EFFECT OF AMINO ACID SUBSITITUTION IN SET1 ON HISTONE H3 METHYLATION AND GENE SILENCING IN *SAACCHAROMYCES CEREVISIAE*

A Senior Honors Thesis

by

MORGAN CHATEAU

Submitted to the Office of Honors Programs
Texas A&M University
In partial fulfillment of the requirements of the

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOWS

April 2008

Majors: Biochemistry and Genetics

ABSTRACT

Effect of Amino Acid Substitution in Set1 on Histone H3 Methylation and Gene Silencing in *Saccharomyces cerevisiae* (April 2008)

Morgan Chateau
Department of Biochemistry and Biophysics
Texas A&M University

Fellows Advisor: Associate Professor Dr. Mary Bryk
Department of Biochemistry and Biophysics

Chromosomal DNA in our cells is wrapped around a histone protein octomer like thread on a spool, forming a structure called a nucleosome. Series of nucleosomes form the nuclear chromosomes found in all eukaryotic organisms. Modifications to histone proteins can change how accessible chromosomal DNA is to protein complexes that act on DNA. DNA sequences that are inaccessible are called silent chromatin and regions that can't be transcribed are subject to "gene silencing." Proper gene silencing is necessary for normal cell development and regulation. Incorrect or missing histone modifications can cause the loss of gene silencing and uncontrolled gene expression similar to the situation in cells of patients with cancer or leukemia. My project focuses on a histone modifying complex COMPASS. COMPASS is composed of eight proteins, one of which is the histone H3 methyltransferase Set1. There are seven Set1 homologs in yeast and over 60 Set1-like proteins in humans, including MLL, which is known to be

associated with human leukemia. Previous studies have shown that Set1 and most COMPASS proteins are essential for gene silencing at the ribosomal DNA locus (rDNA) in yeast. The SET domain is the active site of the Set1 histone methyltransferase, where methyl groups are covalently attached to the fourth lysine residue (K4) of histone H3. My goal is to investigate the effect of six individual amino acid substitutions in the SET domain of Set1; Y967A, I972A, Y993A, H1017L, Y967F, and G951A, on histone H3 methylation and gene silencing. These altered Set1 proteins are being expressed in the yeast *Saccharomyces cerevisiae*. Using Western blots and marker genes, the effect of these mutations are compared to wild type Set1. My data show that there are defects in histone H3 methylation in the amino acid substitution variants of Set1. In five of the mutants there is a complete loss of H3K4 methylation. In the future, we will determine if these altered Set1 proteins are assembled into the COMPASS complex. By characterizing the catalytic domain of Set1 using amino acid substitution variants, we will acquire a better understanding of the related proteins in humans.

ACKNOWLEDGMENTS

I’d like to thank Dr. Mary Bryk and Kelly Williamson.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION[1]

DNA is a long polymer that encodes the genes necessary for life. Most of DNA in eukaryotes (yeast, humans) is contained in the nucleus. During the growth phase of living cells, DNA is kept in a loose chromatin structure by being packaged with histones and other chromatin associated proteins. Chromatin is often described as "beads on a string", where the beads are nucleosomes and the string is linear DNA wrapped around and between each nucleosome.

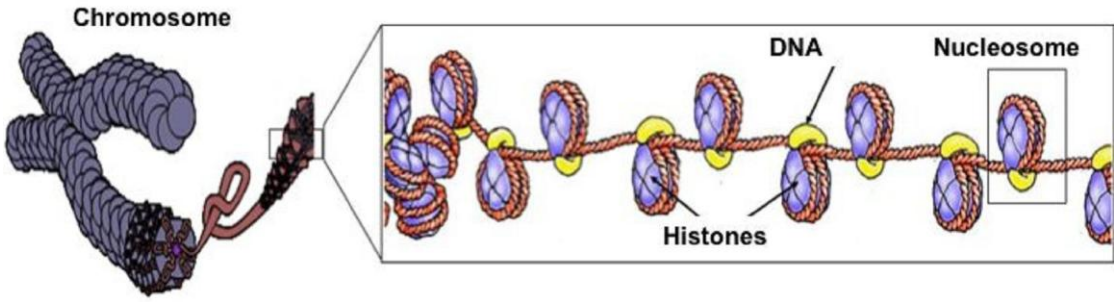
A nucleosome consists of many base pairs of DNA (~150bp) wrapped around a core of eight histones. The histone octamer contains two copies of each of the four types of histones; H2A, H2B, H3, and H4. Linear DNA wraps around a histone octamer about one and three-quarters times to make a nucleosome, like string wrapping around a spool. Several nucleosomes connected in a row resemble the "beads on a string" conformation mentioned before (Figure 1).

---

[1] This thesis follows the style and format of *Journal of Biological Chemistry*

The DNA in some regions of chromosomes is highly compact, such that the genes cannot be transcribed making them silent. For a gene to be expressed, the histones must be released from the DNA. Chemical modification of the histone tails can change the level of compaction of chromatin.

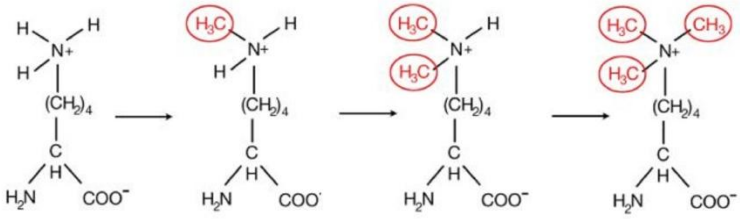
Fig. 1. **A diagram of chromatin in the "beads on a string" conformation**



Histone tail modification is a common regulatory mechanism for gene expression. In *Saccharomyces cerevisiae*, the lysine residue 4 of histone H3 can be methylated by Set1-C or COMPASS (complex associated with Set1) (1). Lysine 4 of histone H3 can be mono-, di-, or trimethylated (1) (Figure 2). The COMPASS complex consists of eight protein subunits, many of which are conserved in higher eukaryotes. Set1 is the histone methyltransferase subunit of COMPASS that is the focus of this project. Set1 is composed of five domains; a RNA-recognition motif, a central domain, an N-SET domain, a catalytic SET domain, and a C-terminal post-SET domain. (3). The project analyzes conserved amino acids in the catalytic SET domain within Set1.

Active chromatin is associated with high levels of K4 methylated histone H3. The level of K4 methylated histone H3 at a specific region of a chromosome varies depending on the transcriptional activity of that region. Figure 2 depicts the different levels of lysine methylation. In eukaryotes, K4 mono- and dimethlylated histone H3 are found throughout the genome at both transcriptionally active and inactive regions (1). In contrast, K4-trimethylated H3 is usually found at active regions within open reading frames (ORFs) that are transcribed by RNA polymerase II (Pol II) (2,7). Silent or transcriptionally inactive domains of chromosomes, such as telomeres and the ribosomal DNA locus (rDNA), have low levels of K4 methylated histone H3 (2, 5).

Fig. 2. **Lysine methylation by COMPASS**



Set1 is required to form an active COMPASS [2]. Without Set1, the cells don't silence genes transcribed by Pol II at the rDNA and telomeres [2]. This project explores the role of individual amino acids in Set1 in histone H3 methylation and gene silencing by making amino acid substitutions in the SET domain of the Set1 protein. Using Western analysis and marker gene expression assays, the

effect of each substitution will be analyzed and compared to wild type Set1 protein.

METHODS

Media

Synthetic complete (SC) and YPADT yeast media were prepared as described (9). SC with minimal adenine is SC media except for the addition of one quarter of the normal amount of adenine (267 μM final concentration).

Mutagenesis and cloning techniques

To make the Set1 variant yeast strains, a plasmid containing a truncated *SET1* gene containing the SET and post-SET domains was amplified using polymerase chain reaction (PCR) with mutagenic oligonucleotide primers containing base substitutions. The mutation-containing plasmids were transformed into *E. coli* (XL1-Blue) and then extracted using a Qiagen plasmid extraction kit. The truncated *SET1* gene was cut out of the plasmid using restriction enzymes ClaI and MfeI, separated on an agarose gel, and purified using a Qiagen Gel Extraction kit. The purified fragment was ligated into a yeast vector pRS406 (10) that had been cut with ClaI and MfeI. The ligation was transformed into *E. coli* (XL1-Blue), ampicillin-resistant transformants were isolated and the correct plasmid was identified by ClaI and MfeI digests. All plasmids were sequenced. The pRS406-*set1* mutant plasmids were linearized by digestion with StuI and then transformed into MBY2269 (*MAT*<u>*a*</u> *ade2Δ::hisG his3Δ200 leu2*$^{-}$*(Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52 set1Δ::TRP1* Ty1*his3AI-236*

ADE2::TELVR) using standard yeast transformation techniques. Genomic DNA was extracted from Ura$^+$ transformants using the technique of Hoffman and Winston (11). Proper insertion of the *set1* mutant plasmid into the *ura3-52* locus in MBY2269 was verified by Southern blotting. The Set1 variant strains are summarized in Table 1.

Table 1. Yeast and *E. coli* Strains

***E. coli* with plasmid**	***S. cerevisiae* (MBY2269 + integrated plasmid)**	**Yeast Genotype**	**Amino Acid Substitution**
MBB546	MBY2382	*MATa ade2Δ::hisG his3Δ200 leu2$^-$ (Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-Y967A set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	Y967A
MBB547	MBY2383	*MATa ade2Δ::hisG his3Δ200 leu2$^-$(Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-I972A set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	I972A
MBB548	MBY2384	*MATa ade2Δ::hisG his3Δ200 leu2$^-$ (Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-Y933A set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	Y993A
MBB549	MBY2385	*MATa ade2Δ::hisG his3Δ200 leu2$^-$ (Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-H1017L set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	H1017L
MBB550	MBY2386	*MATa ade2Δ::hisG his3Δ200 leu2$^-$ (Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-Y967F set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	Y967F
MBB551	MBY2387	*MATa ade2Δ::hisG his3Δ200 leu2$^-$ (Δ0 or Δ1) met15Δ0 trp1Δ63 ura3-52::pRS406-set1-G951A set1Δ::TRP1* Ty1*his3AI-236 ADE2::TELVR*	G951A

Protein extraction and Western blotting

Whole cell extracts of the Set1 variant strains were prepared and then analyzed by Western analysis to determine the levels of K4-methylated histone H3, as described in Mueller *et al*., (2). Primary antibodies used were: α-K4-dimethyl H3 (Lot No. 26335, Millipore, 1:3,333), α-K4-trimethyl H3 (Lot No. 77499, Abcam, 1:3,333), α-histone H3 (Lot No. 206549, Abcam, 1:5,000). HRP-conjugated secondary antibodies (Promega) were used at 1:????.

Analysis of rDNA silencing

Loss of silencing at the rDNA locus was measured using the Pol II-transcribed Ty1*his3AI* gene located in the rDNA locus. To quantify the frequency of Ty1*his3AI* transposition events, a liquid culture assay was used. First, a 5 ml culture in YPADT media was grown overnight at 30°C. Ten microliters of the overnight culture was used to inoculate each of three tubes containing 5 ml YPADT liquid media. After the cultures were incubated with agitation at 20°C for four days, the cell density was estimated by plating a small volume (Be specific - what volume was plated?) of each culture on YPADT plates, incubating the plates at 30°C for three days, and then counting the number of colonies that grew. In addition, cells in each of the three 5 ml cultures were pelleted by centrifugation and plated onto media lacking histidine (SC-His plates). The plates were incubated at 30°C for five days and the number of His^+ colonies was counted.

Analysis of telomeric silencing

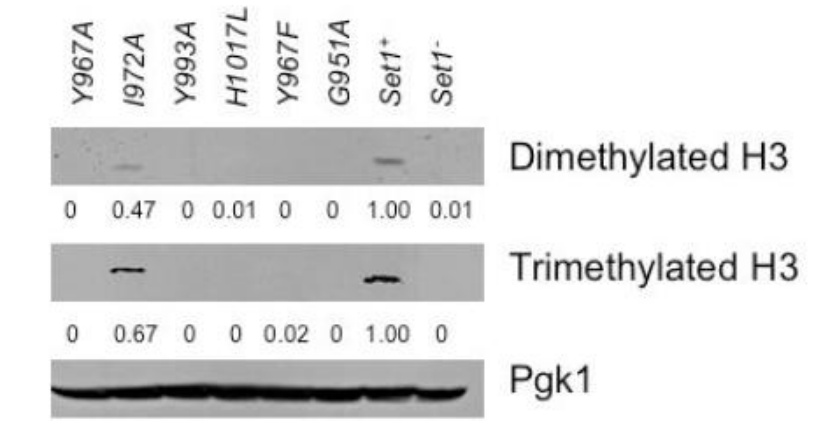
Telomeric silencing was measured in the yeast cells containing the Set1 variants by monitoring expression of an *ADE2* gene inserted in the right telomere on chromosome V. The *ADE2* gene encodes an enzyme necessary for adenine biosynthesis. If the *ADE2* gene is silenced, a red intermediate accumulates in the cells making the colonies appear red or dark pink. If the ADE2 gene is expressed (not silenced) the colonies are white. Each Set2 variant yeast strain was grown overnight at 30°C in 5 ml of YPADT. Twenty microliters of the overnight culture was used to make a series of dilutions. Five microliters of each dilution was spotted onto SC media with minimal adenine. As the colonies grew, they depleted the media of adenine and then began to biosynthesize adenine. Yeast strains that are competent for gene silencing at the telomere are red or dark pink. The strains that have lost silencing at the telomere are white or light pink colored. Figure 4 shows a representative telomeric silencing assay.

RESULTS

Western Analysis to measure the levels of K4 methylated histone H3

If an amino acid substitution affects Set1's ability to methylate histone H3 at K4, we would expect to see a reduction in the levels of K4 methylated histone H3 in whole cell extracts. Western blot analysis was used to test for the levels of K4 methylated histone H3 in whole cell extracts from cells with wild type Set1 and cells with the Set1 amino acid substitution variants (Figure 3.).

Fig. 3. **Loss of K4 methylated H3 in five of six Set1 variant strains**



Numbers under each panel are the ratio of the level of K4-methylated H3 in the Set1 variants relative to wild type cells.

As shown in Figure 3, five of the six Set1 amino acid substitution variants showed a complete loss of K4 dimethylated and K4-trimethylated histone H3. The I972A substitution variant showed a reduction in both forms of K4 methylated histone H3. The controls for the Western analysis are: $Set1^{+}$, extract

from cells with wild type Set1 and Set1$^-$, extract from cells lacking Set1.

Quantitative Analysis of the Loss of Silencing at the rDNA Locus

Loss of silencing at the rDNA locus was measured using a Ty1*his3AI* gene located in the rDNA locus. Ty1*his3AI* is a retrotransposable element that contains a backwards *his3AI* gene. When the Ty1*his3AI* element is expressed, the artificial intron (AI) can be spliced out of the *his3AI* gene. When the spliced Ty1*HIS3* mRNA is reverse-transcribed to make a cDNA, it will contain the sequences for a functional *HIS3* gene (6). The Ty1*HIS3* cDNA then inserts itself into the chromosome at a new location, making His$^-$ cells into His$^+$ cells. The frequency of retrotransposition of Ty1*his3AI* is directly related to the degree of silencing in the rDNA locus (6). The results of the rDNA silencing assays are summarized in Table 2. The frequency of His$^+$ cells was calculated using the following equation.

$$\frac{\text{Number of His}^+ \text{ cells}}{\text{Total number of cells}}$$

Table 2. **Loss of silencing at rDNA locus**

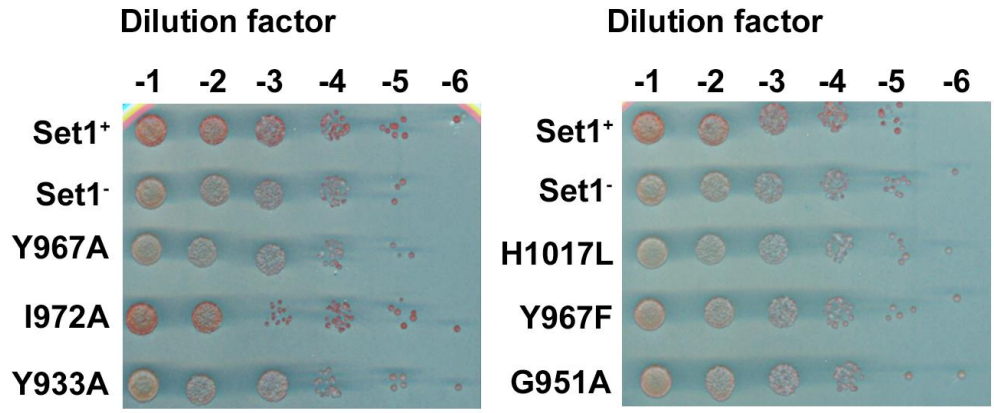
Set1 Variant	Fold Increase in His$^+$ cells (average of three independent experiments)
Set$^+$	**1.00**
Set$^-$	**32.45**
Y967A	**30.10**
I972A	**3.68**
Y993A	**17.80**
H1017L	**33.89**
Y967F	**2.02**
G951A	**14.50**

The frequency of His$^+$ cells in the Set1 variants was normalized to the level in wild-type cells (Set1$^+$). The I972A and Y967F variants show nearly wild type rDNA silencing phenotypes suggesting that these amino acid substitutions do not impede silencing of Pol II transcription at the rDNA locus.

Qualitative Analysis of Silencing at a Telomere and the rDNA Locus

In addition to the Ty1*hisAI* gene at the rDNA locus, all of the strains used in this study have an *ADE2* gene at the right telomere on chromosome V. The strains were grown and serially diluted as described in the Methods section and then plated on the SC-minimal adenine plates (Figure 4).

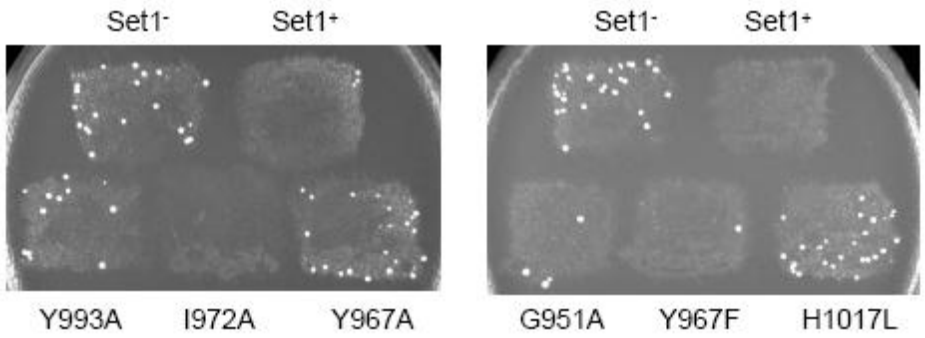
Fig. 4. **Loss of silencing at a telomere in five of six Set1 variants**



The amino acid substitution variant I972A retained wild type telomeric silencing at the right telomere of chromosome V. The five other amino acid substitutions variants showed a loss of silencing of the *ADE2* gene at the telomere.

Qualitative analysis of gene silencing at the rDNA locus was done by patching each strain onto YPADT media plates and incubating the plates for four days at 20°C. Then, the patches of cells were transferred by replica-plating onto a SC-His plate and incubated for three days at 30°C. Cells that have lost gene silencing at the rDNA locus show evidence of transposition of the Ty1*his3AI* element by forming colonies on media lacking histidine (SC-His). Figure 5 shows an image of the SC-His plates.

Fig. 5. **Loss of silencing at rDNA locus in five of the six Set1 variants**



Similar to the changes in telomeric silencing, the amino acid substitution variant I972A has a wild type rDNA silencing pattern. The Y993A, Y967A, and H1017L substitutions show a loss of silencing phenotype at the rDNA locus. The G951A and Y967F substitutions showed a partial loss of silencing that does not agree with the results of the quantitative rDNA silencing assay results. Since this patch assay is qualitative, the results of the quantitative assay are a more reliable representation of the gene silencing phenotype at the rDNA locus.

DISCUSSION

The Set1 protein in budding yeast is a member of the SET domain histone methyltransferease family of enzymes. *S. cerevisiae* has six SET domain containing proteins encoded in its genome(8). Set1 is the closest to the trithorax SET domain found in *Drosophila melanogaster* (8). SET7 and SET9 are histone methyltransferases from the SET domain family found in humans. The human SET7/SET9 proteins only mono and dimethylate histone H3 due to steric hindrances, while the Set1 protein of *S. cerevisiae* can mono-, di- and trimethylate K4 of histone H3 (12).

In four of the Set1 mutants characterized in this study there is a complete loss of methylation of K4 of histone H3. These variants will be analyzed further to determine which step in the Set1 methylation reaction mechanism is defective. For Set1 to function properly it must fold correctly, form a complex with seven other proteins, bind histone H3, and bind a methyl group donor (SAM). A defect in any one of these processes is expected to reduce the methyltransferase activity of Set1.

The H1017L variant has lost the ability to methylate histone H3. It is possible that this protein fails to form a catalytically active tertiary structure. In the human SET7/9 protein, this histidine residue is involved in forming a

"pseudo knot", which forms a binding site for the methyl donor AdoMet (12). A mutation at this residue may affect the structure of the methyl donor binding site in *S. cerevisiae* Set1.

In one of the mutants (I972A), there is a reduction of histone H3 methylation, which suggests that the isoleucine at position 972 is not essential for the catalytic activity of Set1. Since isoleucine, a hydrophobic residue, was replaced by alanine, I expected the protein to fold incorrectly. Hydrophobic amino acid residues avoid exposure to aqueous environments and bury themselves deep inside of proteins. Alanine residues do not do this so we expected this region might be improperly folded. The results show I972 is not critical for the structure of Set1.

The Set1 variants at the Y967 position (Y967A and Y967F) help address a question of specificity. The Y967A variant lost the ability to methylate K4 of histone H3, as shown by Western analysis, and had a silencing defect consistent with loss of silencing at the telomere and the rDNA locus. In the Y967F variant, K4-methylated histone H3 could not be detected, and there was evidence of loss of silencing at the telomere. Interestingly, in the Y967F variant silencing at the rDNA locus was maintained. These results suggest that a phenyl ring is necessary for histone H3 methylation but the hydroxyl group, which may function in substrate specificity or methylation efficiency, is dispensible for Pol

II gene silencing at the rDNA. The telomeric *ADE2* gene must be constantly silenced to allow the accumulation of the red pigment in cells. One possibility is that the level of histone H3 methylation in Y967F Set1 variant is too low to maintain telomeric silencing but is sufficient for gene silencing at the rDNA locus.

In summary, these studies have provided us with new information about the roles of specific amino acids in the SET domain of Set1. By characterizing the catalytic domain of Set1 using amino acid substitution variants, we will acquire a better understanding of Set1 and the Set1-like proteins in human cells.

REFERENCES

1. Dehe, P-M., and Geli. V. (2006) *Biochem. Cell Biol.***84,** 536-548
2. Mueller, J. E., Canze, M., and Bryk, M. (2006) *Genetics.* **173,** 557-567
3. Cairns, B. R., and Schlichter, A. (2005) *The EMBO Journal* **24,** 1222-1231
4. Robzyk, K., and Kassir, Y. A. (1992) *Nucleic Acids Research* **20,** 3790
5. Huang, Y. (2002) *Nucleic Acids Research* **30,** 1465-1482
6. Bryk, M., Banerjee, M., Murphy, M., Knudsen, K. E., Garfinkel, D. J., and Curcio, M. J. (1997) *Genes and Development* **11,** 255-269
7. Li, C., Mueller, J. E., and Bryk, M. (2006) *Mol. Biol. of the Cell* **17,** 3848-3859
8. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W.W.M.P., Wilm, M., asland, R., and Stweart, A.F. (2001) *The EMBO Journal* **20,** 7137-7148
9. Rose, M. D., Winston, F., and Hietter, P. (1990) *Methods in Yeast Genetics: A Laboratory Course Manual.* Cold Spring Harbor Laboratory Press, Plainview, NY.
10. C. B., Brachmann, Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998) *Yeast* **14**, 115–132
11. Hoffman, C. S. and Winston, F. (1987) *Gene* **57**, 267-272
12. Cheng, X. and Zhang, X. (2007) *Mutation Research* **618,** 102-115

CURRICULUM VITA

MORGAN CHATEAU
lukka@tamu.eud

EDUCATION

Texas A&M University	Bachelor's of Science Degree
College Station, Texas	Double Major: Biochemistry and Genetics
Graduation: May 2008	GPR: 3.59/4.0

RESEARCH

TAMU - Department of Molecular and Cellular Medicine

- Thesis: *Altered Amino Acid Sequence Affects Amyloid Formation in Aβ(25-35)*
- TAMU Department for Undergraduate Research
- TAMU Student Research Week 2007 & 2008- Poster Presenter
- Work done under Dr. J.M. Scholtz, January 2006 – May 2007

TAMU - Department of Biochemistry and Biophysics

- Thesis: *Effect of Amino Acid Substituions in Set1 on Histone H3 Methylation and Gene Silencing in Saccharomyces cerevisiae*
- Work done under Dr. M. Bryk , January- May 2007 and August 2007 - present

University of Texas Southwestern Medical Center - Department of Pharmacology

- Summer Undergraduate Research Fellows
- Work done under Dr. G. Suel, June - August 2007

AWARDS & ACCOMPLISHMENTS

TAMU Honors Department

- Honors University Research Fellows
- Foundation Honors
- University Honors

Lechner Scholarship

- Four year scholarship

SKILLS

Familiar with

- PCR, Electrical and Heat Competent Transformations, Gel Electroporesis and Imaging, Ion-exchange and Size Exclusion Column Purification, Spectrofluormetery, Western Analysis, Plasmid and Yeast Genome Extraction