD	+++	+++	+++	++	+++
α –mls cds	++	+	+	++	
α –mls N-term	+	+	+++	+	
α –mls C-term	+	++	+	+	
β –mls	+++	++	+++	++	+++
δ –mls	+++	+++	+++	+++	+++
ε –mls	++	+	+++	++	+++
γ cds	++		+	+	

5 Discussion

While we easily cloned all subunits into the bait and prey plasmids, the sequencing data revealed that several of the plasmids contained mutations in the pcr insert. We determined that if a mutation was seen in both the pGADT7 and pGBKT7, that it was most likely due to a difference between lab strains. This is based on the assumption that it would be almost statistically impossible for a proof-reading DNA polymerase to make the same error in two different pcr reactions. Of course, this could happen if the error occurred in a particularly difficult stretch of the genomic DNA, like a highly repetitive sequence. However, a cursory look at the sequence context for these mutations didn't reveal any obvious problematic DNA regions. Fortunately, most of the changes we identified were silent mutations and were included in the Y2H screen. However, some mutations resulted in amino acids with side chains that have different chemical attributes and these could potentially disrupt the folding of a domain that is involved in protein-protein interactions, resulting in false negatives. Since we had difficulties identifying strong interacting partners in our earliest screens, it would be best to eliminate this potential problem and clone the 520 AD, 6250 AD and BD vectors again. Another factor that could affect protein folding or protein-protein interactions is the inclusion of the mitochondrial localization sequence (MLS), which gets cleaved to create the mature protein once it enters the mitochondria. Since its effect on the recombinant protein is unknown, it will be interesting to analyze specific subunits with or without their experimentally determined MLS and see if the interacting partners vary.

Another factor that could influence the success of the Y2H screen is the protein expression levels in the transformed yeast. While not all published Y2H screens routinely check protein

expression, we wanted to verify that these trypanosomatid specific genes were being translated into stable proteins of expected size. Therefore, we implemented a mild alkali treatment to quickly lyse yeast cells and isolate enough protein to be analyzed by a western blot. Our preliminary results suggest that most yeast colonies express similar levels of protein, which should be sufficient for a Y2H screen. However, there were instances when no epitope tagged protein was detected from one of several colonies obtained from a single transformation plate. Therefore, it needs to be determined if this reproducibility issue between yeast colonies is due to the amount of yeast growth in the overnight culture, the protein isolation procedure or if it will be required to screen individual colonies for protein expression before including them in a Y2H screen. We also observed that colonies stored on an agar plate at 4°C for 6 weeks or more tend to produce less of the recombinant protein, even when they are streaked onto a fresh agar plate. Therefore, it will be imperative to use freshly transformed yeast and plan the extensive Y2H screen when the investigator has a significant block of time to spend in the lab, perhaps during the summer months. Finally, a few of the proteins were observed at unexpected molecular sizes, either due to proteolysis or post-translational modifications. Since this generally occurred for a protein expressed in only one of the two Y2H plasmids, most likely these variances were due to flawed protein preparations and simply need to be repeated to ensure that false negatives don't arise from modified recombinant proteins.

While our various attempts at the Y2H screen did yield any exciting results, we have begun to optimize the procedure to resolve our unique problems. The first concern is if we are successfully mating haploids to create diploids with both Y2H plasmids. The formation of zygotes (visible as "Mickey Mouse" cells or clovers) should be observed under the microscope after 20 hrs of haploid incubation. However, we rarely see cells with this obvious morphology, which makes us question if the conditions are correct for diploid formation. Maybe we are looking too late for zygotes or we are not seeding enough of each haploid, both of these factors will have to be optimized. Still, when the diploids are selected on agar plates comprised of the appropriate dropout media, we see that most of the haploid mattings result in significant growth, suggesting that these yeast do in fact contain both Y2H plasmids. However, it is puzzling that the growth is not consistent between all selected diploids on a plate with up to 96 mattings. Depending on the method of plating, the diminished growth of some diploids could significantly alter the outcome of the screen and result in false negatives. Furthermore, since the mating haploids are incubated in YPDA rich media, it appears that it is best to allow the selected diploids time to acclimate to media lacking leucine and tryptophan before adding additional stress in the form of the Y2H screen.

Even when diploids were successfully selected, we frequently observed that diploids containing one of the empty vectors had significant growth in the Y2H screen, resulting in false negatives. While it is possible for a protein with a hydrophobic domain fused to the Gal4 DNA binding domain to activate transcription of the reporter gene without a binding partner, we saw auto-activation even with the empty BD vector. Therefore, we went back to sequence the empty Y2H vectors again, but the results didn't indicate any problems. Even the protein expression from these plasmids without a pcr insert resulted in a tagged Gal4 protein of the expected size. While the negative control plasmids appeared correct, auto-activation can occasionally be observed from proteins with specific properties. However, we saw robust growth with almost every pairwise combination. This seems to suggest that the growth we saw was not actually due to the activation of the histidine reporter, but rather represented the background growth that can be observed with mutants in this biochemical pathway.

Another factor that needs to be addressed is the reporter assay itself. The His3 gene reporter we chose to use is an enzyme involved in the synthesis of the essential amino acid histidine. However, this biochemical pathway is notorious for producing false positives, so it is routine for screens to include various concentrations of the inhibitor 3-AT to discern which growth is truly dependent on bait-prey interactions. The concentration of 3-AT to add varies on the outcome of each screen. High concentrations (up to 30mM) can be used when there is a problem with auto-activation, but in other screens growth observed on 2-3.5 mM 3-AT is considered to represent strong interactions between two proteins. Since we had difficulties with empty vectors promoting growth in the Y2H screen, we used a wide concentration range of 3-AT (2.5mM, 10mM and 30mM). However, even the lowest concentration severely limited all growth the same, so no positive interactions could be detected. In these initial studies, the 3-AT is always added into the agar plate after the components have cooled slightly, but it is unclear if it would be better to spread 3-AT on top of the solidified plates. This would make it difficult to evenly control the concentration of the inhibitor across the plate and would seem to make more sense if the applied 3-AT had no effect on growth. Hence, when we resume the Y2H screen, we will allow diploids to grow on SD-Leu-Trp plates for 3 days and then streak them out with a wire loop on plates lacking adenine. This way we can utilize the more stringent adenine reporter that doesn't require an additional inhibitor. Any positive interactions identified between bait and prey subunits in this screen can then be used in a β-galactosidase assay, which can quantify the strength of the protein-protein interaction based on the lac2 reporter gene.

6 Conclusion

Based on their hydrophobicity, we selected 14 subunits of the *T. brucei* F_oF_1-ATP synthase complex to include in a Y2H screen. All of these subunits were easily cloned into the pGADT7 and pGBKT7 Y2H plasmids using a commercial enzyme mix to fuse together DNA fragments with terminal end-homology. Sequencing of the gene inserts identified some conserved alterations from the referenced genomic strain (interpreted as lab strain differences), while other mutations that changed the physical properties of an amino acid were most likely introduced during pcr reactions. The former usually resulted in silent mutations and were used for the Y2H screen. Only three plasmids (520 AD and 6250 AD and BD), will need to be re-sequenced and cloned again if the mutations persist.

Each of the bait and prey plasmids were transformed into S. cerevisiae PJ69-4 haploids with either the a or α mating type. Yeast protein expression from each plasmid was verified by western blot analysis using antibodies that recognized the fused epitope tag. Most of the yeast produced significant and consistent amounts of the tagged protein migrating at the correct size. A few yeast will need to be analyzed again since the expression levels detected were quite low, or in the case of the γ AD fusion protein, where the molecular weight of the protein was significantly less than expected. Several attempts to mate haploids in various pairwise combinations and test the resulting diploids for protein-protein interactions resulted with inconclusive data. However, our preliminary trials suggest that after a few modifications (e.g. allowing diploids several days to acclimate to SD-Leu-Trp media before being screened for protein-protein interactions, streaking out the diploids rather than using a high-throughput stamp and using the more stringent adenine reporter), the Y2H screen can be helpful in deciphering the function of *T. brucei* F_oF_1-subunits without any known homologues outside of trypanosomatids.

7 References

- ALI, Syed Hamid and James A. DECAPRIO. Cellular transformation by SV40 large T antigen: interaction with host proteins. *Seminars in Cancer Biology* [online]. 2001, **11**(1), 15-23 [cit. 2016-03-31]. DOI: 10.1006/scbi.2000.0342. ISSN 1044579x. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1044579X00903420
- Amino acid analysis analysis. *Laboratory of Aquaculture & Artemia Reference Center* [online]. 2003 [cit. 2016-02-01]. Available from: http://www.aquaculture.ugent.be/Education/coursematerial/online%20courses/ATA/analysis/prot_aaa.htm
- BRINGAUD, Frédéric, Loïc RIVIÈRE and Virginie COUSTOU. Energy metabolism of trypanosomatids: Adaptation to available carbon sources. *Molecular and Biochemical Parasitology* [online]. 2006,**149**(1), 1-9 [cit. 2016-03-27]. DOI: 10.1016/j.molbiopara.2006.03.017. ISSN 01666851. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0166685106001150
- BESTEIRO, Sébastien, Michael P. BARRETT, Loïc RIVIÈRE and Frédéric BRINGAUD. Energy generation in insect stages of Trypanosoma brucei: metabolism in flux. *Trends in Parasitology* [online]. 2005,**21**(4), 185-191 [cit. 2016-03-28]. DOI: 10.1016/j.pt.2005.02.008. ISSN 14714922. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1471492205000486
- HANNAERT, Véronique, Frédéric BRINGAUD, Fred R OPPERDOES and Paul AM MICHELS. Evolution of energy metabolism and its compartmentation in Kinetoplastida. *Kinetoplastid Biology and Disease* [online]. 2003, **2**(1), 11- [cit. 2016-03-28]. DOI: 10.1186/1475-9292-2-11. ISSN 14759292. Available from: http://www.kinetoplastids.com/content/2/1/11
- HOLMES, Peter. Tsetse-transmitted trypanosomes – Their biology, disease impact and control. *Journal of Invertebrate Pathology*[online]. 2013, **112**, S11-S14 [cit. 2016-04-18]. DOI: 10.1016/j.jip.2012.07.014. ISSN 00222011. Available from: http://linkinghub.elsevier.com/retrieve/pii/S002220111200184X
- Human African trypanosomiasis: The parasite. *World Health Organization* [online]. [cit. 2016-03-28]. Available from: http://www.who.int/trypanosomiasis_african/parasite/en/
- JAMES, Phil, John HALLADAY and Elizabeth A. CRAIG. Genomic Libraries and a Host Strain Designed for Highly Efficient Two-Hybrid Selectioni n Yeast. *Genetics* [online]. 1996, (144), 1425-1436 [cit. 2016-03-30].
- KIMBALL, John W. Budding Yeast: Saccharomyces cerevisiae. In: *Kimball's Biology Pages* [online]. 2003 [cit. 2016-02-03]. Available from: http://pvtridvs.net/pool/users.rcn.com/jkimball.ma.ultranet/BiologyPages/Y/Yeast.html
- LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* [online]. 1970-8-15, **227**(5259), 680-685 [cit. 2016-02-01]. DOI: 10.1038/227680a0. ISSN 0028-0836. Available from: http://www.nature.com/doifinder/10.1038/227680a0
- LUNDBLAD, Victoria and Kevin STRUHL. Yeast. *Current Protocols in Molecular Biology* [online]. Hoboken, NJ, USA: John Wiley & Sons, Inc, 2008 [cit. 2016-03-30].

DOI: 10.1002/0471142727.mb1300s82. ISBN 0471142727. Available from: http://doi.wiley.com/10.1002/0471142727.mb1300s82

- Mordacq, J. and R. Ellington. 2008. The yeast two-hybrid assay. Pages 167-187, in Tested Studies for Laboratory Teaching, Volume 29 (K.L. Clase, Editor). Proceedings of the 29th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 433 pages.
- MÜLLER-ESTERL, Werner. UNTER MITARB. VON ULRICH BRANDT .*Biochemie: eine Einführung für Mediziner und Naturwissenschaftler*. [Nachdr.]. Heidelberg: Elsevier, Spektrum Akad. Verl, 2007. ISBN 3827405343.
- OPENWETWARE CONTRIBUTORS. SDS-PAGE sample buffer (Morris formulation). In: *OpenWetWare* [online]. OpenWetWare, 2013 [cit. 2016-02-04]. Available from: http://openwetware.org/index.php?title=SDS-PAGE_sample_buffer_%28Morris_formulation%29&oldid=686255
- DR OSWALD, Nick. How SDS-PAGE works. In: *BitesizeBio* [online]. 2008 [cit. 2016-02-01]. Available from: http://bitesizebio.com/580/how-sds-page-works/
- Pierce™ ECL Western Blotting Substrate. *Thermo Fisher Scientific* [online]. 2013 [cit. 2016-02-01]. Available from: https://www.thermofisher.com/order/catalog/product/32106
- SCHNAUFER, Achim, G Desmond CLARK-WALKER, Alodie G STEINBERG and Ken STUART. The F 1 -ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function. *The EMBO Journal* [online]. 2005, **24**(23), 4029-4040 [cit. 2016-03-28]. DOI: 10.1038/sj.emboj.7600862. ISSN 02614189. Available from: http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7600862
- SINO BIOLOGICAL INC. SDS PAGE. In: *Assay Protocol* [online]. Sino Biological Inc., 2013 [cit. 2016-02-04]. Available from: http://www.assay-protocol.com/molecular-biology/electrophoresis/denaturing-page
- STOCK, D. *Molecular Architecture of the Rotary Motor in ATP Synthase* [online]. 1999, [cit. 2016-03-28]. DOI: 10.1126/science.286.5445.1700. ISBN 10.1126/science.286.5445.1700. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.286.5445.1700
- ŠUBRTOVÁ, Karolína, Brian PANICUCCI, Alena ZÍKOVÁ and Achim SCHNAUFER. ATPaseTb2, a Unique Membrane-bound FoF1-ATPase Component, Is Essential in Bloodstream and Dyskinetoplastic Trypanosomes. *PLOS Pathogens* [online]. 2015-2-25, **11**(2), e1004660- [cit. 2016-03-28]. DOI: 10.1371/journal.ppat.1004660. ISSN 1553-7374. Available from: http://dx.plos.org/10.1371/journal.ppat.1004660
- Trypanosomiasis, human African (sleeping sickness). *World Health Organization* [online]. 2016 [cit. 2016-03-28]. Available from: http://www.who.int/mediacentre/factsheets/fs259/en/
- UltraPure™ Agarose. *Thermo Fisher Scientific* [online]. [cit. 2016-03-20]. Available from: www.thermofisher.com
- YANG, Ping-Chang and Tahrin MAHMOOD. Western blot: Technique, theory, and trouble shooting. *North American Journal of Medical Sciences* [online]. 2012, **4**(9), 429- [cit. 2016-02-01]. DOI: 10.4103/1947-2714.100998. ISSN 1947-2714. Available from: http://www.najms.org/text.asp?2012/4/9/429/100998

- Yeast Protocol Handbook. In: *Clontech* [online]. 2009 [cit. 2016-01-29]. Available from: http://www.clontech.com/xxclt_ibcGetAttachment.jsp?cItemId=17602
- ZÍKOVÁ, Alena, Achim SCHNAUFER, Rachel A. DALLEY, Aswini K. PANIGRAHI, Kenneth D. STUART and John C. BOOTHROYD. The F0F1-ATP Synthase Complex Contains Novel Subunits and Is Essential for Procyclic Trypanosoma brucei. *PLoS Pathogens* [online]. 2009-5-15, **5**(5), e1000436- [cit. 2016-03-28]. DOI: 10.1371/journal.ppat.1000436. ISBN 10.1126/science.286.5445.1700. ISSN 1553-7374. Available from: http://dx.plos.org/10.1371/journal.ppat.1000436