12-2011

DEFINITION OF THE LANDSCAPE OF CHROMATIN STRUCTURE AT THE FRATAXIN GENE IN FRIEDREICH’S ATAXIA

Eunah Kim

Follow this and additional works at: http://digitalcommons.library.tmc.edu/utgsbs_dissertations

Part of the Biochemistry Commons, Cell Biology Commons, Molecular and Cellular Neuroscience Commons, and the Molecular Biology Commons

Recommended Citation
http://digitalcommons.library.tmc.edu/utgsbs_dissertations/206
DEFINITION OF THE LANDSCAPE OF CHROMATIN STRUCTURE AT THE FRATAXIN GENE IN FRIEDREICH’S ATAXIA

by

Eunah Kim, M.S.

APPROVED:

_________________________________________
Supervisory Professor: Sharon Y. R. Dent, Ph.D.

_________________________________________
Michelle C. Barton, Ph.D.

_________________________________________
Pierre McCrea, Ph.D.

_________________________________________
Xiaobing Shi, Ph.D.

_________________________________________
Min Gyu Lee, Ph.D.

APPROVED:

_________________________________________
Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston
DEFINITION OF THE LANDSCAPE OF CHROMATIN STRUCTURE AT THE FRATAxin GENE IN FRIEDREICH’S ATAXIA

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston and
The University of Texas M. D. Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

By

Eunah Kim, M.S.
Houston, Texas

December 2011
Acknowledgements

I would like to first thank my advisor Dr. Sharon Dent. I deeply appreciate her support throughout my graduate career. Her silent and strong support let me always think a big picture for my life. Her supports were not only training of research, writing or presentation, but also making me being supported by somebody for personal life in out of home country. Finally, I am extremely happy that I have my role model for my life.

I am extremely grateful to committee members Dr. Michelle Barton and Dr. Pierre McCrea for both advisory and supervisory, Dr. Jean-Pierre Issa and Dr. Stephanie Watowich for advisory, Dr. Gregory May, Dr. Andreas Bergmann, Dr. Gilbert Cote, and Dr. Zheng Chen for examinary, Dr. Xiaobing Shi, Dr. Min-Gyu Lee for supervisory. I would like to especially thank both Dr. Michelle Barton and Dr. Pierre McCrea for not only their insight during committee meetings, but also caring me outside of committee meetings. The wonderful letters of recommendation they have provided me over the years gave me smooth transition from a student to a postdoc position.

I would like to thank all the past and current members of the Dent lab. I am very grateful to Medelene Commbe, a past lab manager, provided to me all the help when I joined the lab. I am grateful to Dr. Renee Chosed who was a postdoc of the Dent lab when I rotated and taught me yeast genetics as a rotation project. She was the second person in the Dent lab that I decide to join the Dent lab with. I would like to specially thank Dr. Marek Napierala who taught me so much about FRDA and was so willing to answer my stupid questions for my Ph.D. project. With all his help and troubleshooting, I could learn how
to manage my research. I am extremely grateful to “Dr.” EJ Chen for not only her help, but also her friendship as my bench and graduation mate. I am indebted to Jill Butler for all her insights, reading of so many ugly drafts, hearing of many complaints and crying, and other help too countless to list. I would also like to thank other Dent Lab members: Jay Zhang, Lou Ramagli, Elizabeth McIvor, Calley Hirsch, John Latham, Rebecca Lewis and Marenda Wilson-Pham for their friendship and advice. I would like to thank the current Dent lab members Andy Salinger, Boyko Atanassov, Ula Polak, Lia Koutelou for their help and friendship, and Andria Schibler for joyful chattings as well as teaching me American culture, and the lab members in Science Park for their remote support.

I am sincerely grateful to the Genes and Development program and its coordinator, Elisabeth Lindheim for all her support to make me comfortable in outstanding program. I am extremely appreciative to the Schissler Family Foundation for their funding support during my graduate studies.

Finally, I would like to thank my family for all their love and support. Without their love and support, I could not go through the last 4 years along in foreign country and have this glorious moment for my life, which I had thought it never happen.
DEFINITION OF THE LANDSCAPE OF CHROMATIN STRUCTURE AT THE
FRATAxin GENE IN FRIEDREICH’S ATAXIA

Publication No.__________

Eunah Kim, M.S.

Supervisory Professor: Sharon Y. R. Dent, Ph.D.

Friedreich’s ataxia (FRDA) is caused by the transcriptional silencing of the frataxin
(FXN) gene. FRDA patients have expansion of GAA repeats in intron 1 of the FXN gene
in both alleles. A number of studies demonstrated that specific histone deacetylase
inhibitors (HDACi) affect either histone modifications at the FXN gene or FXN
expression in FRDA cells, indicating that the hyperexpanded GAA repeat may facilitate
heterochromatin formation. However, the correlation between chromatin structure and
transcription at the FXN gene is currently limited due to a lack of more detailed analysis.
Therefore, I analyzed the effects of the hyperexpanded GAA repeats on transcription
status and chromatin structure using lymphoid cell lines derived from FRDA patients.
Using chromatin immunoprecipitation and quantitative PCR, I observed significant
changes in the landscape of histone modifications in the vicinity of the GAA tract in
FRDA cells relative to control cells. Similar epigenetic changes were observed in GFP
reporter construct containing 560 GAA repeats. Further, I detected similar levels of FXN
pre-mRNA at a region upstream of hyperexpanded GAA repeats in FRDA and control cells, indicating similar efficiency of transcription initiation in FRDA cells. I also showed that histone modifications associated with hyperexpanded GAA repeats are independent of transcription progression using the GFP reporter system. My data strongly support evidence that \textit{FXN} deficiency in FRDA patients is consequence of defective transition from initiation to elongation of \textit{FXN} transcription due to heterochromatin-like structures formed in the proximity of the hyperexpanded GAAs.
# Table of Contents

Chapter 1. Introduction and Background................................................................. 1

1. FRDA.................................................................................................................... 2
   - Clinical and pathological aspects........................................................................ 2
   - The role of FXN in mitochondria ...................................................................... 3
   - Differential expression of the FXN gene .......................................................... 5
   - Mutation at the FXN gene in FRDA patients (Expansion of GAA repeats at the FXN gene in FRDA)................................................................. 8
   - Expansion of the GAA triplet............................................................................. 11
   - The effect of the GAA repeats on gene expression............................................ 12
   - Treatments of FRDA ....................................................................................... 13
   - Models for study of FRDA ................................................................................ 14

2. Epigenetics regulation of FXN expression......................................................... 21
   - Regulation of Chromatin structures .................................................................. 22

Chapter 2. Materials and Methods........................................................................ 37

Chapter 3. Defining of chromatin structure in the vicinity of the hyperexpanded GAA repeats at the FXN gene ................................................................. 46
   - GAA repeat expansion and FXN deficiency in the lymphoid cell lines............ 47
   - Distribution of histone modification at the FXN gene in the lymphoid cell lines... 53
   - Figure 13 ............................................................................................................ 68
   - Distribution of histone modification in the GFP_GAA reporter system........... 69

Chapter 4. Determination of the effect of the GAA repeats in transcription in the context of chromatin structure ........................................................................ 79
   - In lymphoid cell lines ...................................................................................... 80
   - In the GFP_GAA system .................................................................................. 100

Chapter 5. Discussion .............................................................................................. 109

Chapter 6. Significance and Future Directions..................................................... 119

Bibliography ........................................................................................................... 133

Vita.......................................................................................................................... 152
List of Figures

Figure 1. Schematic diagram of the human \(FXN\) gene......................................................... 6

Figure 2. Mutation of the \(FXN\) gene in FRDA patients................................................................. 9

Figure 3. The GFP_GAA system ........................................................................................................ 19

Figure 4. Diagram of histone modifications at the amino-terminal tails of histone H3 and H4 in each chromatin structure (euchromatin and heterochromatin)................. 23

Figure 5. Distribution of histone modifications across genes to define chromatin structures. ............................................................................................................................... 28

Figure 6. Hyperexpanded GAA repeats at the \(FXN\) gene in the lymphoid cell lines 49

Figure 7. \(FXN\) transcription is impaired in FRDA lymphoid cell lines. .................. 51

Figure 8. Fragmented chromatin extracted from the lymphoid cell lines and the GFP_GAA reporter system.................................................................................................................. 54

Figure 9. Hyperexpanded GAA repeats have no association with distribution of total histone H3 ................................................................. 56

Figure 10. H3K9/14ac is significantly repressed at the \(FXN\) gene in FRDA cells.... 59

Figure 11. H3K9me3 is significantly increased at the \(FXN\) gene in FRDA cells...... 61

Figure 12. H4K20me3 is significantly increased at the \(FXN\) gene in FRDA cells..... 63

Figure 13. H3K27me3 is enriched at the \(FXN\) gene in FRDA cells. ............................ 67

Figure 14. Validation of the presence of 560 GAA repeats in the GFP_GAA system ................................................................................................................................. 70

Figure 15. \(GFP\) transcription is impaired in the GFP_(GAA•TCC) \(_{560}\) cells. ....... 72
Figure 16. H3K9/14ac is significantly decreased at the GFP gene in the presence of GAA repeats. ................................................................. 75

Figure 17. H3K9me3 is enriched at the GFP gene in the presence of longer GAA repeats. ................................................................................................................. 77

Figure 18. FXN deficiency appears at the region downstream of the hyperexpanded GAA repeats in FRDA cells ................................................................................................................. 81

Figure 19. Distribution of total RNA pol II is repressed at the FXN gene in FRDA cells. ................................................................................................................................. 84

Figure 20. H3K4me2 is decreased at the region upstream of the GAA repeat, not promoter region in FRDA cells................................................................................................................. 87

Figure 21. Levels of H3K4me3 are not affected at promoter region in FRDA cells. 89

Figure 22. H3K4me3 at the region upstream of the GAA repeats in FRDA cells is associated with impaired FXN transcription elongation................................................. 92

Figure 23. H3K36me3 are repressed at the entire FXN gene in FRDA cells relative to unaffected cells..................................................................................................................... 96

Figure 24. H3K79me2 is significantly decreased in FRDA cells........................................ 98

Figure 25. H3K4me3 is affected at the region downstream of the GAA repeats in the presence of the GAA repeats............................................................................................................. 101

Figure 26. H3K79me2 is affected at the region downstream of the GAA repeats in the GFP_(GAA•TTC)_560 cells.................................................................................................................. 103

Figure 27. The tract of GAA repeats is sufficient to form the heterochromatin-like structure in the absence of transcription, which lead to inhibition of transcription elongation.............................................................................................. 107
Figure 28. A summary of distributed histone modifications at the human $FXN$ gene.

........................................................................................................................................................................ 121

Figure 29. A working model of heterochromatin-like structure formation at the $FXN$ gene in FRDA ........................................................................................................................................................................ 131
List of Tables

Table 1. Primers used for analyses in the GFP_GAA reporter system ...................... 44

Table 2. Primers used for analyses in the lymphoid cell lines ................................. 45
Chapter 1. Introduction and Background

Some parts of this dissertation were modified from the following journal article.
1. FRDA

Clinical and pathological aspects

Friedreich’s ataxia (FRDA) is an degenerative neuro-muscular disease that affects 1 in 50,000 individuals in the United States (Delatycki, Williamson and Forrest, 2000). The features of FRDA are recessively inherited with an incidence of one in 60 to 100 depending on the ethnic group (Cossee, Schmitt, Campuzano, Reutenauer, Moutou, Mandel and Koenig, 1997). FRDA is characterized by progressive ataxia, including uncoordinated gait and limb movements, weakened muscle strength, and diminished senses of position and vibration. FRDA involves degeneration of the posterior columns of the spinal cord that is extended to the brain stem. FRDA patients also exhibit atrophy in the spinocerebellar tract, a set of axonal fibers in the spinal cord terminating in the cerebellum. In the dorsal root ganglion within the spinal cord, large primary sensory neurons are also degenerated in FRDA patients. Quantitative analysis reveals that contact between Purkinje cells in the cerebellum and dendrites are decreased, while the cerebellar cortex is not affected in patients of FRDA. These phenomena indicate that FRDA patients are most likely defective in providing information from the spinal cord through the spinocerebellar tract, which lead loss of position and vibration sense. Additionally, hypertropic cardiomyocytes are observed in approximately 80% of FRDA patients (Campuzano, Montermini, Lutz, Cova, Hindelang, Jiralerspong, Trottier, Kish, Faucheux, Trouillas, Authier, Durr, Mandel, Vescovi, Pandolfo and Koenig, 1997). Such cardiomyocytes contain abnormal intracellular iron deposits, suggesting that this feature of FRDA is associated with defective function of iron metabolism.
The role of FXN in mitochondria

The FRDA disease gene is genetically mapped to chromosome 9q13 and encodes the Frataxin protein (FXN), a nuclear encoded mitochondrial protein (Campuzano, Montermini, Lutz, Cova, Hindelang, Jiralerspong, Trottier, Kish, Faucheux, Trouillas, Authier, Durr, Mandel, Vescovi, Pandolfo and Koenig, 1997; Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996). FRDA is characterized by insufficient levels of FXN mRNA and protein in all examined tissues (Campuzano, Montermini, Lutz, Cova, Hindelang, Jiralerspong, Trottier, Kish, Faucheux, Trouillas, Authier, Durr, Mandel, Vescovi, Pandolfo and Koenig, 1997). Most functions of FXN have been demonstrated from yeast since yeast is viable without the yeast FXN homolog, Yfh1p, while FXN is an essential protein in mouse (Adinolfi, Iannuzzi, Prischi, Pastore, Iametti, Martin, Bonomi and Pastore, 2009; Bulteau, O'Neill, Kennedy, Ikeda-Saito, Isaya and Szweda, 2004; Cossee, Puccio, Gansmuller, Koutnikova, Dierich, LeMeur, Fischbeck, Dolle and Koenig, 2000). Yfh1p plays a role in the storage of iron within mitochondria (Babcock, de Silva, Oaks, Davis-Kaplan, Jiralerspong, Montermini, Pandolfo and Kaplan, 1997). During iron homeostasis, FXN acts as a chaperone (Bulteau, O'Neill, Kennedy, Ikeda-Saito, Isaya and Szweda, 2004). Negatively charged FXN binds ferrous iron to balanced levels of iron (Dhe-Paganon, Shigeta, Chi, Ristow and Shoelson, 2000). FXN has a ferroxidase activity when the concentration of ferrous iron is low in mitochondria leading to iron detoxification. Insufficient levels of FXN expression result in abnormal biosynthetic processes through a defective utilization of iron such as
decreased activities of iron-sulfur cluster (ISC) enzymes, accumulation of iron in mitochondria, and increased susceptibility to oxidative stress (Wilson, 2003). Oxidative stress is caused by oxygen excess in mitochondria due to an impaired ability to maintain homeostasis by detoxifying oxidative molecules. Decreased activity of ISC is associated with a dysfunction of respiratory chain leading to exacerbation of oxidative stress conditions (Jensen, Sanchez, Srinivasan, Valentine and Culotta, 2004). Oxidative stress observed in FRDA patient possibly results from FXN deficiency through impaired ISC in mitochondria. However, the underlying mechanisms of FRDA have not yet been fully elucidated. FXN is evolutionarily conserved from yeast to humans, in particular in the C-terminal region. Interestingly, an exogeneous protein consisting of the C-terminus of human FXN fused to the N-terminus of Yfh1p complements the phenotype of Yfh1p deficient yeast (Wilson and Roof, 1997). Yfh1p precursor is processed to a 122-amino acid protein through cleavage by the mitochondrial processing protease (MPP) (Gordon, Shi, Dancis and Pain, 1999). Human FXN is similarly processed to a 128-amino acid protein from a 210-amino acid precursor. A N-terminal portion of FXN contains a mitochondrial signal sequence that is removed by the MPP. Unlike Yfh1p, FXN is processed through two-step cleavages (Schmucker, Argentini, Carelle-Calmels, Martelli and Puccio, 2008). The first cleavage occurs between Gly41 and Leu42 by MPP, generating a 169-amino acid intermediate protein in mitochondria. The second cleavage between Lys80 and Ser81 brings into being the 130-amino acid mature form of FXN. Only the final mature 130-amino acid form is functional in mitochondria. Ectopic expression of the mature form of FXN efficiently rescues the lethality of FXN null fibroblasts (Schmucker, Argentini, Carelle-Calmels, Martelli and Puccio, 2008).
Consistently, residues 90-210 of FXN have been demonstrated to be rigid by crystallography and NMR structure (Dhe-Paganon, Shigeta, Chi, Ristow and Shoelson, 2000; Musco, Stier, Kolmerer, Adinolfi, Martin, Frenkiel, Gibson and Pastore, 2000). This suggests that C-terminal region of FXN is functional as mentioned above while N-terminal region is responsible for localization of FXN into mitochondria. Indeed, residues 90-210 correspond to an $\alpha_1$-helix and $\beta_1$-sheet of the core structure of FXN that has potential to either function to donate iron to its partner or to form multimers (Dhe-Paganon, Shigeta, Chi, Ristow and Shoelson, 2000). Although the processed forms of $FXN$ are associated with FRDA pathology, it is noteworthy that defects in the process of FXN maturation are not. This emphasizes that enhancement of FXN transcription will be critical for curing FRDA.

**Differential expression of the $FXN$ gene**

Since there is no evidence that FXN RNA is abnormally processed to make mature $FXN$ RNA in FRDA, insufficient levels of FXN in FRDA likely result from the stage of $FXN$ transcription. The $FXN$ gene consists of 7 exons (Figure 1). The major transcripts contain exon 1 to 5a and are pronounced in all examined tissues of the patients (Kostrzewa, Klockgether, Damian and Muller, 1997). Other isoforms are occasionally detected in the heart. The $FXN$ gene is differentially expressed in various tissues. The human $FXN$ gene is mainly expressed in primary sites of degeneration in FRDA patients such as liver, heart and CNS, particularly in the spinal cord (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996).
Figure 1. Schematic diagram of the human \textit{FXN} gene.

Black bars represent exons (exon1, 2, 3, 4, 5a, and exon 5b). Light gray bar represents non-coding exon that is involved in regulating the expression of the gene and does not lead to additions to the protein sequence. Exons 5a and 5b are alternatives and result in slightly different protein sequences. Exon1–5a produce major \textit{FXN} transcripts. The \textit{FXN} gene is displayed in correct proportion.
Human Frataxin (Chromosome 9: 71,650,175-71,715,094, forward strand)
The mouse *FXN* gene is differentially expressed during development. Expression is gradually increased from embryonic day 10 (E10) to later stages, appearing the highest at E14. In the spinal cord, a characteristic pattern of *FXN* expression becomes evident at E12.5, particularly in the thoracolumbar region that is most likely composed of large neuronal cells in the DRG (Jiralerspong, Liu, Montermini, Stifani and Pandolfo, 1997). In the adult brain, expression is restricted to the ependymal layer, chorid plexus, and granular layer of the cerebellum. This pattern indicates that FXN expression in various tissues is correlated with FRDA pathology, suggesting that changes in FXN transcription is likely responsible for FRDA.

**Mutation at the *FXN* gene in FRDA patients (Expansion of GAA repeats at the *FXN* gene in FRDA)**

Approximately 98% of FRDA patients are characterized by abnormal GAA repeat expansions within intron 1 of the *FXN* gene in both alleles (Figure 2), while 2% of FRDA patients have either point mutations or missense mutations in the *FXN* gene that act as null mutations (Bidichandani, Ashizawa and Patel, 1998; Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996; Ohshima, Sakamoto, Labuda, Poirier, Moseley, Montermini, Ranum, Wells and Pandolfo, 1999). The GAA repeat at the *FXN* gene is detected in the middle of an *Alu* repeat element. In addition, a polymorphic mononucleotide tract of adenines (poly A tract) was identified in
Figure 2. Mutation of the FXN gene in FRDA patients.

Expansion of GAA repeats is located within intron 1 of the FXN gene in FRDA patients. Black bars represent exon 1 and 2, while light gray bar represents non-coding exon 6. More than 1,700 GAA repeats are observed in FRDA patient, whereas less than 66 GAA repeats is detected in unaffected individuals. The region between exon 1 and exon 2 is displayed in correct proportion.
Figure 2

Unaffected  

FRDA  

> 1,700 GAA repeats  

66 GAA repeats  

1Kb
the vicinity of the GAA triplet repeat (Monticelli, Giacchetti, De Biase, Pianese, Turano, Pandolfo and Cocozza, 2004). Reduced levels of the $FXN$
gene expression in FRDA patients are correlated with the hyperexpanded tract of
repeated GAA triplets (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996; Grabczyk and Usdin, 2000a). In FRDA patients, the GAA
tract frequently consists of more than 1500 triplets, whereas unaffected individuals have
66 or fewer repeats at the $FXN$ gene (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996; Pandolfo, 1998; Pianese, Turano, Lo Casale, De Biase, Giacchetti, Monticelli, Criscuolo, Filla and Cocozza, 2004). About
85% of unaffected individuals have fewer than 12 GAA repeats whereas the remainder
has 12 to 33 repeats.

**Expansion of the GAA triplet**

Expansion or contraction in size of the GAA repeat tracts is commonly observed during
intergenerational transmission from parents to children. Although the underlying
mechanism for how the size of the repeats is expanded has not yet been demonstrated, it
has been suggested that DNA metabolism such as repair, replication, and recombination
is involved in the flexibility of the GAA triplet size among generations. The mismatch
repair (MMR) pathway is involved in inheritance and stability of the pathological repeats
among generations (Krasilnikova, Kireeva, Petrovic, Knijnikova, Kashlev and Mirkin,
In longer GAA repeats, single-stranded DNA is exposed to form hairpin DNA structures. MMR machinery is recruited to the hairpin structure, which leads to repeat expansion (Krasilnikova, Kireeva, Petrovic, Knijnikova, Kashlev and Mirkin, 2007). During DNA replication, exposed single stranded DNA containing GAA repeats enables folding into secondary DNA structures, which cause expansion of GAA repeats through misalignment of repetitive DNA (Ohshima and Wells, 1997). Recombination pathways also contribute to expansion of the GAA repeats. Longer GAA repeats are associated with a high frequency of recombination in E.coli, which results in expansion of the GAA repeats (Napierala, Dere, Vetcher and Wells, 2004). However, the detailed mechanisms are not clear.

The effect of the GAA repeats on gene expression

Previous studies demonstrate that the long GAA tracts cause transcriptional pausing in *in vitro* as well as in cell culture systems (Grabczyk and Usdin, 2000b; Krasilnikova, Kireeva, Petrovic, Knijnikova, Kashlev and Mirkin, 2007). Oligonucleotides containing longer GAA repeats show abnormal migration mobility on agarose gels when compared to linear DNA consisting of random sequences (Ohshima, Montermini, Wells and Pandolfo, 1998). Abnormal DNA structures of constructs containing longer GAA repeats have been observed by nuclear magnetic resonance (LeProust, Pearson, Sinden and Gao, 2000). Polypurine/polypyrimidine DNA sequences containing mirror repeat symmetry form distinctive non-canonical structures. Either a pyrimidine or a purine rich single-stranded DNA is intertwined into the major groove of a double helix consisting of the hyperexpanded GAA repeats forming intermolecular interactions between the strands in
in vitro systems (Sakamoto, Chastain, Parniewski, Ohshima, Pandolfo, Griffith and Wells, 1999). Integration of oligodeoxyribonucleotides into longer GAA repeats block formation of the non-canonical DNA structure leading to restoration of full-length FXN transcripts in a concentration-dependent manner (Grabczyk and Usdin, 2000a). Small molecules recognizing specific DNA sequences such as polyamides de-repress FXN expression in cellular systems (Dervan and Edelson, 2003; Gottesfeld, 2007; Gottesfeld, Turner and Dervan, 2000). T7 RNA polymerase is impeded in the vicinity of GAA repeats in in vitro systems(Grabczyk and Usdin, 2000b). These results explain the correlation between size of the GAA repeats and defective transcription in FRDA.

Treatments of FRDA

Although the mechanism of how hyperexpanded GAA repeats down-regulate FXN expression in FRDA patient remains unclear, a number of studies have been carried out to identify compounds that enhance the expression of FXN in FRDA. To date, several compounds have been reported to increase FXN expression in cellular models, including hemin, butyric acid (Sarsero, Li, Wardan, Sitte, Williamson and Ioannou, 2003), and erythropoietin (Sturm, Stupphann, Kaun, Boesch, Schranzhofer, Wojta, Goldenberg and Scheiber-Mojdehkar, 2005) by unknown mechanisms.

As FXN deficiency in FRDA results in increasing free radicals accompanied by abnormal accumulation of iron in mitochondria, both antioxidants and iron chelators have been applied to patients as a therapy to alleviate symptoms of FRDA. To date, only idebenone, a synthetic analogue of coenzyme Q10, has been demonstrated in a phase III trial as free
radical chelator (a treatment pipeline in the Friedreich's Ataxia Research Alliance (www.fara.org)). Unfortunately, a recent study has reported that idebenone does not improve neurological symptoms of FRDA in pediatric patients (Lynch, Perlman and Meier, 2010).

Although the treatments mentioned above have improved the progress of FRDA, there are no effective treatments for curing FRDA. Enhancement of FXN expression from FXN cDNA harboring viral vectors partially alleviates FXN deficiency and defective response to oxidative stress (Fleming, Spinoulas, Zheng, Cunningham, Ginn, McQuilty, Rowe and Alexander, 2005; Gomez-Sebastian, Gimenez-Cassina, Diaz-Nido, Lim and Wade-Martins, 2007; Lim, Palomo, Mauritz, Gimenez-Cassina, Illana, Wandosell and Diaz-Nido, 2007). These studies suggest that restoring FXN levels is likely to be more effective for FRDA treatment than targeting downstream consequences of FXN deficiency. These studies make the point that understanding the underlying mechanisms of how FXN expression is affected by hyperexpanded tracts of GAA repeats in patients with FRDA is critical to developing therapeutic targets for curing FRDA.

Models for study of FRDA

As previously mentioned, an effective treatment has not yet developed for curing FRDA. Nevertheless, various systems have been applied to investigate the molecular pathology of FRDA. Each system has different purpose for study.

- Escherichia coli (E.coli)
E. Coli has been used to investigate the propensity of longer GAA repeats to form non-B-DNA structures that are associated with abnormal replication and defective transcription in a length-dependent manner (Grabczyk and Usdin, 2000b). Moreover, persistent RNA-DNA hybrids in the presence of longer GAA repeats was first identified in E. coli (Grabczyk, Mancuso and Sammarco, 2007). RNA transcribed from long GAA repeats is hybridized with exposed single-stranded DNA to form persistent RNA-DNA hybrids, which inhibit transcription. Therefore, E. coli has been useful to define how non-B DNA structures consisting of longer GAA repeats act as cis-inhibitors of DNA metabolism.

**Yeast**

Yeast is an appropriate model system to investigate the function of FXN since YFH1 deficient yeast survive (Wilson and Roof, 1997). Moreover, similar to human and mouse FXN, yeast FXN homologue contains a potential mitochondrial targeting sequence in the N-terminus of YFH1. Therefore, deletion of YFH1 in yeast strain provided evidence to demonstrate that FXN acts as an iron chaperone in mitochondria (Babcock, de Silva, Oaks, Davis-Kaplan, Jiralerspong, Montermini, Pandolfo and Kaplan, 1997).

**Mouse models**

Mouse models for FRDA have been generated to study the mechanism of the pathology and to test pharmacological therapies.

- FXN null/conditional KO/inducible KO

FXN null mice that have constitutive inactivation of FXN by homologous recombination are embryonic lethal, indicating that FXN is essential for development (Cossee, Puccio,
However, these mice did not exhibit defective iron metabolism during development, suggesting that lethality of the null embryos is independent of the iron metabolism defects caused. To further understand the lethality caused by knockout (KO) of FXN, conditional KO mice were generated using Cre-lox system specific to skeletal and cardiac muscle (Puccio, Simon, Cossee, Criqui-Filipe, Tiziano, Melki, Hindelang, Matyas, Rustin and Koenig, 2001). The conditional FXN KO mice were viable and resembled the phenotype of FRDA patients such as large sensory neuron dysfunction and deficient activity of mitochondria. Although this model did not exhibit susceptibility to oxidative stress, this phenomenon is also controversial in FRDA patients. This result suggests that the susceptibility to oxidative stress is not a major consequence of FXN deficiency. Therefore, these mice models that have insufficient levels of FXN without GAA repeats are appropriate to investigate the role of FXN and the pathological process of FRDA independent of the GAA repeats.

- GAA in mice

In an effort to define the effects of the GAA repeats in the pathology of FRDA, a transgenic mouse model (knock in-knock in (KIKI) mice) that contains a tract of GAA repeats was generated by insertion of 230 GAA into the first intron of the mouse Fxn gene and then mated with Fxn -/- mice (Miranda, Santos, Ohshima, Smith, Li, Bunting, Cossee, Koenig, Sequeiros, Kaplan and Pandolfo, 2002). The offspring (Fxn^{230/-}) have 25~36% reduction of FXN expression relative to wild type. Unfortunately, these mice did not phenocopy a FRDA in despite of reduced levels of Fxn caused by the insertion of the GAA repeats. This finding suggests that 230 GAA repeats are insufficient to stimulate
pathology of FRDA. Moreover, this study also suggests that the appropriate location of the GAA repeat at a given gene might be a major factor in exhibition of FRDA pathology, as the location of the repeat in the mouse Fxn gene is not comparable to the location of the GAA repeats in human FXN gene. Nevertheless, this is a first FRDA mouse model to exhibit FXN deficiency caused by the presence of the GAA repeat.

- YAC-GAA in mice

To overcome embryonic lethality of FXN KO model, transgenic mice were generated by introducing 190 or 190/90 GAA repeats in intron 1 of the human FXN gene within a human YAC clone in the background of FXN null mice (Al-Mahdawi, Pinto, Ruddle, Carroll, Webster and Pook, 2004; Pook, Al-Mahdawi, Carroll, Cossee, Puccio, Lawrence, Clark, Lowrie, Bradley, Cooper, Koenig and Chamberlain, 2001). These transgenic mice exhibited typical FRDA phenotypes such as coordination deficits and a progressive decrease in motor activity in younger mice (3 month old), indicating that location of the GAA repeat is considerable when compared to KIKI mice. Histological studies in this model showed similarities to FRDA patients such as large vacuoles in dorsal root ganglia cells and iron deposits in cardiomyocytes, whereas electrophysiological studies showed a mild, progressive peripheral neuropathy (Al-Mahdawi, Pinto, Varshney, Lawrence, Lowrie, Hughes, Webster, Blake, Cooper, King and Pook, 2006). Moreover, this model showed tissue-specific and age-dependent somatic instability specifically in the cerebellum and dorsal root ganglia (Clark, De Biase, Malykhina, Al-Mahdawi, Pook and Bidichandani, 2007). Especially, a HDACi enhances FXN expression accompanying with increased levels of acetylation on histones in the KIKI mice. This outcome is meaningful that removal of marks associated heterochromatin-like structures has potential to alleviate
FXN deficiency in FRDA. However, it is unclear that the observed phenomenon resulted from the GAA repeats rather than from the construct *per se*.

- **Lymphoid cell lines**

Patient-derived lymphoid cell lines immortalized by Epstein-Barr virus available in the Coriell Institute (http://www.coriell.org). Moreover, cells from unaffected control individuals that contain a range of the GAA repeats are available for comparison studies. Therefore, the lymphoid cell lines are perhaps the best system to address my hypothesis since the presence of expanded GAA repeats at the *FXN* locus does not result in degeneration of these lymphoid cell lines.

- **The GFP_GAA system**

To avoid effects from inherent characteristics of the *FXN* gene so that changes induced by the hyperexpanded GAA repeats *per se* can be determined, the GAA repeat derived from a FRDA patient (the GM16210 cell line from Coriell Institute) was engineered into an intronic region originating from the human *Pem1* gene that is known to be spliced efficiently during RNA processing (Figure 3). The *Pem* intronic region containing the GAA repeat is embedded into the middle of the GFP coding region to divide the GFP coding region. The CMV promoter controls expression of this construct to avoid inherent
Figure 3. The GFP_GAA system.

Schematic diagram of the GFP gene containing GAA repeats in the GFP_{(GAA\cdot TTC)} reporter system. The GFP gene was divided into two exons (gray boxes) by the intron containing 560 GAA repeats. This reporter system is expressed under control of the CMV promoter (black box) and the tetracycline operator/repressor (TetO)\textsuperscript{2} (white boxes) that allows regulating transcription initiation by tetracycline. BGH pA represent bovine growth hormone and polyadenylation signal that is used during mRNA processing.
Figure 3
characteristics of the FXN promoter. A tetracycline operator element was also inserted in the promoter region so that GFP expression is regulated by tetracycline treatment, reducing the possibility of genomic instability associated with transcription of repetitive DNA sequences. This reporter construct genetically recapitulates the location of the GAA repeats as characteristics of the endogenous FXN gene. Moreover, the GFP\_GAA system is validated as a model system to investigate molecular pathway associated with FXN expression changes (Soragni, Herman, Dent, Gottesfeld, Wells and Napierala, 2008). The insertion of the GAA repeats results in reduction of GFP mRNA. The intronic region in the proximity of the GAA repeat has increased histone modifications marking heterochromatin and decreased modifications marking euchromatin. Furthermore, small drugs that affect modifications on chromatin structure restore GFP expression in the GFP\_((GAA\textbullet\TTC)^{560}\) cell line.

Section 2 of this chapter will describe general information about epigenetics, which might play a role in down-regulating FXN transcription through the hyperexpanded GAA repeats in FRDA.

2. Epigenetics regulation of FXN expression

In eukaryotes, the genome is organized into arrays of nucleosomes composed of histones and DNA. The histones are classified as H1, H2A, H2B, H3, and H4. Two copies of H2A, H2B, H3, and H4 compose a core octamer around which is wrapped around 146 base pair (bp) DNA to form a nucleosome (Luger, Mader, Richmond, Sargent and Richmond, 1997). Chromatin is categorized into either euchromatin or heterochromatin,
depending on the conformation of structure, status of transcription, and post-translation modifications of histones. Configuration of chromatin structure is less packed in euchromatin. Moreover, transcription occurs more in euchromatin. In contrast, heterochromatin is characterized by condensed chromatin structure, rich in repetitive DNA, which is rarely transcribed in eukaryotes and plants (Grewal and Jia, 2007). Heterochromatin comprises most portions of both centromeres and telomeres in mammals. Genes localized in heterochromatin are most likely silenced (Grewal and Jia, 2007). This phenomenon suggests that gene expression is regulated by not only trans through transcription machinery but also cis through organization of chromatin structures.

**Regulation of Chromatin structures**

- **Post-translational modifications of Histones**

  Regulation of chromatin structure is involved in various cellular functions such as DNA repair, replication as well as transcription by specific covalent marks in the N-terminal tails of histones such as acetylation (ac) and methylation (me) (Li, Carey and Workman, 2007). Although a number of lysine residues have been identified as targets for modifications, this thesis will focus on well characterized sites of acetylation and methylation of histone H3 and H4 (Figure 4).

  Lysine residues in the amino-terminal regions of the histones are acetylated dependent on the status of transcription and on temporal and spatial conditions such as tissue specificity and development. Most acetylated histones are readily detected at transcriptionally active
Figure 4. Diagram of histone modifications at the amino-terminal tails of histone H3 and H4 in each chromatin structure (euchromatin and heterochromatin).

Each chromatin structure has distinct combination of histone modifications as shown in diagram. DNA wrapped around histone octamer consisting of pairs of H2A, H2B, H3 and H4. Each histone exhibits a flexible histone tail that can be dynamically modified. Shown here are histone modifications examined in this study. Blue oval represent a methyl group on lysine residue. Green oval represent acetyl group on lysine residue.
Figure 4
genes, particularly, in regulatory regions. Acetylation of lysines in histones neutralize the positive charge on histones and thereby decrease interactions between histones and DNAs. Therefore, transcription machinery is better able to access chromatin structures that are hyper acetylated. In contrast, lack of acetyl group on histone results in compaction of chromatin structure leading to gene silencing. Inhibition of histone deacetylation using HDACi is promising to enhance expression of a given gene. Indeed, many diseases such as neurodegenerative diseases and cancers have been treated with HDACi to alleviate symptoms through regulation of histone acetylation. In heterochromatin, acetylation is rarely detected as compared to euchromatin. Moreover, higher eukaryotic heterochromatin is associated with heterochromatin protein (HP1), first identified at pericentric heterochromatin in *Drosophila*. HP1 contributes to formation and maintenance of heterochromatin by multimerization of HP1 through ‘chromo domain’ at the N-terminal and ‘chromo shadow domain’ at the C-terminal domain of the HP1. Gene silencing associated with HP1 is relieved by inhibition of histone deacetylase, TSA, suggesting that hypoacetylation plays a role in HP1 mediated heterochromatinization. However, HP1 is also observed in euchromatin where genes are transcribed (Ayyanathan, Lechner, Bell, Maul, Schultz, Yamada, Tanaka, Torigoe and Rauscher, 2003), suggesting that additional proteins or histone modifications may be involved to HP1-mediated gene silencing. Acetylation of lysine of proteins can serve as binding sites for effector molecules that contain ‘bromodomains’. GCN5 is an acetyltransferase that is found in SAGA (Spt-Ada-Gcn5-acetyltransferase) complex. The GCN5 bromodomain-mediated interaction with acetylated histones promotes stability of SAGA complex at target genes, which maintain chromatin structure as euchromatin along
with increased levels of acetylation on histones. These results suggest a role of bromodomain as a modulator of chromatin modifying/remodeling complexes (Li and Shogren-Knaak, 2009).

Methylation of histones is involved in either active or repressive chromatin structures, depending on either the number of methyl groups present on the targeted residue and which residue is targeted in the histones (Li, Carey and Workman, 2007). While acetylation occurs only at lysine residues, either lysine or arginine residues can be methylated. Methylation at lysine 9 of histone H3 is enriched in heterochromatin, while acetylation at lysine 9 of histone H3 is enriched in euchromatin, where genes are actively transcribed. Methylation at lysine 9 of histone H3 prevents acetylation, promoting gene silencing. Methylation at lysine 9 of histone H3 at heterochromatin provides a binding site for HP1 protein to stimulate condensing of chromatin structures through the chromo domain and chromo shadow domain (Jenuwein and Allis, 2001). HP1 at the region where H3K9me3 are enriched act as anchor to recruit chromatin-remodeling complex to further modify histones. Although the majority of H3K9me3 is associated with repressed and silenced regions throughout the genome, this histone modification is also found in the coding regions of transcriptionally active genes (Vakoc, Mandat, Olenchock and Blobel, 2005). In centromeric heterochromatin, where gene expression is repressed, enrichment of H4K20me3 is accompanied by enrichment of H3K9me3 (Kourmouli, Jeppesen, Mahadevhaiah, Burgoyne, Wu, Gilbert, Bongiorni, Prantera, Fanti, Pimpinelli, Shi, Fundele and Singh, 2004; Schotta, Lachner, Sarma, Ebert, Sengupta, Reuter, Reinberg and Jenuwein, 2004).
During development, genes are appropriately either turning on or off at a given stage. Methylation at lysine 27 of histone H3 is well-characterized and is important in the regulation of gene expression during development. H3K27me3 is mainly detected in repressive chromatin structures such as the inactive X chromosome as well as in inducible genes that are required for a given stage of development. In embryonic stem cells, H3K4me3 is co-localized with H3K27me3 at the promoter of some genes, called bivalent genes. The silenced gene expression at the bivalent gene is reversed by the removal of H3K27me3, and then H3K4me3 becomes pronounced for these genes as cells differentiate. To switch off gene expression, methylation of H3K4 is removed by a demethylase such as a lysine-specific demethylase (LSD) 1 leading to over representation of H3K27me3 at the bivalent gene during differentiation or development (Seenundun, Rampalli, Liu, Aziz, Palii, Hong, Blais, Brand, Ge and Dilworth, 2010). This phenomenon is also observed in non-developmental inducible genes such as a heat-shock gene, Hsp70, in Drosophila (Gilmour and Lis, 1986). RNA polymerase II (RNA pol II) are accompanying with H3K4me3 at the promoter. Once the repressive mark is removed, paused RNA pol II is released from the promoter to produce full-length transcripts.

During ongoing transcription, H3K4me3 is mainly observed at the 5’ end of genes, representing initiation of transcription (Barski, Cuddapah, Cui, Roh, Schones, Wang, Wei, Chepelev and Zhao, 2007). The presence of H3K4me3 accompanying with H3S10p is preventing recruitment of an enzyme that methylases a lysine 9 of histone H3 as a repressive mark (Dormann, Tseng, Allis, Funabiki and Fischle, 2006). During elongating transcript, the lysine 36 of histone H3 are methylated by a Set domain containing methyltransferase that is recruited to an active form of RNA pol II phosphorylated at
Figure 5. Distribution of histone modifications across genes to define chromatin structures.

This pattern is generated by genome scale screening by several groups of histone modifications in yeast and mammalian cells by ChIP-seq. The middle represents euchromatin that is enriched in H3K4me2/3 and H3K9ac at TSS of genes during transcription initiation and H3K79me2 and H3K36me3 in the body of genes during transcription elongation. Both ends represent heterochromatin that is compacted more than euchromatin as shown here. As I mentioned before, H3K9me3 and H4K20me3 are mainly enriched in heterochromatin.
Figure 5
distributed across genes and are associated with transcriptional status (Figure 5). This discrete enrichment of each histone modifications at a given gene suggests that coordination of histone modifications is essential for regulation of gene expression.

- **DNA methylation**

DNA methylation is an epigenetic event that is frequently observed in silenced gene expression. The C-5 position of cytosine is predominantly methylated in CG dinucleotide rich regions that are mainly located within regulatory regions such as gene promoters. These CG rich regions are generally unmethylated, but become methylated upon gene silencing. There are two possibilities to explain how DNA methylation is associated with gene silencing. First, during transcription, the accessibility of regulatory elements embedded into DNA sequences to either transcription machinery or transcription factors is inhibited by DNA methylation. Alternatively, there are proteins that specifically recognize methylated DNA such as the methyl-CpG-binding domain (MBD) containing protein, MeCP2. MBD proteins exist in mammals and bind to methylated CG residues in *in vitro* (Ng, Zhang, Hendrich, Johnson, Turner, Erdjument-Bromage, Tempst, Reinberg and Bird, 1999). Secondly, the methylated DNA may influence the configuration of chromatin structures. Indeed, high levels of DNA methylation affect the preference of histones to incorporate into certain DNA sequence in *in vitro* (Davey, Pennings and Allan, 1997).

Interaction between histone methyltransferase and DNA methyltransferase supports that DNA methylation is associated with configuration of chromatin structure (Cedar and Bergman, 2009). In embryonic stem cells, SUV39H1 and SUV39H2 responsible for
H3K9me3 recruit DNA methyltransferase 3A and 3B to add methyl groups on CpG residues in satellite repetitive DNA (Lehnertz, Ueda, Derijck, Braunschweig, Perez-Burgos, Kubicek, Chen, Li, Jenuwein and Peters, 2003). The methylated CpG is possibly recognized by MBD proteins, which recruit chromatin modifying/remodeling complexes leading to further heterochromatinization (Jones, Veenstra, Wade, Vermaak, Kass, Landsberger, Strouboulis and Wolffe, 1998). However, the relation between chromatin structures and DNA methylation in gene expression is not yet completely clear.
Suggested effect of the GAA repeats at the FXN gene leading to FXN deficiency in FRDA patients.

Among various factors contributing to heterochromatin formation, high density of repetitive DNA, such as found in transposons, is a main stimulator across the genome (Grewal and Jia, 2007). Hyperexpanded GAA repeats have been reported to adopt a heterochromatin-like structure that is characterized by increased levels of di- and trimethylated lysine 9 of histone H3 (H3K9me2/3) and decreased acetylation on H3 and H4 at the FXN gene (Al-Mahdawi, Pinto, Ismail, Varshney, Lymperi, Sandi, Trabzuni and Pook, 2008; Rai, Soragni, Jenssen, Burnett, Herman, Coppola, Geschwind, Gottesfeld and Pandolfo, 2008; Saveliev, Everett, Sharpe, Webster and Festenstein, 2003; Soragni, Herman, Dent, Gottesfeld, Wells and Napierala, 2008). Inhibition of histone deacetylases restores levels of FXN expression in FRDA primary lymphocytes and in a murine model (Herman, Jenssen, Burnett, Soragni, Perlman and Gottesfeld, 2006; Rai, Soragni, Jenssen, Burnett, Herman, Coppola, Geschwind, Gottesfeld and Pandolfo, 2008), and manipulating histone modifications, particularly acetylation, can partially enhance expression of the FXN gene. Therefore, these studies suggest that changes in chromatin structure associated with the hyperexpanded GAA repeats induce FXN repression.

However, it is an open question whether the heterochromatin-like structure induced by the hyperexpanded tract of GAA repeats impacts initiation and/or elongation of FXN transcription. Some studies show that the heterochromatin-like structure induced by the hyperexpanded GAA repeats impact promoter activity, which affect initiation of FXN transcription (De Biase, Chutake, Rindler and Bidichandani, 2009; Kumari, Biacsi and Usdin). Marks enriched in heterochromatin such as H3K27me3, H3K9me3 and
heterochromatin protein (HP1) are detected at the transcription start site (TSS) of the $FXN$ gene in FRDA fibroblast lines, leading to the failure of $FXN$ transcription initiation. These repressive marks may also influence expression of $FXN$ antisense transcripts in the proximity of the TSS, thereby interfering with initiation of $FXN$ sense transcripts in FRDA patients (De Biase, Chutake, Rindler and Bidichandani, 2009). Other studies suggest that both defective initiation and transcript elongation is affected in FRDA (Kumari, Biacsi and Usdin, 2011b). Levels of both H3K4me3, indicator of active transcription initiation at the TSS of the $FXN$ gene and H3K36me3, an indicator of transcription elongation, are decreased at the $FXN$ gene in FRDA cell lines.

Besides altered histone modifications in the proximity of the hyperexpanded GAA repeats at the $FXN$ gene in FRDA, high levels of DNA methylation have been reported in the region upstream of the GAA repeats in FRDA cell lines (Castaldo, Pinelli, Monticelli, Acquaviva, Giacchetti, Filla, Sacchetti, Keller, Avvedimento, Chiariotti and Cocozza, 2008; Greene, Mahishi, Entezam, Kumari and Usdin, 2007). Moreover, the levels of DNA methylation correlate with the number of the GAA repeats (Castaldo, Pinelli, Monticelli, Acquaviva, Giacchetti, Filla, Sacchetti, Keller, Avvedimento, Chiariotti and Cocozza, 2008). Taken together, these observations suggest that epigenetic alterations induced by the hyperexpanded GAA repeats are associated with regulation of FXN expression in FRDA.
Current knowledge gap in molecular mechanisms to understand \textit{FXN} deficiency via the hyperexpanded GAA repeats in FRDA patients.

Epigenetic alterations caused by the hyperexpanded GAA repeats are one of the potential therapeutic targets for curing FRDA. Mouse models containing a longer GAA repeats tract exhibit sensitivity to epigenetic modifiers, suggesting that epigenetic regulation by the GAA repeats is likely a good target for potential drugs in restoring \textit{FXN} expression (Festenstein, Sharghi-Namini, Fox, Roderick, Tolaini, Norton, Saveliev, Kioussis and Singh, 1999; Saveliev, Everett, Sharpe, Webster and Festenstein, 2003). Moreover, a number of studies showed that specific HDACi are capable of \textit{FXN} expression and increasing histone acetylation in FRDA cells (Herman, Jenssen, Burnett, Soragni, Perlman and Gottesfeld, 2006; Rai, Soragni, Jenssen, Burnett, Herman, Coppola, Geschwind, Gottesfeld and Pandolfo, 2008; Soragni, Herman, Dent, Gottesfeld, Wells and Napierala, 2008). In contrast, levels of H3K9me3, a repressive mark, observed in the proximity of the long GAA repeats are maintained upon HDACi treatment (Herman, Jenssen, Burnett, Soragni, Perlman and Gottesfeld, 2006; Rai, Soragni, Jenssen, Burnett, Herman, Coppola, Geschwind, Gottesfeld and Pandolfo, 2008). Moreover, \textit{FXN} expression has no association with inhibition of H3K9 methylation using a small molecule in FRDA cells (Grabczyk and Usdin, 2000a). These results suggest that two or more epigenetic silencing pathways may be simultaneous involved in regulation of \textit{FXN} expression. Additionally, although DNA methylation is one of major factors associated with gene silencing, investigations of how DNA methylation might be involved in \textit{FXN} deficiency have been limited to date. Therefore, this thesis provides a detailed definition of the spectrum of chromatin structure changes associated with hyperexpanded GAA
repeats in order to further understand the molecular mechanisms underlying \textit{FXN} deficiency in FRDA cells.
The purpose of this study is to determine whether the hyperexpanded GAA repeats foster epigenetic changes associated with the formation of heterochromatin-like structures leading to inhibition of \( FXN \) transcription in FRDA.
Chapter 2. Materials and Methods

Some parts of this dissertation were modified from the following journal article:

**Cells**

Lymphoid cell lines were purchased from NIGMS Human Genetic Cell Repository at The Coriell Institute for Medical Research, Camden, NJ, USA. I selected three cell lines derived from FRDA patients: GM15850 harboring two expanded alleles of approximately 650 and 1030 GAA repeats; GM16798 harboring two expanded alleles of approximately 750 and 1000 GAA repeats and GM16209 harboring approximately 800 GAA repeats on both alleles of the FXN gene. As controls, I also used three lymphoid cell lines (GM15851, GM03928 and GM05152) derived from healthy, unaffected individuals containing short GAA repeat tracts within a normal range. Lymphoid cell lines were grown in RPMI 1640 medium supplemented with 15% FBS, 100U/ml penicillin and 100 μg/ml streptomycin at 37°C in 5% CO2. The GFP_GAA systems [GFP_(GAA•TTC)_0 and GFP_(GAA•TTC)_560] were generated by integration of a tract of GAA repeats derived from the GM16210 affected cells as described previously (Soragni, Herman, Dent, Gottesfeld, Wells and Napierala, 2008). A DNA fragment harboring approximately 560 GAA repeats was PCR amplified from genomic DNA isolated from the GM16210 cell line (NIGMS Human Genetic Cell Repository at The Coriell Institute for Medical Research, Camden, NJ, USA) as described in (Soragni, Herman, Dent, Gottesfeld, Wells and Napierala, 2008). PCR product was cleaved by Bsu36I and BssHII endonucleases and the GAA repeat containing fragment was cloned into intron 1 (1.2Kbp from the exon1/intron1 junction) of the *GFP* gene of the pGFP_Int plasmid (Seluanov, Mittelman, Pereira-Smith, Wilson and Gorbunova, 2004; Wilson and Roof, 1997). Plasmids containing 0 and 560 GAA repeats were integrated by site-specific recombination into the genome of the HEK293Flp-InT-Rex cell line (Invitrogen). The use of identical sites of
integration for both GFP_{(GAA\cdot TTC)₀} and GFP_{(GAA\cdot TTC)₅₆₀} constructs allows direct comparison between cell lines and eliminates any potential bias resulting from random integration events in different chromosomal contexts. Integrants were selected using hygromycin (200 µg/ml); individual hygromycin-resistant colonies were isolated and analyzed for repeat size and GFP expression level. Correct splicing of the GFP mRNA was determined as described earlier (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996). All constructs were sequenced prior to as well as after establishing the stable cell lines. The GFP_{GAA} lines were selectively maintained in Dulbecco’s Modified Eagle’s Medium with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin at 37°C in 5% CO₂ supplemented with hygromycin (200 µg/ml) and blasticidin (5 µg/ml). To induce transcription of the GFP minigene in the GFP_{GAA} system, cells were treated with 0.1 µg/ml tetracycline for 24h.

**Polymerase chain reaction (PCR)**

The size of the hyperexpanded tract of GAA repeats in lymphoid cell lines were determined by conventional PCR using primers 2500F and 629 under previously described conditions (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996). The insertion of the GAA repeats at an intron of the GFP gene in the GFP_{GAA} system was detected by the conventional PCR using the primers
listed in Table 1. The PCR products were analyzed on 0.8% agarose gels. To perform quantitative reverse transcriptase PCR (qRT-PCR), total RNAs from lymphoid cell lines and the GFP_GAA system were isolated using the RNeasy Mini Kit (Qiagen). DNase I (TURBO DNA-free; Ambion) was added to remove genomic DNA contamination from isolated RNAs. The qRT-PCR was conducted using the Power SYBR Green RNA-to-C<sub>T</sub> 1-Step Kit (7500 Fast Real Time-PCR System, Applied Biosystems). As a control, reactions were also performed without reverse transcriptase to confirm a removal of genomic DNA. To calculate levels of FXN expression, the delta Ct value was generated by subtraction of the Ct value of GAPDH from the Ct value of either FXN mRNA or 4 different region transcripts of each cell line. Levels of FXN mRNA from each cell line were normalized to the delta Ct of GM15851. For levels of FXN pre-mRNA expression, each delta Ct value of 4 different regions at the FXN gene was normalized to the delta Ct value of the region upstream of the GAA repeat from GM15851. All primers used in this study are listed in Tables 1 and 2.

**Chromatin immunoprecipitation (ChIP)**

ChIP was performed according to the EZ ChIP instructions (Upstate Biotechnology Inc.). Briefly, proteins and DNA were cross-linked with 1% formaldehyde for 10min (15min in the case of RNA pol II ChIP) at room temperature. The cross-linking reaction was quenched with 125mM glycine for 5min (10min in the case of RNA pol II ChIP). Whole lysates were prepared by using a cell lysis buffer (50mM Tris-HCl at pH8.0, 10mM EDTA, and 1% SDS) supplemented with protease inhibitor (Mini EDTA-free protease inhibitor cocktail, Roche) and sonicated to obtain 100 to 300 bp DNA fragments using a
Bioruptor Sonicator (Diagenode). The fragmented chromatin was diluted 10 times with dilution buffer (16.7mM Tris-HCl at pH 8.0, 167mM NaCl, 1.2mM EDTA, 1.1% Triton X-100, and 0.01% SDS). The chromatin fragments were pre-cleared for 1h using Protein A agarose (Millipore) pre-blocked with BSA and salmon sperm DNA. The equivalent of 5 x 10^6 (1 x 10^7 cells in the case of RNA pol II ChIP) cells was immunoprecipitated overnight with 5µg of antibodies recognizing histone modifications (15µg of RNA pol II antibody in the case of RNA pol II ChIP). For total histone H3 and H4, the equivalent of 5 x 10^6 cells was immunoprecipitated with 3µg of antibodies recognizing histone H3 and H4 after estimating saturation point between antibody for antibody and antigen. The immunoprecipitates were immobilized on the pre-blocked protein A agarose beads for 1h. Subsequently, the beads were washed with buffers containing low salt, high salt, and LiCl, chromatin was eluted from the beads with elution buffer (100mM NaHCO_3 and 1% SDS) at room temperature for 15min twice and subsequently subjected to the reverse cross-linking reaction using 5M NaCl at 65°C for at least 5h. The DNAs from chromatin complexes were isolated with Tris-EDTA buffer (80mM Tris-HCl at pH 6.5, and 20mM EDTA) supplemented with Proteinase K and RNase A at 37°C for 30min followed by 42°C for 1h. DNA fragments were purified using phenol/chloroform extraction before quantitative real-time PCR (qPCR). The qPCR was conducted using the Power SYBR Green-C_T Kit (7500 Fast Real Time-PCR System, Applied Biosystems). The qPCR was carried out as follows: 10 min at 94°C, 50 cycles of 30s at 94°C followed by 60s at 60°C. Since changes in nucleosome occupancy is associated with gene activation and gene silencing (Bryant, Prabhu, Floer, Wang, Spagna, Schreiber and Ptashne, 2008), it is worthy to notice that the changes of histone modifications (referred to as the “relative
fold over total H3”) determined by qPCR was analyzed by normalizing the quantity of the immunoprecipitated sample to the quantity of total histone H3 (total H4 in the case of H4K20me3) after normalization to inputs.

**Antibodies**

The antibodies used in this study were anti-rabbit IgG as a negative control (Cell Signaling), anti-total H3 (Cell Signaling or Active Motif) and anti-total H4 (Abcam), anti-H3K9/14ac (Upstate), anti-H3K4me2 (Active Motif), anti-H3K4me3 (Abcam or Active Motif), anti-H3K9me3 (Upstate or Active Motif), anti-H3K27me3 (Abcam), anti-H3K36me3 (Upstate), anti-H3K79me2 (Upstate or Active Motif), anti-H4K20me3 (Active Motif), and an antibody against the large subunit of RNA polymerase II (N20; Santa Cruz Biotechnology).

**D-ribofuranosylbenzimidazole (DRB) treatments**

To optimize condition of DRB treatments in the lymphoid cell lines, various concentration, 1, 50, and 100 µM of DRB was applied to 1 x 10^6 cells for 5 hr. H3K4me3 distribution was determined using ChIP assay upon inhibition of transcription elongation by DRB in the lymphoid cell lines as described in the ChIP section of Materials and Methods.

**Statistics**
A 2-way analysis of variance followed by the Bonferroni post-hoc test were performed to
determine the statistical significance of the results of qRT-PCR and ChIP analysis
between unaffected and FRDA cell lines.
Table 1. Primers used for analyses in the GFP_GAA reporter system.

<table>
<thead>
<tr>
<th>Primer (5' to 3')</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Repeat</strong></td>
<td>CTTCCCTTTACACAACGTTTGGGT</td>
<td>GTACTGTGTTTGGATTCAGTGAGGGACT</td>
</tr>
<tr>
<td><strong>Ex1/Ex2</strong></td>
<td>GCGACGTAACGGCCACAAGTT</td>
<td>ATGCCCTTCAGCTCGATGCGGT</td>
</tr>
<tr>
<td><strong>Ex1/Int</strong></td>
<td>GACGACGGCAACTACAAGACC</td>
<td>CTAGGACAAAGGTGCCTAAGACC</td>
</tr>
<tr>
<td><strong>Up</strong></td>
<td>AATAGCCTCCTGACCACAGATCCTT</td>
<td>CCATGTGACATCTAGCCCCGCA</td>
</tr>
<tr>
<td><strong>Down</strong></td>
<td>CCCACAGGCCTGAAACACT</td>
<td>TTCATCGTGCTAGGGGTAAA</td>
</tr>
<tr>
<td><strong>Int/Ex2</strong></td>
<td>CCCTAGCAGCGATGAACC</td>
<td>ATGCCCTTTCAGCTCGATGCGGT*</td>
</tr>
</tbody>
</table>

* This primer is the same as the GFP Ex1/Ex2 reverse primer.
Table 2. Primers used for analyses in the lymphoid cell lines.

<table>
<thead>
<tr>
<th>Primer (5’ to 3’)</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex3/Ex4</td>
<td>CCTTGCAGACAAGCCATACA</td>
<td>GGTCCACTGGATGGAGAAGA</td>
</tr>
<tr>
<td>In1Ex2</td>
<td>AGCACTCGGTACAGGCACT</td>
<td>GCCCAAAGTTCCAGATTTCC</td>
</tr>
<tr>
<td>−242</td>
<td>CGCATTTTATAAAACAAGGCACA</td>
<td>GTATGTTGGGCAGGAGAC</td>
</tr>
<tr>
<td>Pro (or -133)</td>
<td>CCCCCCATACCCAACCTACTG</td>
<td>GCCCGCCGCTTCTAAAATTC</td>
</tr>
<tr>
<td>+48</td>
<td>AAGCAGGCTCTCCATTATTG</td>
<td>CCGCAGGCACTCTTTCTGT</td>
</tr>
<tr>
<td>Up (or +1231)</td>
<td>GAAACCCAAGAATGGCTGTG</td>
<td>TTCCCTCCTCGTGAAACACC</td>
</tr>
<tr>
<td>+1394</td>
<td>GGTACGCCCCATGTTAGG</td>
<td>GCAACCAATCCCCAAGTTTC</td>
</tr>
<tr>
<td>Down</td>
<td>CTGGAAAAATTAGGCAAGTGG</td>
<td>CAGGGGTGGAAGCCCAATAC</td>
</tr>
<tr>
<td>In2Ex3</td>
<td>GGTAATCATGTTTTGGTTTTGTGC</td>
<td>AGTCCTCAAACGTGTATGGCTTGTC</td>
</tr>
</tbody>
</table>

Number on the name of primers denotes the first base of forward primer on the target region with respect to the TSS, +1.
Chapter 3. Defining of chromatin structure in the vicinity of the hyperexpanded GAA repeats at the \textit{FXN} gene

Some parts of this dissertation were modified from the following journal article:
GAA repeat expansion and FXN deficiency in the lymphoid cell lines

In order to discriminate lengths of the GAA repeats at the FXN gene in the lymphoid cell lines from unaffected (GM15851, GM03928 and GM05152) and FRDA patients (GM15850, GM16798 and GM16209), conventional PCR was carried out using a primer set that recognizes sequences surrounding where the GAA repeats at the FXN gene as previously reported (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996). A band around 500 bp is detected in unaffected controls, whereas FRDA lymphoid cell lines exhibit expanded GAA repeats up to 3Kbp (Figure 6). This pattern is consistent with number of the GAA repeats reported in NIGMS Human Genetic Cell Repository at The Coriell Institute for Medical Research as described in Materials and Methods.

To confirm insufficient levels of FXN transcription in FRDA cell lines, FXN mature RNA was measured by qRT–PCR using a primer set recognizing exon 3 as a forward primer and exon 4 as a reverse primer (Figure 7). Clinically affected FRDA patients express as little as 5% and as much as 35% of FXN relative to unaffected individuals (Campuzano, Montermini, Molto, Pianese, Cossee, Cavalcanti, Monros, Rodius, Duclos, Monticelli, Zara, Canizares, Koutnikova, Bidichandani, Gellera, Brice, Trouillas, De Michele, Filla, De Frutos, Palau, Patel, Di Donato, Mandel, Cocozza, Koenig and Pandolfo, 1996; Pianese, Turano, Lo Casale, De Biase, Giacchetti, Monticelli, Criscuolo, Filla and
Cocozza, 2004). In order to establish a common epigenetic profile in FRDA cell lines, I selected three cell lines expressing low (GM16798), medium (GM15850), and high (GM16209) levels of the $FXN$ mRNA corresponding to 8, 17 and 29% of the $FXN$ mRNA found in the GM15851, respectively.
Figure 6. Hyperexpanded GAA repeats at the FXN gene in the lymphoid cell lines.

Hyperexpanded tract of GAA repeats (indicated by the asterisk) was confirmed by conventional PCR in lymphoid cell lines derived from patients with FRDA but not in cell lines derived from unaffected individuals.
Figure 6
Figure 7. *FXN* transcription is impaired in FRDA lymphoid cell lines.

*FXN* mRNA was analyzed in 3 control and 3 FRDA cell lines by qRT-PCR with primers complementary to the exons 3 and 4. White bars designate control cell lines, while gray bars designate FRDA cell lines. Each cell line has different textures. *FXN* expression was normalized to the expression level determined for control GM15851 cells. The experiment was conducted in triplicate. All data are expressed as mean ± SD. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 7

![Graph showing FXN Mature RNA expression in Unaffected and FRDA samples.](image-url)
Distribution of histone modification at the FXN gene in the lymphoid cell lines

To determine if the hyperexpanded GAA repeats might influence chromatin structures, chromatin immunoprecipitation (ChIP) was used to examine histone modifications in the lymphoid cell lines. Chromatin from the lymphoid cell lines was fragmented to obtain a majority of chromatin fragment under 500 bp (Figure 8). All six lymphoid cell lines, including three unaffected cells and three FRDA, were applied to examine each histone modification in the following text and figures. Additionally, to demonstrate that the hyperexpanded GAA repeats is sufficient to induce locally repressive chromatin structure, I took advantage of the GFP_GAA reporter system that is engineered to harbor 560 GAA repeats at the GFP gene introduced into HEK293 cells. Each value in the ChIP experiments discussed below is represented as the relative fold over total H3 of DNA recovery against each histone modifications when normalized to total H3. Occupancy of total H3 is not affected by the presence of the hyperexpanded GAA repeats in either the lymphoid cell lines or the GFP_GAA system (Figure 9).

H3K9/14ac

To confirm the decreased levels of acetylation of histones at the FXN gene in FRDA observed in other reports, I performed ChIP for H3K9/14ac in the lymphoid cell lines. H3K9/14ac is well-known to be acetylated by GCN5 and p300/CBP-associated factor (PCAF), chromatin remodelers, in eukaryotes (Bhaumik, 2011). Four primer sets were used to define the distribution of H3K9/14ac at the FXN gene. I observed that the levels of H3K9/14ac, when normalized for nucleosomal occupancy, were decreased at the FXN
Figure 8. Fragmented chromatin extracted from the lymphoid cell lines and the GFP_GAA reporter system.

Fragmented chromatin was detected using a 0.8% agarose gel. The majority of resulting fragmentation is between 100 and 500 bp in length.
Figure 8
Figure 9. Hyperexpanded GAA repeats have no association with distribution of total histone H3.

Total histone H3 in lymphoid cell lines and the GFP_GAA reporter system was determined by ChIP. DNA from chromatin immunoprecipitated using antibodies specific for total histone H3 was subjected to qPCR using primers amplifying (A) promoter region, fragments upstream (Up) and downstream (Down) of the GAA repeats and junction between intron 2 and exon 3 (In2Ex3) in the lymphoid cell liens, and (B) Ex1/Int, a junction between exon 1 and the intron; Up, region upstream of the GAA repeats; Down, region downstream of the GAA repeats; Int1/Ex2, a junction between the intron and exon2 for the GFP_GAA reporter system. Total histone H3 levels are normalized to 1% of the input (5x10^6 cells). White bars designate control cell lines, while gray bars designate FRDA cell lines. Each cell line has different textures. No major differences of total histone H3 distribution are observed at the entire FXN gene in the lymphoid cell lines.
Figure 9

A. 

(GAA)$_N$

E1  Up  Down  In2Ex3

Promoter  (-133)  Up (+1231)  Down

% Input

2.5

GM15851
GM03928
GM05152
GM15950
GM16798
GM16209

0.0  0.5  1.0  1.5  2.0  2.5

Promoter  Up  Down  In2Ex3

B. 

(GAA)$_0 / 560$

Exon 1  Up  Down  Int/Ex2

% Input

1.0

GM15851
GM03928
GM05152
GM15950
GM16798
GM16209

0.0  0.2  0.4  0.6  0.8  1.0

Exon 1/Int  Up  Down  Int/Ex2

GFP$_{(GAA\+TTC)\_0}$

GFP$_{(GAA\+TTC)\_560}$
gene in FRDA cell lines as expected. Specifically the levels of H3K9/14ac were the lowest at the region upstream of the GAA repeats in FRDA cells when compared with unaffected cells (Figure 10). On the contrary, in promoter region of the FXN gene, the levels of H3K9/14ac were similar in both groups. This suggests that the hyperexpanded GAA repeats are associated with repressive chromatin structure, but this structure does not extend to the promoter in FRDA cell lines.

**H3K9me3**

To add more evidence that the FXN gene in patients with FRDA has a repressive chromatin structure, I examined levels of H3K9me3 at the FXN gene by ChIP (Figure 11). High levels of H3K9me3 are observed within heterochromatin such as is found at centromeres (Wang, Schones and Zhao, 2009). This methylation is most likely associated with silencing of gene expression. At the FXN gene, H3K9me3, when normalized for nucleosomal occupancy, was greatly enriched in the region upstream of the GAA repeats where lower levels of H3K9/14ac were observed in FRDA cell lines.

**H4K20me3**

Although H3K9me3 is preferentially enriched at heterochromatin, this histone modification is also detected in the coding regions of transcriptionally active genes (Vakoc, Mandat, Olenchock and Blobel, 2005). In centromeric region, enrichment of H4K20me3 is accompanied by enrichment of H3K9me3.
Figure 10. H3K9/14ac is significantly repressed at the FXN gene in FRDA cells.

Distribution of H3K9/14ac at the FXN gene was defined by ChIP assay. DNA from chromatin immunoprecipitated using antibodies specific for H3K9/14ac was subjected to qPCR using primers amplifying promoter region, fragments upstream (Up) and downstream (Down) of the GAA repeats and junction between intron 2 and exon 3 (In2Ex3). The experiment was conducted in triplicate using three FRDA and three control cell lines. Data are expressed as mean ± SEM. P values were calculated by comparing the averages of the relative fold over total H3 between unaffected and FRDA lines. Levels of H3K9/14ac was significantly decreased in at the region upstream of the GAA repeats in FRDA cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 10

H3K9/14ac

$p < 0.001$
Figure 11. H3K9me3 is significantly increased at the *FXN* gene in FRDA cells.

Distribution of H3K9me3 at the *FXN* gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA and three control cell lines. Data are expressed as mean ± SEM. The region upstream of the GAA repeats in FRDA cells is enriched with levels of H3K9me3. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 11

H3K9me3

Ratio (H3K9me3/H3)

$\rho < 0.001$

GM15851
GM03928
GM05152
GM15850
GM16798
GM16209

Promoter  Up  Down  In2Ex3
Figure 12. H4K20me3 is significantly increased at the FXN gene in FRDA cells.

Distribution of H4K20me3 at the FXN gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA and three control cell lines. Data are expressed as mean ± SEM. The region upstream of the GAA repeats in FRDA cells is enriched with levels of H4K20me3. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 12

H4K20me3

$\text{Ratio (H4K20me3/H4)}$

$p < 0.001$

GM15851
GM03928
GM05152
GM15850
GM16798
GM16209

Promoter  Up  Down  In2Ex3
(Kournouli, Jeppesen, Mahadevhaiah, Burgoyne, Wu, Gilbert, Bongiorni, Prantera, Fanti, Pimpinelli, Shi, Fundele and Singh, 2004; Schotta, Lachner, Sarma, Ebert, Sengupta, Reuter, Reinberg and Jenuwein, 2004). To ascertain whether high levels of H3K9me3 at the \(FXN\) gene in FRDA cells are associated with the heterochromatin-like structure induced by proximate regions of the hyperexpanded GAA repeats, I determined the status of H4K20me3 at the \(FXN\) gene in FRDA and control cell lines (Figure 12). The levels of H4K20me3, when normalized for nucleosomal occupancy, were the highest at the region upstream of the GAA repeats in FRDA cell lines. The results of ChIP experiments indicate that enrichment of H3K9me3 was associated with those of H4K20me3 at the region upstream of the hyperexpanded GAA repeats in FRDA cells. These data demonstrate that the region upstream of the hyperexpanded GAA repeats facilitate a heterochromatin-like structure in FRDA cell lines.

**H3K27me3**

Repressive chromatin structures, such as the inactive X chromosome and silenced inducible genes such as the bivalent gene, are frequently associated with enrichment of H3K27me3 (Plath, Fang, Mlynarczyk-Evans, Cao, Worringer, Wang, de la Cruz, Otte, Panning and Zhang, 2003; Seenundun, Rampalli, Liu, Aziz, Palii, Hong, Blais, Brand, Ge and Dilworth, 2010). To further define change of chromatin configuration in the presence of the hyperexpanded GAA repeats, I examined distribution of H3K27me3 at the \(FXN\) gene in the lymphoid cell lines (Figure 13). ChIP assay revealed over representation of H3K27me3, when normalized for nucleosomal occupancy, in the region upstream of the
GAA repeats in FRDA cell lines, co-localizing with both H3K9me3 and H4K20me3 enrichment.

Taken together, I found that the expansion of GAA repeats induces a particular constellation of histone modifications associated with heterochromatin-like structures at the $FXN$ gene in FRDA cells.
Figure 13. H3K27me3 is enriched at the FXN gene in FRDA cells.

Distribution of H3K27me3 at the FXN gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA and three control cell lines. Data are expressed as mean ± SEM. The region upstream of the GAA repeats in FRDA cells is enriched with levels of H3K27me3. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 13

H3K27me3

Ratio (H3K27me3/H3)

$p < 0.05$

GM15851
GM03928
GM05152
GM15850
GM16798
GM16209
Distribution of histone modification in the GFP_GAA reporter system

To investigate whether the alteration of histone modifications at the FXN gene results from either inherent characteristics of the FXN gene or the hyperexpanded GAA repeats per se, I used a GFP_GAA reporter system that harbors a single copy of the GFP reporter gene containing 560 or 0 GAA repeats in the intron (Figure 14). I determined an effect of integrated tract of the GAA repeats on transcription using qRT-PCR. Mature RNA of GFP is decreased in the GFP_{(GAA•TTC)_{560}} cells compared to the GFP_{(GAA•TTC)_{0}} cells under the tetracycline treatment (Figure 15), indicating that the GFP_GAA system genetically resembles FRDA.

To assess changes in chromatin structures caused by presence of longer GAA repeats in the context of the GFP gene, I analyzed the distribution of histone modifications in the vicinity of the GAA repeats in the GFP_GAA reporter system.

H3K9/14ac and H3K9me3

In order to validate that chromatin structure flanking the hyperexpanded GAA repeats has characteristics of a heterochromatin-like structure, I examined distribution of H3K9/14ac as a mark for open chromatin (Figure 16) and H3K9me3 as a mark for closed chromatin in the GFP_GAA system (Figure 17). To further define whether the tract containing 560 GAA repeats has ability to bidirectionally spread across the GFP gene as characteristics of heterochromatin, 2 more primer sets, a junction between exon1 and the intron and a junction between the intron and exon 2 were used to perform qPCR following
Figure 14. Validation of the presence of 560 GAA repeats in the GFP_GAA system.

The presence of 560 GAA repeats in the GFP_GAA reporter system repeats was confirmed by conventional PCR. The region flanking 560 GAA repeats was amplified by primer sets, Up and Down for the GFP gene listed in Table 1. The GFP_(GAA•TCC) 560 cells exhibit approximately 3Kb of PCR product, indicating that appropriate interstion and maintenance of 560 GAA repeats at the GFP gene in the GFP_(GAA•TCC) 560 cells.
Figure 14
Figure 15. *GFP* transcription is impaired in the GFP_(GAA•TCC) 560 cells.

*GFP* mRNA was analyzed in the GFP_GAA reporter system with primers complementary to the exons 1 and exon 2. White bars designate control cell lines (the GFP_(GAA•TCC) 0 cells), while gray bars designate the GFP_(GAA•TCC) 560 cells. Each cell line has different textures. The experiment was conducted in triplicate. All data are expressed as mean ± SD. Asterisk indicates statistical significance (*p* < 0.05). The longer GAA repeats is associated with *GFP* deficiency in the GFP_(GAA•TCC) 560 cells.
Figure 15
immunoprecipitation of chromatin against each histone modification. I confirmed
decreased levels of H3K9/14ac and increased levels of H3K9me3, when normalized for
nucleosomal occupancy, in the vicinity of the GAA repeats at the GFP gene in the
GFP\_\((\text{GAA}\_\text{TTC})_{560}\) cells as shown in previous study (Soragni, Herman, Dent,
Gottesfeld, Wells and Napierala, 2008). Furthermore, I observed that both H3K9/14ac
and H3K9me3, when normalized for nucleosomal occupancy, extended up to the junction
between exon 1 and the intron as well as the junction between the intron and exon 2 of
the GFP gene in the GFP\_\((\text{GAA}\_\text{TTC})_{560}\) cells. These results demonstrate that the
introduced GAA repeat induces condensed chromatin structures in the proximity of the
GAA repeats. Moreover, these results suggest that the hyperexpanded tract of GAA
repeats \textit{per se} has the ability to induce spreading of condensed chromatin structures.
Figure 16. H3K9/14ac is significantly decreased at the *GFP* gene in the presence of GAA repeats.

Upon tetracycline treatment, the levels of H3K9/14ac in the GFP_GAA reporter system were determined by ChIP assay. White bars represent the GFP_(GAA•TTC)_{0} cells. Gray bars represent the GFP_(GAA•TTC)_{560} cells. Data are expressed as mean ± SEM.

Asterisk indicates statistical significance (one: $p < 0.05$, two: $p < 0.01$, three: $p < 0.001$). The abundance of histone modifications is shown relative to input DNA in the PCR and normalized to a total H3 for each region. Levels of H3K9/14 are less enriched at the entire GFP gene in the GFP_(GAA•TTC)_{560} cells relative to the GFP_(GAA•TTC)_{0} cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 16
Figure 17. H3K9me3 is enriched at the GFP gene in the presence of longer GAA repeats.

The levels of H3K9me3 in the GFP_GAA reporter system were determined in active transcription. White bars represent the GFP_(GAA•TTC)_0 cells. Gray bars represent the GFP_(GAA•TTC)_560 cells. Data are expressed as mean ± SEM. Asterisk indicates statistical significance (one: \( p < 0.05 \), two: \( p < 0.01 \), three: \( p < 0.001 \)). The abundance of histone modifications is shown relative to input DNA in the PCR and normalized to a total H3 for each region. H3K9me3 are significantly increased in the GFP_(GAA•TTC)_560 cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 17

![Graph showing H3K9me3 expression levels at different genomic locations for two conditions: GFP_(GAA*TTC) 0 and GFP_(GAA*TTC) 560. The x-axis represents genomic regions: Ex1/Int, Up, Down, Int/Ex2. The y-axis represents H3K9me3/H3 expression levels.](image)
Chapter 4. Determination of the effect of the GAA repeats in transcription in the context of chromatin structure

Some parts of this dissertation were modified from the following journal article:
In lymphoid cell lines

FXN transcription in lymphoid cells

In order to assess the effect of the hyperexpanded GAA repeats on the transcription progression at the FXN gene, I examined levels of FXN pre-mRNA in one upstream (from +1231 to +1344 within the intron 1) and in three regions downstream (variable positions, depending on repeat tract length) of the GAA repeats (Figure 18). Importantly, all primers pairs used in this experiment recognize FXN pre-mRNA to avoid a bias from the FXN mRNA. I found that levels of FXN pre-mRNA at the region upstream of the hyperexpanded GAA repeats have no differences in all three unaffected control cell lines and two of the three FRDA lymphoid cell lines. The GM16798 cell line exhibited a lower level of the FXN pre-mRNA upstream of the GAA repeats as compared to the remaining five cell lines. The GM16798 also exhibited the lowest expression of FXN mRNA (Figure 7). Additionally, FRDA patient from whom the GM16798 cell line was derived reported the earliest onset of the disease (5 yr vs. 10 and 11 yrs for the other FRDA individuals). In contrast to the similar levels of FXN pre-mRNA upstream of the GAA repeats, levels of FXN pre-mRNA downstream of the GAA repeats were reduced approximately 4-fold in all three FRDA lymphoid cell lines compared to unaffected cell lines. This result suggests that levels of the FXN pre-mRNA downstream expression are tightly associated with the pathology of FRDA.
Figure 18. FXN deficiency appears at the region downstream of the hyperexpanded GAA repeats in FRDA cells.

Relative level of FXN pre-mRNA was determined in various regions of the transcript. One ‘Up’ and three ‘Down’ primers amplify regions in the immediate vicinity of the GAA repeats. In1Ex2 and In2Ex3 anneal to the junctions between the corresponding introns and exons of FXN pre-mRNA. White bars designate control cell lines, while gray bars designate FRDA cell lines. Each cell line has different textures. FXN pre-mRNA expression was normalized to the level of the fragment amplified upstream of the GAA region in the control GM15851 cells. The experiment was performed in triplicate. The region downstream transcript is significantly decreased in FRDA cells compared to unaffected cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 18
Distribution of total RNA pol II at the FXN gene

Gene expression is dependent on the accessibility of the DNA sequences to RNA polymerase, transcription factors, and chromatin remodeling complexes. In an effort to further investigate how the hyperexpanded GAA repeats influence progression of FXN transcription, I analyzed the distribution of total RNA pol II at the FXN gene by ChIP assay (Figure 19). Immunoprecipitation of total RNA pol II include both unphosphorylated and phosphorylated forms of the large subunit of RNA pol II. Seven primer sets were used to map the distribution of RNA Pol II in the FXN gene. Based on prior studies showing transcriptional pausing in the proximity of the long GAA tracts in

in vitro and in cell culture system (Grabczyk and Usdin, 2000b; Krasilnikova, Kireeva, Petrovic, Knijnikova, Kashlev and Mirkin, 2007), I expected that RNA pol II progression would be inhibited in the region upstream of the GAA repeats, resulting in accumulation of transcriptional machinery in the region upstream of the GAA repeats in FRDA cell lines. However, I detected repressed levels of RNA pol II across the entire FXN gene in FRDA cell lines relative to unaffected cell lines. Importantly, levels of RNA pol II were not significantly changed at promoter region (from -242 to -12) in FRDA cell lines compared to unaffected cell lines, suggesting that FXN promoter activity is not affected in FRDA cell lines. These results correlate with similar levels of FXN pre-mRNA upstream of the GAA repeats observed in both FRDA and unaffected control cell lines as shown in Figure 17.

Interestingly, RNA pol II is mostly enriched at a FXN promoter-proximal region (from +48 to +150) in unaffected cell lines, even though a number of the GAA repeats is short
Figure 19. Distribution of total RNA pol II is repressed at the FXN gene in FRDA cells.

Level of total RNA pol II was determined by immunoprecipitating the large subunit of RNA pol II in FRDA and unaffected cells. Average occupancy of RNA pol II across the FXN gene in all three FRDA and three control cells is shown using gray circles and black squares, respectively. The position of the first nucleotide for the forward primer relative to transcription start site (TSS) is indicated below the X axis. The specificity of RNA pol II antibody for ChIP was verified by RNA pol II pausing at the +55 bp of the GAPDH gene. Error bars represent standard error of the mean. The experiment was conducted in triplicate using three FRDA and three control cell lines. The $P$ value was generated by comparing the average % Input of total RNA pol II between unaffected and FRDA lymphoid cell lines. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 19
in these cells. The accumulation at (+) 48 bp is significant as compared to either (-)242 bp or (-)133 bp ($p < 0.05$). This phenomenon is also observed in some genes as shown at the GAPDH gene (Guenther, Levine, Boyer, Jaenisch and Young, 2007), where RNA pol II is known to pause. No RNA pol II Pausing was detected in the FRDA cell lines. These results suggest that defective transcriptional in FRDA cells originate at a post-initiation step of $FXN$ transcription, which result in decreased enrichment of the transcription machinery at the promoter-proximal region as well as throughout the $FXN$ gene.

**H3K4me2 and H3K4me3**

To investigate changes of active chromatin structure caused by the hyperexpanded GAA repeat, histone modifications relative to active chromatin structures was examined by ChIP in the lymphoid cell lines. As indicators of active transcription initiation, levels of both H3K4me2 and H3K4me3 exhibit a peak at transcription start site (TSS) of a given gene that is recently transcribed. Although H3K4me2 and H3K4me3 levels varied among lymphoid cell lines, levels of H3K4me2 and H3K4me3, when normalized for nucleosomal occupancy, were not changed at the promoter region (from −138 to −12) of the $FXN$ gene in FRDA cell lines relative to unaffected controls (Figure 20 and Figure 21). These results are correlated with our quantitative $FXN$ pre-mRNA analyses showing similar levels of $FXN$ pre-mRNA upstream of GAA repeats in both unaffected and FRDA cell lines. Interestingly, H3K4me2 and H3K4me3 were significantly diminished at a region upstream of the GAA repeats (from +1228 to +1344) of the $FXN$ gene in FRDA cell lines, while levels of H3K4me2 and H3K4me3 cells, when normalized for nucleosomal occupancy, were increased at this region in the unaffected cells. Enrichment
Figure 20. H3K4me2 is decreased at the region upstream of the GAA repeat, not promoter region in FRDA cells.

Distribution of H3K4me2 at the FXN gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA (gray bars) and three control cell lines (white bars). Each cell line has different texture. Data are expressed as mean ± SEM. Levels of H3K4me2 are decreased at the region upstream of the GAA repeat in FRDA cell, while unaffected show great levels of H3K4me2. No differences of H3K4me2 were detected in both FRDA and unaffected cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 20

H3K4me2

Ratio (H3K4me2/H3)

$\* p < 0.001$

- Promoter
- Up
- Down
- In2Ex3

- GM15851
- GM03928
- GM05152
- GM15850
- GM16798
- GM16209
Figure 21. Levels of H3K4me3 are not affected at promoter region in FRDA cells.
Distribution of H3K4me3 at the FXN gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA (gray bars) and three unaffected cell lines (white bars). Each cell line has different texture. Data are expressed as mean ± SEM. Levels of H3K4me3 are not changed at promoter region in FRDA cells compared to unaffected cells. However, H3K4me3 is decreased at the region upstream of the GAA repeats in FRDA cells compared to unaffected cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 21
of H3K4me2 has been associated with non-methylated CpG residues more tightly than transcriptional activity (Okitsu and Hsieh, 2007; Okitsu, Hsieh and Hsieh). In previous studies from other groups, increased levels of DNA methylation in the region upstream of the GAA repeats have been reported in FRDA cell lines relative to unaffected cells (Castaldo, Pinelli, Monticelli, Acquaviva, Giacchetti, Filla, Sacchetti, Keller, Avvedimento, Chiarotti and Cocozza, 2008; Greene, Mahishi, Entezam, Kumari and Usdin, 2007). Thus, the lower levels of H3K4me2 at the region upstream of the GAA repeats in FRDA cell lines could imply hyper DNA methylation, particularly at this region in FRDA patients. Enrichment of H3K4me3 in coding regions is associated with the efficient transition between initiation and elongation of transcription during ongoing transcription (Sims, Millhouse, Chen, Lewis, Erdjument-Bromage, Tempst, Manley and Reinberg, 2007). Therefore, the decreased levels of H3K4me3 in FRDA cell may reflect that the hyperexpanded GAA repeats influence the efficiency of post-initiation of FXN transcription.

To further examine whether levels of H3K4me3 at the region upstream of GAA repeats in unaffected cells is associated with the activity of transcription elongation rather than transcription initiation, I measured levels of H3K4me3 upon inhibition of transcription elongation by D-ribofuranosylbenzimidazole (DRB) (Figure 22). DRB, a pharmacologic inhibitor of CDK9, decreases RNA pol II phosphorylation at Serine-2, a hallmark feature of the elongating form of RNA pol II during transcription. Following DRB treatment, the control cell line (GM15851) exhibited a significant decrease of H3K4me3 at the region upstream of the GAA repeats, which was very similar to that observed in FRDA cell line
Figure 22. H3K4me3 at the region upstream of the GAA repeats in FRDA cells is associated with impaired FXN transcription elongation.

H3K4me3 distribution were determined by ChIP assay upon inhibition of transcription elongation (50 µM DRB for 5 hr). The ChIP experiment was conducted in twice using one FRDA cell (gray bars: GM15850) and one unaffected cell line (white bars: GM15851). Data are expressed as mean ± SEM. H3K4me3 is decreased at the region upstream of the GAA repeats in FRDA cells upon inhibition of transcription elongation.
Figure 22

![Graph showing H3K4me3/H3 levels for different conditions](image)
(GM15850), when normalized for nucleosomal occupancy. This result suggests that the levels of H3K4me3 at the region upstream of the GAA repeats are associated with transcription elongation rather than transcription initiation. Taken together, these results suggest that the hyperexpanded GAA repeat does not affect chromatin structures at the *FXN* promoter region but does result in impaired elongation of *FXN* transcription in FRDA.

**H3K36me3 and H3K79me2**

To further understand the effect of the hyperexpanded GAA repeats on the transcription process throughout the *FXN* gene, I examined the levels of H3K36me3 and H3K79me2, indicators of transcription elongation (Figure 23 and Figure 24). Increasing enrichment of H3K36me3 to the 3′-end of a gene is indicating ongoing transcription elongation. In unaffected cell lines, H3K36me3 levels, when normalized for nucleosomal occupancy, progressively increased toward the 3′-end of the *FXN* gene. However, enrichment of H3K36me3 was repressed across the *FXN* gene in FRDA cell lines. The distribution of H3K36me3 is consistent with total RNA pol II progression shown in Figure 19. These results indicate defective transcription elongation of the *FXN* gene in FRDA cell lines. A genome-wide study revealed that enrichment H3K79me2 at regions downstream of the TSS accompanying with H3K4me3 at the TSS is associated with full-length transcripts of genes (Kanhere, Viiri, Araujo, Rasaiyaah, Bouwman, Whyte, Pereira, Brookes, Walker, Bell, Pombo, Fisher, Young and Jenner, 2010). In contrast, enrichment of H3K4me3 alone is feature for genes that experience transcription initiation alone, without elongation (Rahl, Lin, Seila, Flynn, McCuine, Burge, Sharp and Young, 2010). I found that
H3K79me2, when normalized for nucleosomal occupancy, were dramatically decreased throughout the $FXN$ gene in FRDA cells, while H3K79me2 is pronounced in the vicinity of the GAA repeats in the unaffected control cells. As shown in Figure 20, since enrichment of H3K4me3 has no differences at the $FXN$ promoter region in FRDA cell lines compared to unaffected cells, these results further support the that $FXN$ deficiency in FRDA cells results from defective elongation step of $FXN$ transcription after successful initiation.
Figure 23. H3K36me3 are repressed at the entire $FXN$ gene in FRDA cells relative to unaffected cells.

Distribution of H3K36me3 at the $FXN$ gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA (gray bars) and three unaffected cell lines (white bars). Each cell line has different texture. Data are expressed as mean ± SEM. Levels of H3K36me3 are decreased at the entire $FXN$ gene in FRDA cells, while they are gradually increased toward 3’ end of the $FXN$ gene in unaffected cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 23

H3K36me3

Ratio (H3K36me3/H3)

Promoter  Up  Down  In2Ex3

$p < 0.001$

GM15851  GM03928  GM05152  GM15850  GM16798  GM16209
Figure 24. H3K79me2 is significantly decreased in FRDA cells.

Distribution of H3K79me2 at the FXN gene was defined by ChIP assay. The ChIP experiment was conducted in triplicate using three FRDA (gray bars) and three unaffected cell lines (white bars). Each cell line has different texture. Data are expressed as mean ± SEM. Levels of H3K79me2 are decreased at the entire FXN gene in FRDA cells, while they are enriched in the vicinity of the GAA repeats at the FXN gene in unaffected cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 24

![Graph showing the ratio of H3K79me2 to H3 across different genomic regions. The graph indicates significant differences marked with p < 0.001 for certain regions.]
In the GFP_GAA system
H3K4me3 and H3K79me2

To further investigate the effect of the presence of the GAA repeats in progression of transcription independently of FXN promoter or FXN intronic sequences, I examined enrichments of H3K4me3 (transcription initiation) and H3K79me2 (transcription elongation) at the GFP gene in the GFP_GAA reporter system using ChIP assay when normalized for nucleosomal occupancy (Figure 25 and Figure 26). Levels of H3K4me3 were not changed at the region upstream of the GAA repeats including a junction between exon 1 and the intron and the region upstream of the GAA repeats (named ‘Up’) in the GFP_(GAA•TTC)_{560} cells compared to the GFP_(GAA•TTC)_{0} cells that lack the GAA repeats. However, H3K4me3 levels were significantly decreased in the region downstream of the GAA repeats in the GFP_(GAA•TTC)_{560} cells. This distribution of H3K4me3 at the GFP gene in GFP_(GAA•TTC)_{560} cells is comparable to that found at the FXN gene in the FRDA cell lines. In addition, I found that enrichment of H3K79me2 was significantly decreased at the region downstream of the GAA repeats at the GFP gene in the GFP_(GAA•TTC)_{560} cells when compared to the corresponding region in the GFP_(GAA•TTC)_{0} cells. These results clearly show that the presence of the hyperexpanded GAA repeats not only stimulate epigenetic changes in the proximate chromatin independently of the DNA context, but also are sufficient to impede transcription elongation. More importantly, the observations of histone modification showing impediment of transcription elongation in the GFP_(GAA•TTC)_{560} cells, without affecting initiation, is consistent with alteration of histone modifications
Figure 25. H3K4me3 is affected at the region downstream of the GAA repeats in the presence of the GAA repeats.

The levels of H3K4me3 in the GFP_GAA reporter system were determined by ChIP assay. White bars represent the GFP_(GAA•TTC)_0 cells. Gray bars represent the GFP_(GAA•TTC)_560 cells. Data are expressed as mean ± SEM. Asterisk indicates statistical significance (p < 0.05). The abundance of histone modifications is shown relative to input DNA in the PCR and normalized to a total H3 for each region. Upon tetracycline treatment, H3K4me3 are affected at the region downstream of the GAA repeats in the GFP_(GAA•TTC)_560 cells compared to the GFP_(GAA•TTC)_0 cells.

Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 25

H3K4me3

![Bar graph showing H3K4me3 levels in different regions with error bars and legend for different conditions.](image-url)
Figure 26. H3K79me2 is affected at the region downstream of the GAA repeats in the GFP\_(GAA•TTC)_{560} cells.

The levels of H3K79me2 in the GFP\_GAA reporter system were determined by ChIP assay. White bars represent the GFP\_(GAA•TTC)_{0} cells. Gray bars represent the GFP\_(GAA•TTC)_{560} cells. Data are expressed as mean ± SEM. Asterisk indicates statistical significance (one: $p < 0.05$, three: $p < 0.001$). The abundance of histone modifications is shown relative to input DNA in the PCR and normalized to a total H3 for each region. Upon tetracycline treatment, H3K79me3 are decreased at the region upstream of the GAA repeats in the GFP\_(GAA•TTC)_{560} cells compared to the GFP\_(GAA•TTC)_{0} cells. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 26

H3K79me2

---

<table>
<thead>
<tr>
<th>Region</th>
<th>GFP(GAA•TTC) 0</th>
<th>GFP(GAA•TTC) 560</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex1/Int</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Int/Ex2</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

---

H3K79me2/H3
observed in FRDA cell lines. These results strongly suggest that elongation is the major step hampered by the hyperexpanded GAA repeats in FRDA patients.

Longer GAA repeats is associated with inhibiting progression of RNA polymerases in \textit{in vitro} and \textit{in vivo} systems (Grabczyk and Usdin, 2000b; Krasilnikova, Kireeva, Petrovic, Knijnikova, Kashlev and Mirkin, 2007). Moreover, transcription arrest is associated with alteration of histone modifications (Balakrishnan and Milavetz, 2007; Li, Carey and Workman, 2007; Nagata, Ito, Arimitsu, Koyama and Sekimizu, 2009). To assess whether RNA pol II arrest associated with the presence of the GAA repeats can prompt the cascade of silencing events, I used the GFP\textsubscript{GAA} reporter system allowing us to precisely control initiation of transcription using tetracycline regulated operator/repressor (Figure 27). Generally, the GFP\textsubscript{GAA} reporter system has been cultured in the absence of tetracycline to maintain cell lines. Both the GFP\textsubscript{(GAA\textsubscript{TTC})_{560}} and the GFP\textsubscript{(GAA\textsubscript{TTC})_{0}} cells were sub-cultured for less than 10 passages prior to the tetracycline treatment. In order to investigate the effect of transcription in alteration of histone modifications by the hyperexpanded GAA repeats, the cells were prepared in parallel in the absence or presence of the tetracycline. Upon the absence of tetracycline to inhibit transcription, levels of H3K9me3 were pronounced in regions both upstream and downstream of the GAA repeats in the GFP\textsubscript{(GAA\textsubscript{TTC})_{560}} cells, not the GFP\textsubscript{(GAA\textsubscript{TTC})_{0}} cells. After initiation of transcription by the tetracycline treatment, no significant changes of H3K9me3 levels were observed in the proximity of the GAA repeats in the GFP\textsubscript{(GAA\textsubscript{TTC})_{560}} cells when compared to the GFP\textsubscript{(GAA\textsubscript{TTC})_{560}} cells that is not treated with tetracycline. However, enrichments of H3K79me2 were
dramatically reduced at the region downstream of the GAA repeats in the
GFP_{(GAA•TTC)}_{560} cells upon inhibition of transcription as I expected. This result
strongly suggests that the hyperexpanded GAA repeats per se are the underlying cause of
the heterochromatin-like structure formation and the arrest of RNA pol II is unlikely to
serve as a signal instigating epigenetic changes associated with the longer GAA repeats.
Taken together, my results demonstrate that insufficient levels of FXN expression in
FRDA patients are a consequence of impeded transcription elongation through the
heterochromatin-like structure induced by the hyperexpanded GAA repeats.
Figure 27. The tract of GAA repeats is sufficient to form the heterochromatin-like structure in the absence of transcription, which lead to inhibition of transcription elongation.

Effect of transcription on the formation of the heterochromatin-like structure was analyzed using ChIP assay in the GFP_GAA reporter system. Enrichment of H3K9me3 and H3K79me2 at the region upstream (Up) and the region downstream (Down) were determined in the absence and presence of tetracycline (0.1 µg/mL for 24 h). White bars represent results obtained from cell line lacking the GAA repeats in intron of the GFP gene while gray bars representing cells harboring intronic 560 of GAAs. Data are expressed as mean ± SEM. The experiment was conducted in triplicate and $P$ values were calculated from the average of all three determinations for the GFP_(GAA•TTC)_{560} and the GFP_(GAA•TTC)_{0} cells. H3K9me3 is enriched in the GFP_(GAA•TTC)_{560} in the absence and presence of tetracycline. H3K79me2 are impaired in the GFP_(GAA•TTC)_{560} upon tetracycline treatment to resume transcription. Permission to re-use by Oxford University Press on behalf of the Nucleic Acids Research.
Figure 27
Chapter 5. Discussion

Some parts of this dissertation were modified from the following journal article:

In my study, defining altered histone modifications and repressive RNA pol II
distribution at the \textit{FXN} gene in FRDA cell lines provides strong evidence that \textit{FXN}
insufficiency in FRDA cell lines begin at a post-initiation step of \textit{FXN} transcription. I
also detected an accumulation of total RNA pol II at the promoter-proximal region of the
\textit{FXN} gene in unaffected cells, but not in FRDA cells. Several studies using genome-wide
analysis have suggested that post-initiation events are rate-limiting to regulate
transcription of some genes (Guenther, Levine, Boyer, Jaenisch and Young, 2007;
Saunders, Core and Lis, 2006). This phenomenon is well-described for a heat shock–
inducible gene, \textit{Hsp70}, in \textit{Drosophila} (Gilmour and Lis, 1986). Increased levels of RNA
pol II are observed at the promoter of the \textit{Hsp70} gene upon uninduced condition, and this
RNA pol II is released from the promoter by heat shock stimuli, leading to launching of
\textit{Hsp70} transcription (Gilmour and Lis, 1986). Genome-wide analysis of transcripts using
nuclear run-on assay reveals paused Pol II on a large number of genes in human cells
(Core, Waterfall and Lis, 2008). This phenomenon prompts me to hypothesize that \textit{FXN}
gene may be categorized as an inducible gene that is regulated by tissue-specific factors.
The \textit{FXN} gene is positively regulated by two transcription factors, peroxisome
proliferator-activated receptor (PPAR) gamma and hypoxia-inducible factor-2 alpha
(Marmolino, Acquaviva, Pinelli, Monticelli, Castaldo, Filla and Cocozza, 2009; Oktay,
Dioum, Matsuzaki, Ding, Yan, Haller, Szweda and Garcia, 2007). PPAR gamma is more
highly expressed in FRDA cells compared to unaffected individuals (Marmolino,
Acquaviva, Pinelli, Monticelli, Castaldo, Filla and Cocozza, 2009). Imbalance of iron as
detected in FRDA patients may also affect the induction of \textit{FXN} expression, as proposed
in a previous study (Li, Besse, Ha, Kouvunovych and Rouault, 2008). The longer GAA
repeats at the *FXN* gene in FRDA patient may recruit particular factors that inhibit induction of FXN expression. In contrast, non-B DNA structure adopted by the longer GAA repeat *per se* may inhibit to recruit particular factors that induce of FXN expression.

According to my results, paused RNA pol II is unlikely to reflect ongoing transcription, at least at the *FXN* gene. Although *FXN* pre-mRNA levels are similar in both unaffected and FRDA cells, total RNA pol II, as measured by ChIP, only accumulates at the promoter-proximal region of the *FXN* gene in FRDA cells. Moreover, this accumulation is restricted to the promoter-proximal region, not other regions including promoter. Therefore, *FXN* pre-mRNA seems to reflect FXN promoter activity more than chromatin structure at the region. This finding is consistent with those of a previous study that indicated promoter activity is not affected at the *FXN* gene in FRDA cells, although altered histone modifications were observed at the promoter-proximal region (Kumari, Biacsi and Usdin, 2011b). In an effort to confirm correlation between the levels of RNA pol II and FXN transcription, I tried to measure pre-mature FXN RNA at the promoter-proximal region (+48bp). Unfortunately, the *FXN* promoter-proximal region is quite CG-rich, so that qRT-PCR assays were not successful. Both H3K4me2 and H3K4me3 are greatly decreased at this region in FRDA cells when compared to unaffected cells. This result suggests that the paused RNA pol II at the promoter-proximal region in unaffected cells may facilitate chromatin remodeling for euchromatin as shown in my results to trigger additional transcription progression. The paused RNA pol II may help to stabilize the binding of transcription machinery assisting further
rounds of transcription initiation. Indeed, the accumulated RNA pol II in the promoter-proximal region is associated with the enrichment of active histone marks such as H3K9/14ac and H3K4me3 in the human genome, regardless of transcription status (Guenther, Levine, Boyer, Jaenisch and Young, 2007). FRDA cells exhibit a slow rate of $FXN$ mRNA synthesis when compared to unaffected cell lines (Punga and Buhler, 2010).

The decreased levels of RNA pol II in the coding region of genes is associated with decreased levels of H3K4me2 and H3K4me3, hyper-methylated CpG residues, and increased levels of H3K9me (Okitsu and Hsieh, 2007; Okitsu, Hsieh and Hsieh, 2010). Consistently, levels of total RNA pol II is repressed in the region upstream of the GAA repeats in FRDA cells compared to unaffected cells. Moreover, the region upstream of the GAA repeats display less enrichment of H3K4me3 and H3K4me2 in FRDA cells compared to unaffected cells. Enrichment of H3K4me3 in the coding region of genes is associated with the efficiency of transition between initiation and elongation in transcription (Okitsu and Hsieh, 2007; Okitsu, Hsieh and Hsieh, 2010), and contributes to recruit chromatin-remodeling complexes (Li, Ilin, Wang, Duncan, Wysocka, Allis and Patel, 2006; Sims, Millhouse, Chen, Lewis, Erdjument-Bromage, Tempst, Manley and Reinberg, 2007). Therefore, the decreased levels of H3K4me3 at the region upstream of GAA repeats in FRDA cells may impair recruitment of chromatin-modifying complexes or lead to $FXN$ deficiency resulting from impediment of $FXN$ transcription elongation in FRDA patients.

Histone methylation status represents both an important regulatory element of gene
expression and an indicator of the transcriptional status of a gene. Enzymes that add or remove methyl groups on specific sites of histones regulate the balance between methylation and demethylation in the cells. Activity of a large group of histone demethylases containing Jumonji C-terminal domain depends on the presence of non-heme iron (Hahn, Bose, Edler and Lengeling, 2008). Several reports show that low levels of FXN, an iron chaperone, results in accumulated irons in the mitochondria and depleted iron in the other cellular compartments leading to a severe imbalance of iron in FRDA cells (Adinolfi, Iannuzzi, Prischi, Pastore, Iametti, Martin, Bonomi and Pastore, 2009; Bulteau, O’Neill, Kennedy, Ikeda-Saito, Isaya and Szweda, 2004). This suggests that the imbalance of iron in FRDA patients may be associated with the activity of iron-dependent demethylases responsible for levels of H3K27me3. In FRDA cells, compared to controls, I detected significant increase of H3K27me3, histone methylation marks potentially affected by iron-dependent histone demethylases such UTX. Indeed, over expression of UTX increases expression of FXN mRNA in KYSE180 cancer cells (van Haaften, Dalgliesh, Davies, Chen, Bignell, Greenman, Edkins, Hardy, O'Meara, Teague, Butler, Hinton, Latimer, Andrews, Barthorpe, Beare, Buck, Campbell, Cole, Forbes, Jia, Jones, Kok, Leroy, Lin, McBride, Maddison, Maquire, McLay, Menzies, Mironenko, Mulderrig, Mudie, Pleasance, Shepherd, Smith, Stebbings, Stephens, Tang, Tarpey, Turner, Turrell, Varian, West, Widaa, Wray, Collins, Ichimura, Law, Wong, Yuen, Leung, Tonon, DePinho, Tai, Anderson, Kahoski, Massie, Khoo, Teh, Stratton and Futreal, 2009). Moreover, genome-wide approach reveals that FXN deficiency is associated with down regulation of genes involved in DNA packaging and nucleosome assembly (Huang, Becker, Whitnall, Rahmanto, Ponka and Richardson, 2009).
In an effort to alleviate transcriptional block facilitated by the GAA repeat expansions in FRDA patients, dissecting the mechanism of heterochromatin-like structure induced by the hyperexpanded GAA repeats is critical for logical design of therapeutic approaches. Currently, HDACis are the most promising small molecules that restore $FXN$ transcription and alter the chromatin status in the proximity of the GAA repeats, but not the promoter region of the $FXN$ gene in FRDA cells (Herman, Jenssen, Burnett, Soragni, Perlman and Gottesfeld, 2006). However, the question of either the initial event or the cascade of epigenetic changes in the GAA repeats region remains to be addressed. Perhaps the hyperexpanded GAA repeats per se or non-B DNA structures facilitated by the hyperexpanded GAA repeats employ particular proteins that affect modifications on histones in the chromatin environment in the vicinity of the GAA repeats. Moreover, paused RNA pol II at the c-$Myc$ gene is observed in the vicinity of a triplex forming repeat structure (Belotserkovskii, De Silva, Tornaletti, Wang, Vasquez and Hanawalt, 2007), suggest that non-canonical DNA structures induced by the hyperexpanded GAA repeats might be a primary origin of deficient $FXN$ transcription in FRDA.

In fission yeast and Drosophila, and possibly other eukaryotes, the RNA interference (RNAi) pathway is involved in heterochromatin formation (Wassenegger, 2005). Endogenous short interfering RNA (siRNA) originates from repetitive sequences such as centromeric repeats, transposable elements or rDNA and act as a trigger the RNAi pathway (Fagegaltier, Bouge, Berry, Poisot, Sismeiro, Coppee, Theodore, Voinnet and Antoniewski, 2009). A genome-wide approach revealed that there are antisense
transcripts at the *FXN* locus in humans (He, Vogelstein, Velculescu, Papadopoulos and Kinzler, 2008). Recently, more antisense transcripts (*FAST*-1) were observed in the proximity of TSS of the *FXN* gene in FRDA cells (De Biase, Chutake, Rindler and Bidichandani, 2009). Antisense *FXN* transcripts can be involved in prompting silencing of *FXN* transcription via RNAi-mediated heterochromatin formation through double stranded RNA generated from repetitive DNA sequences. In an effort to assess the relative contribution of antisense *FXN* transcripts in regulation of *FXN* expression, we also performed strand-specific RT-PCR to detect antisense transcription in the *FXN* locus. Unfortunately, we were unable to detect antisense transcripts at the regions specified in the previous study, the proximity of TSS of the *FXN* gene as well as region upstream of the GAA repeats (data not shown). Therefore, all *FXN* transcripts examined in this study represent sense FXN RNA transcripts.

Most diseases associated with repetitive DNA sequences exhibit distinct patterns of repetitive element instability in neuronal as compared to non-neuronal cells (Pearson, Nichol Edamura and Cleary, 2005). Indeed, although GAA contractions occur in all tissues, expansions occur only in the brain of FRDA patients (Cleary and Pearson, 2003). The tissue specificity could contribute to disease progression (Swami, Hendricks, Gillis, Massood, Mysore, Myers and Wheeler, 2009). DNA metabolism such as replication, repair, recombination and transcription, is important for instability of the repetitive DNA unit. Notably, abnormal function of mismatch repair proteins, such as MSH2, is associated with GAA instability in yeast (Shishkin, Voineagu, Matera, Cherng, Chernet, Krasilnikova, Narayanan, Lobachev and Mirkin, 2009) and in inducible pluripotent stem
(iPS) cells from FRDA patients (Ku, Soragni, Campau, Thomas, Altun, Laurent, Loring, Napierala and Gottesfeld, 2010). Similar to FRDA, base excision repair factors are associated with CTG repeat stability (Liu, Prasad, Beard, Hou, Horton, McMurray and Wilson, 2009). Tissue specific levels of these trans-acting factors may contribute to increased instability of repetitive DNA sequence in neuronal cells of the patients. Also, cis-elements, such as the CTCF DNA binding motif, are associated with tissue and developmental specificity of increased repeat instability (Pearson, Nichol Edamura and Cleary, 2005). CTCF DNA binding motifs flank numerous genes that contain repetitive DNA units and are known to modulate chromatin structure and transcription in several neurological disorders (Chang, Zhang, Heath, Galjart, Wang and Milbrandt, 2010; Pearson, Nichol Edamura and Cleary, 2005). Methylation of cytosine or mutation of CTCF-binding motif at the SCA7 locus inhibits CTCF binding, thereby increasing tissue specific repeat instability in SCA7 (CAG repeats) (Libby, Hagerman, Pineda, Lau, Cho, Baccam, Axford, Cleary, Moore, Sopher, Tapsccott, Filippova, Pearson and La Spada, 2008). In FRDA, more CTCF occupancy are accompanied with heterochromatinized structures at the FXN genes, which lead to transcribe antisense FXN RNAs (De Biase, Chutake, Rindler and Bidichandani, 2009). However, distinct levels of CTCF in neuronal cells have not yet been reported in neurodegenerative diseases affected by repetitive DNA sequences. Moreover, occupancy of CTCF is not correlated with overt transcription in mammals (Splinter, Heath, Kooren, Palstra, Klous, Grosveld, Galjart and de Laat, 2006). This suggests that upstream tissue-specific factors that regulate CTCF may contribute to tissue specificity of repeat instability. Methylations at the promoter of the BDNF gene, a growth factor that plays important roles in regulating neurogenesis,
synaptic plasticity, and neuronal survival, are associated with neurodegenerative disorders, including Huntington's disease (Zuccato and Cattaneo, 2009). Hypermethylated CpG residues in the BDNF gene in the cortical neuron in neurodegenerative conditions inhibits CTCF binding leading to silence of BDNF expression (Chang, Zhang, Heath, Galjart, Wang and Milbrandt, 2010). This report may shed light on the mechanistic relationship between pathology of neurodegenerative diseases and repeat instability in neuronal cell types.

Moreover, CTCF is associated with chromatin looping and altered histone modifications (Splinter, Heath, Kooren, Palstra, Klous, Grosveld, Galjart and de Laat, 2006). I found that pausing of RNA pol II occurred at the promoter-proximal region, not the region upstream of the GAA repeat at the FXN gene. This phenomenon might suggest looping of the chromosome from the TSS to a region upstream of the GAA repeats at the FXN gene in FRDA cells. The hyperexpanded GAA repeats influence the looping in FRDA cells, thereby affect not only chromatin structure but also FXN transcription. In reconstitution assays, nucleosomes are less well assembled into double stranded DNA sequences containing longer GAA repeats (Ruan and Wang, 2008). The hyperexpanded GAA repeat that have less nucleosome assembly might have higher opportunity to be looped between TSS and the region immediately upstream of the GAA repeats, thereby affecting configuration of chromatin structure and distribution of RNA pol II at the FXN gene in FRDA cells.
Expansions of repetitive DNA sequences are correlated with human neurological and neuromuscular diseases. To date, at least 15 neurodegenerative disorders are identified as the repeats-expansion diseases (Wells, 1996). These lengths of repetitive DNA unit are dynamic in among individuals, generations, and tissues of the patients. There are many evidences that these repeat expansion diseases have common epigenetic features such as heterochromatin associated histone modifications, the presence of short/antisense RNA, and hypermethylated CpG residue, and CTCF binding in the proximity of the repeats (Al-Mahdawi, Pinto, Ismail, Varshney, Lymperi, Sandi, Trabzuni and Pook, 2008; Cho, Thienes, Mahoney, Analau, Filippova and Tapscott, 2005; Filippova, Thienes, Penn, Cho, Hu, Moore, Klesert, Lobanenkov and Tapscott, 2001; Pietrobono, Tabolacci, Zalfa, Zito, Terracciano, Moscato, Bagni, Oostra, Chiurazzi and Neri, 2005; Verdel, Jia, Gerber, Sugiyama, Gygi, Grewal and Moazed, 2004). These similarities in epigenetic pathologies on silenced alleles of each repeat expansion diseases suggest similar therapeutic approaches to enhance the gene expression in an effort to alleviate disease symptom. However, some of features are excluded when epigenetic pathology of each repeat expansion disease defines, suggesting that distinct mechanisms responsible for repeat-mediated chromatin remodeling in each repeat expansion disease may be present. Therefore, my study that carefully define chromatin structures adopted by the hyperexpanded GAA repeat will help not only understand shared epigenetic features in repeat expansion diseases but also distinguish epigenetic features of FRDA from other repeats expansion diseases.
Chapter 6. Significance and Future Directions

Some parts of this dissertation were modified from the following journal article:

My studies define comprehensive distribution of histone modification at the \textit{FXN} gene in multiple lymphoid cell lines derived from FRDA patients and unaffected controls (Figure 28). I found that the hyperexpanded GAA repeats influence chromatin structure in the regions flanking the GAA repeats, but these differences do not spread to the promoter. On the contrary, the levels of RNA pol II and histone methylations representing transcription elongation were significantly decreased at the \textit{FXN} gene in FRDA cell lines. These results strongly demonstrate that the transcription machinery is inhibited after successful initiating \textit{FXN} transcription in FRDA cells. Moreover, I first show that RNA pol II is predominant at the \textit{FXN} promoter-proximal region, appears to be paused, in unaffected cell lines, but not FRDA cell lines. Comparison of altered H3K4me3 and H3K79me2 levels at the \textit{FXN} gene revealed that \textit{FXN} deficiency is associated with defective efficiency of transition from initiation to elongation of \textit{FXN} transcription.

Taken together, my study provides strong indication that \textit{FXN} insufficiency in FRDA patients is caused by a defective transition at a post-initiation step of \textit{FXN} transcription. Perhaps heterochromatin-like structures in the vicinity of the hyperexpanded GAA repeats act as an obstacle to produce full length \textit{FXN} mRNA.

Regulatory elements located in the first intron affect expression of several genes (De Jaco, Camp and Taylor, 2005; Lee, Dahi, Mahimkar, Tulloch, Alfonso-Jaume, Lovett and Sarkar, 2005). Studies using deletion mapping of the \textit{FXN} intron 1 revealed that the intron 1 is capable of regulating \textit{FXN} expression (Greene, Mahishi, Entezam, Kumari and Usdin, 2007). Results of ChIP assay in this study demonstrate that the intronic region upstream of the hyperexpanded GAA repeats exhibit the most significant differences of
Figure 28. A summary of distributed histone modifications at the human *FXN* gene.

In FRDA cells, the *FXN* gene enriched with constellation of H3K9me3, H3K27me3, and H4K20me3, strongly supporting that the *FXN* gene in FRDA cells has heterochromatin-like structures due to possibly the hyperexpanded tract of GAA repeats. This phenomenon demonstrated in the GFP_GAA reporter system. H3K4me2, H3K36me3, and H3K79me2 associated with euchromatin are decreased at the *FXN* gene in FRDA cells as well as the *GFP* gene in the GFP_(GAA•TTC)_560 cells of the GFP_GAA reporter system. Distribution of RNA pol II is repressed at the entire FXN gene in FRDA cells compared to unaffected cells. RNA pol II seems to be poised at the downstream of the transcription start site (TSS) of the *FXN* gene in unaffected cells, not FRDA cells. Although the detailed mechanisms is unclear, heterochromatin-like structure at the vicinity of the GAA repeats induced by the hyperexpanded GAA repeats lead to consequence of *FXN* deficiency at the region downstream of the GAA repeats in FRDA cells.
Figure 28
histone modifications between FRDA and control cell lines. The cytosine residue in this region is preferentially methylated in FRDA patients (Greene, Mahishi, Entezam, Kumari and Usdin, 2007). The levels of DNA methylation proportionally correlate with the length of the GAA repeats (Castaldo, Pinelli, Monticelli, Acquaviva, Giacchetti, Filla, Sacchetti, Keller, Avvedimento, Chiariotti and Cocozza, 2008). One of the methylated cytosine is localized into an E-box motif (CANNTG). The c-Myc, a bHLH transcription factor that recognize the E-box motif, interacts with the E-box resided in the region upstream of the repeats in P493 cells (Fernandez, Frank, Wang, Schroeder, Liu, Greene, Cocito and Amati, 2003). Moreover, the induction of c-Myc is associated with enhancement of FXN transcription (Schuhmacher, Kohlhuber, Holzel, Kaiser, Burtscher, Jarsch, Bornkamm, Laux, Polack, Weidle and Eick, 2001). Myc proteins allow the paused RNA pol II to be released from promoter to activate gene expression (Rahl, Lin, Seila, Flynn, McCuine, Burge, Sharp and Young, 2010). c-Myc physically interact with components from chromatin modifying/remodeling complexes such as GCN5 to activate gene expression (Flinn, Wallberg, Hermann, Grant, Workman and Wright, 2002; Kenneth, Ramsbottom, Gomez-Roman, Marshall, Cole and White, 2007). In the unaffected cells, enrichments of H3K4me3, H3K4me2 and H3K79me2 in the region upstream of the GAA repeats encompassing E-box motif potentially facilitate binding of myc proteins in the E-box. Therefore, hypermethylated DNA sequences in the region upstream of the GAA repeats in FRDA cells can influence the interactions between myc proteins and the E-box motif (Guccione, Martinato, Finocchiaro, Luzi, Tizzoni, Dall'Olio, Zardo, Nervi, Bernard and Amati, 2006; Perini, Diolaiti, Porro and Della Valle, 2005). My study here emphasizes the region upstream of the repeats is valuable as the
regulatory element leading to appropriate \( FNX \) transcription. It is likely that epigenetic changes in the intron 1 of the \( FNX \) gene induced by the hyperexpanded GAA repeats interfere with transcriptional enhancers. Subsequently, reduced progression of transcriptional machinery result in defective \( FNX \) expression in FRDA patients. Additionally, c-Myc has been shown to interact with GCN5, an acetyltransferase responsible for H3K9/14ac in eukaryotes (Flinn, Wallberg, Hermann, Grant, Workman and Wright, 2002; Kenneth, Ramsbottom, Gomez-Roman, Marshall, Cole and White, 2007). The decreased levels of H3K9/14ac at the region upstream of the GAA repeats in FRDA cells might result from failure of GCN5 recruitment at the region upstream of the GAA repeats in FRDA cells. Reduced occupancy (or interaction) of C-Myc and GCN5 in FRDA cells may lead to reduced recruitment of transcription machinery containing RNAPII. This impaired recruitment may result in particular constellation of histone methylations marking repressive chromatin structures at \( FNX \) gene in FRDA patients as I observed. In order to investigate this possibility, distribution of c-Myc and GCN5 could be defined by ChIP-seq in the lymphoid cell lines. Unfortunately, there are no appropriate antibodies for GCN5 and c-Myc ChIP. Alternatively, distribution of histone modification at the \( FNX \) gene can be determined in either the absence or over expression of either c-Myc or GCN5 in lymphoid cell lines. However, the lymphoid cell lines exhibit technically low efficiency of transfection. It would take longer time to set a condition to optimize transfection activity. Alternatively, other cell lines such as K562 and HEK293T cells as general cell lines can be utilized to investigate effect of c-Myc or GCN5 in \( FNX \) transcription. However, this system will exclude effect of the hyperexpanded GAA repeats to understand the roles of c-Myc or GCN5 in \( FNX \) transcription. The GFP_GAA
system will be a suitable system. However, this system also has limitations that the artificial condition in the GFP_GAA system may not be correlated with actual cellular function to the heterochromatin-like structure at the \textit{FXN} gene in FRDA patients. Moreover, since several studies suggest that intronic region of the \textit{FXN} gene have regulatory function in \textit{FXN} transcription, the GFP_GAA system has weakness that have non-F\textit{FXN} intronic sequence to regulate either GFP transcription and chromatin structure change.

During transcription consisting of series of phase such as initiation, elongation, and termination of RNA transcript by RNA pol II, the promoter-proximal pausing of RNAPII is a rate-limiting step to produce full-length transcription(Guenther, Levine, Boyer, Jaenisch and Young, 2007). I found that levels of RNA pol II is impaired at the promoter-proximal region of the \textit{FXN} gene in FRDA cells. More interestingly, the SAGA complex that contains GCN5 and the Mediator complex have important roles in inducing appropriate levels of gene expression from poised promoters in budding yeast (Lee, Fletcher, Zhang, Chen, Fischbeck and Stargell, 2010). Moreover, c-Myc is involved in the release of paused RNA pol II from the promoter-proximal region of a given gene (Rahl, Lin, Seila, Flynn, McCuine, Burge, Sharp and Young, 2010). These results raise the possibilities that c-Myc bound to the E-box in the region upstream of the GAA repeats might be involved in release of paused RNA pol II at the promoter-proximal region of the \textit{FXN} gene.

To further confirm that paused RNA pol II regulates FXN expression, I would examine
additional transcription factors, DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) as well as P-TEFb using conventional ChIP technique. DSIF and NELF are sufficient to stall RNA pol II at the promoter-proximal region of genes (Wu, Yamaguchi, Benjamin, Horvat-Gordon, Washinsky, Enerly, Larsson, Lambertsson, Handa and Gilmour, 2003). The paused RNA pol II is released from the promoter-proximal region by phosphorylation of NELF by P-TEFb (Rahl, Lin, Seila, Flynn, McCuine, Burge, Sharp and Young, 2010). The P-TEFb is necessary to release the paused RNA pol II by phosphorylation of NELF (Peterlin and Price, 2006). Inhibition of P-TEFb induces accumulated RNA pol II at the most of genes, suggesting that majority of gene is regulated by pausing of RNA pol II (Cheng and Price, 2007; Peterlin and Price, 2006). C-myc is associated with activation of P-TEFb to release paused RNA pol II (Rahl, Lin, Seila, Flynn, McCuine, Burge, Sharp and Young, 2010). Successful release of the paused RNA pol II from the promoter-proximal region of genes is necessary to elongate full-length transcripts (Saunders, Core and Lis, 2006). Artificial recruitment of P-TEFb at the Hsp70 gene under uninduced conditions increases basal Hsp70 expression (Lis, Mason, Peng, Price and Werner, 2000). This result suggests that the FXN gene in FRDA may have less P-TEFb compared to unaffected controls. To investigate this possibility, I could first treat pharmacological inhibitor of P-TEFb in lymphoid cell lines followed by ChIP for total RNA pol II at the FXN gene.

Phosphorylation at serine 2 in the CTD of RNA pol II by P-TEFb is indicator of ongoing elongation of full-length transcripts (Peterlin and Price, 2006). Unfortunately, I couldn’t detect appreciative levels of phosphorylated serine 2 of RNA pol II by conventional
ChIP. Instead, nuclear run on assay is possibly able to detect elongating full-length $FXN$ transcripts. If paused RNA pol II is associated with regulation of FXN transcription, I would detect increased phosphorylated RNA pol II at serine 2 upon the treatment of pharmacological P-TEFb inhibitor in FRDA cells.

Levels of H3K9me3 have been less correlated with $FXN$ transcription, suggesting that levels of H3K9me3 are not the primary means to induce the heterochromatin-like structure facilitated by the GAA repeats. This suggestion allows me to hypothesize that hyper DNA methylation at the region upstream of the GAA repeats in FRDA cells is associated with altered histone modifications as is observed in my studies. Although the relationship between DNA methylation and histone modifications has not been fully understood, these two systems may play a part in regulating repressive chromatin structures. To investigate the effect of the DNA methylation in $FXN$ transcription as well as altered chromatin structure at the $FXN$ gene in FRDA cells, I would examine distribution of more histone modifications and total RNA pol II after removal of DNA methylation using 5-aza-2-deoxycytidine, an inhibitor of DNA methylation. If there is no change in histone modifications, it would further demonstrate that hyper DNA methylation in FRDA patients is involved in a different mechanism to silence $FXN$ transcription in FRDA patients. If there are changes in histone modifications associated with the transcription process after 5-aza-2-deoxycytidine treatments, I will be able to determine which step of $FXN$ transcription is affected by high levels of the DNA methylation at the region upstream of the GAA repeats in FRDA patients. To further demonstrate that the observations after removal of DNA methylation result from the
hyperexpanded GAA repeats *per se*, not either inherent characteristics of the *FXN* gene, I will take advantage of the GFP_{GAA} system. First of all, I will examine the levels of DNA methylation at the region upstream of the GAA repeats in the GFP_{GAA} cell lines to see the expanded GAA repeats induce DNA methylation in the vicinity of the repeats using either MS-PCR or bisulphate sequencing. If there are no DNA methylation changes in the GFP_{GAA} cells, one possibility is that the GAA repeats do not induce DNA methylations. The histone modification changes by the GAA repeats are primary factors to restore *FXN* transcription in FRDA. Another possibility is that 560 GAA repeats insufficient to induce DNA methylation to silence gene expression since the DNA methylation is correlated with the length of the GAA repeats in FRDA patients. If the GFP_{(GAA\cdot TTC)\_{560}} cells have more DNA methylation compared to the GFP_{(GAA\cdot TTC)\_{0}} cells, I should determine which factor between DNA methylation and histone modification can be more important/first targets to restore *FXN* transcription by additional experiments.

Although we detected no antisense *FXN* transcripts at the *FXN* locus, we cannot exclude possibility that there are double strand RNAs originating from the hyperexpanded GAA repeats to form RNAi-mediated heterochromatin-like structures at the *FXN* gene in FRDA cells. Moreover, the presence of H3K9me3 overlaps with RNAi-mediated heterochromatin formation in fission yeast (Cam, Sugiyama, Chen, Chen, FitzGerald and Grewal, 2005). Although I observed significant levels of H3K9me3 at the region upstream of the GAA repeat in FRDA cells, I could not detect any isoforms of HP1 protein. This suggest different pathway to form heterochromatin-like structure at the *FXN*
gene. The argonaute protein is a key player to participate in RNAi-mediated gene silencing as part of the RNA-Induced Transcriptional Silencing (RITS) complex mediating RNAi pathway to heterochromatin assembly (Buker, Iida, Buhler, Villen, Gygi, Nakayama and Moazed, 2007). The identifying small RNAs and proteins associated with the RITS complex at the \(FXN\) gene in FRDA will determine if the hyperexpanded GAA repeat is involved in RNAi-mediated \(FXN\) silencing via assembly of heterochromatin in FRDA.

Although expansion of the GAA repeats is the primary factor driving chromatin changes, the possibilities that metabolic changes resulting from frataxin deficiency in FRDA cells contribute to shaping of epigenetic landscape at the \(FXN\) gene has not been excluded. Frataxin is associated with protestation of cells against cytotoxic stress that might be involved in the fine-tuning regulation of FXN expression (Jiralerspong, Ge, Hudson and Pandolfo, 2001).

Moreover, HDACi treatments have no association of H3K9me3 levels at the \(FXN\) gene in FRDA cells in despite of enhancing \(FXN\) expression (Herman, Jenssen, Burnett, Soragni, Perlman and Gottesfeld, 2006; Punga and Buhler, 2010). This result suggests that inhibition of acetylation on non-histone is involved in the regulation of \(FXN\) transcription as well as chromatin structure changes at the \(FXN\) gene in FRDA cells. Therefore, global analyses of nucleosome position and histone modifications in FRDA are necessary to reveal the full spectrum of epigenetic consequences of reduced frataxin levels.
Overall, my study will help to clarify underlying mechanisms of how epigenetic changes observed in the vicinity of the hyperexpanded GAA repeats at the \textit{FXN} gene are associated with \textit{FXN} deficiency in FRDA patients. Furthermore, this study can potentially be developed to investigate therapeutic targets (\textit{e.g.} methyltransferase or demethylase responsible for altered histone modifications observed in my studies) for curing FRDA.
**Figure 29. A working model of heterochromatin-like structure formation at the FXN gene in FRDA.**

HDACs are recruited to the FXN gene and remove acetyl group at the N-terminal tail of the histones. Histone lysine methyltransferase add methyl group at the N-terminal tail of the histones in the vicinity of the hyperexpanded GAA repeats in FRDA cells. According to the distribution of RNA pol II in the lymphoid cell lines, the fate of RNA pol II is first likely to be determined at the promoter-proximal region (+48 bp) to regulate either FXN expression or chromatin structures. At the region upstream of the GAA repeats, RNA pol II would be inhibited to elongate FXN transcription through one of mechanisms in following, which lead to FXN deficiency in FRDA patients. One is that either non-B DNA structure or heterochromatin-like structure induced by the hyperexpanded tract of GAA repeats inhibit recruitment of chromatin modifying/remodeling complexes or DNA methyltransferase. Another is that methylated CpG residue at the E-box (CACGTG) in FRDA inhibits binding of transcription factors such as c-Myc, which lead to low levels of RNA pol II and altered histone modifications resulting in FXN deficiency in FRDA. Moreover, antisense FXN transcripts from the hyperexpanded GAA repeats in FRDA patients can be involved in small RNA-mediated heterochromatin formation.
Figure 29

Unaffected

FRDA

| RNA polymerase II (RNAPII) | Lysine methyltransferase (KMT) | Lysine acetyltransferase (KAT) | Lysine deacetyltransferase (KDAT) | Other Lysine methyltransferase (KMT) | Nucleosome | Transcription factors | DNA methylation |
Bibliography


binding factor (CTCF)/cohesin binding and transcription at the BDNF locus. Proc Natl Acad Sci U S A 107, 21836-21841.


Eunah Kim was born in Seoul, South Korea on March 19th, 1979. Eunah is the daughter of Sam-san Kim and Geum-jeong Yoon. After completing high school in 1998, she attended Hanyang University, South Korea majoring in Biochemistry & Molecular Biology. Eunah graduated from Hanyang University in February 2002 with a degree of Bachelor of Arts in Biochemistry & Molecular Biology. In graduate school, she performed research in Dr. Young-Gyu Chai’s laboratory in the Department of Biochemistry. Eunah earned a degree of Master of Science in Biochemistry in February 2004. During her undergraduate and graduate, she received scholarships in Department of Biochemistry & Molecular biology for Fall 1998, Fall 1999 and 2002 - 2003. After completing a degree of Master of Science, as a research assistant, she performed researches in Dr. Eun-Kyoung Choi’s laboratory in the Ilsong Life Science Institute at Hallym University, South Korea. In August of 2007, she entered the University of Texas Graduate School of Biomedical Sciences in Houston, Texas. In May of 2008, she joined the laboratory of Dr. Sharon Dent at the University of Texas M.D. Anderson Cancer Center for her dissertation research. In graduate school, she was awarded the Schissler Fellowship in Human Genetics.

Publications:


*Jang, B. and Kim, E. contributed equally to this paper.*

Permanent address:

2222 Maroneal st. #732

Houston, TX 77030