[Frontiers in Bioscience 12, 4731-4741, May 1, 2007]

Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease

Yuh-Hwa Wang

Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, NJ 08854, USA.

TABLE OF CONTENTS

1. ABSTRACT

In eukaryotic cells, chromatin structure organizes genomic DNA in a dynamic fashion, and results in regulation of many DNA metabolic processes. The CTG/CAG and CGG/CCG repeating sequences involved in several neuromuscular degenerative diseases display differential abilities for the binding of histone octamers. The effect of the repeating DNA on nucleosome assembly could be amplified as the number of repeats increases. Also, CpG methylation, and sequence interruptions within the triplet repeats exert an impact on the formation of nucleosomes along these repeating DNAs. The two most common triplet expansion human diseases, myotonic dystrophy 1 and fragile X syndrome, are caused by the expanded CTG/CAG and CGG/CCG repeats, respectively. In addition to the expanded repeats and CpG methylation, histone modifications, chromatin remodeling factors, and noncoding RNA have been shown to coordinate the chromatin structure at both myotonic dystrophy 1 and fragile X loci. Alterations in chromatin structure at these two loci can affect transcription of these disease-causing genes, leading to disease symptoms. These observations have brought a new appreciation that a full understanding of disease gene expression requires a knowledge of the structure of the chromatin domain within which the gene resides.

2. INTRODUCTION

Chromatin structure can have dominant effects on gene activities and, if a negative consequence arises, may lead to human diseases. The two most common human triplet expansion diseases, myotonic dystrophy and fragile X syndrome, have provided examples in which altered chromatin structure plays a role in disease pathogenesis. In the human cell, DNA including expanded triplet repeats is packaged progressively beginning with the wrapping of 146 base pairs of DNA around a histone protein octamer to form nucleosomes (1). This string of nucleosomes is further condensed by the binding of histone H1 into chromatin fibers ~30 nm in diameter (2). The 30 nm fibers are organized as a series of loops occurring about once every 40~100 kb, and bound to the nuclear matrix or scaffold by sequences called matrix or scaffold attachment regions (MARs or SARs), creating distinct chromosomal domains. Further compaction driven by proteins such as the condensing complex and topoisomerase II, gives rise to the condensation levels in metaphase chromosomes (3). The organization of chromosome architecture can be regulated at many levels. Since chromatin structure is critical to the proper regulation of DNA metabolic processes (4), factors influencing the formation of chromatin structure could alter those processes.

The primary sequence of the DNA itself has been shown to affect assembly of nucleosomes. Extensive studies have revealed that disease-causing triplet repeats possess differential abilities for the binding of histone octamers. For example, we and, subsequently, others have found the repeating CTG/CAG sequence to be the strongest known natural nucleosome-positioning element (5-7). This unusual ability of triplet repeats to assemble nucleosomes could be greatly amplified as the expansion of repeats increases. Since the ability of nucleosome assembly fundamentally determines the effect of chromatin structure on many DNA processes (transcription and replication), the net result of expansion of pure repeats could be dramatic. In contrast, sequence interruptions within repeat tracts could diminish such an outcome by preventing the repeat phenomena (8). In addition to primary sequence, modification of DNA bases, such as CpG methylation, can create binding sites for proteins to alter chromatin structure. The CGG/CCG repeats contain intrinsic CpG sites for methylation, and these sites are indeed methylated in the expanded forms (9). Other disease-causing triplet repeats such as CTG/CAG and GAA/TTC, cannot be methylated themselves, but have CpG sites in the vicinity, and their methylation could influence chromatin structure as well (10, 11). Moreover, several classes of proteins - histone variants, histones with covalent modifications, and chromatin remodeling factors, are known to have an impact on chromatin structure. Many of these have been shown to associate with triplet repeat-containing loci. Further, noncoding RNA from diverse repetitive sequences has been found to direct the formation of heterochromatin, resulting in transcriptional gene silencing (12, 13). RNA sequences derived from several disease-causing triplet repeats can fold on themselves to form very stable duplex RNA (14, 15, 16). Therefore, it is intriguing that expanded triplet repeats might have the potential to manipulate chromatin structure through an RNA-targeted mechanism. This review aims to discuss the involvement of many of the factors described above in the pathogenesis of triplet expansion diseases.

3. EXPANDED CTG/CAG REPEATS AND MYOTONIC DYSTROPHY

Expansion of repeating tracts of CTG/CAG nucleotide triplets has been implicated in neuromuscular diseases such as Huntington's disease, myotonic dystrophy 1 (DM1), and spinocerebellar ataxia (SCA8), through both loss-of-function and gain-of-function mechanisms. In Huntington's diseases and at least eight other ataxias, the CTG/CAG expansions code for additional polyglutamines (CAG codes for glutamine) that are incorporated into a protein whose altered function elicits neurological dysfunction in cells (17). This gain-of-function effect is not limited to the protein level, and expanded CTG/CAG repeats located in the 3'-untranslated region in *SCA8* and *DM1* loci give rise to toxic CUG RNAs (18, 19). In addition, in DM1, a multi-system disease, the expanded repeat exerts its pathological effect also via loss-of-function mechanisms by shutting down the expression of *DMPK* (20-24), *SIX5* (25-27) and *DMWD* (28, 29) genes. Several studies have suggested that chromatin structure directed by the CTG/CAG repeats contributes to the gene repression at the *DM1* locus (24-27). Furthermore, the sizes of the expanded tracts are relatively small (36 to 127 copies) with polyglutamine diseases and SCA8, in contrast to the much larger expanded tracts observed in DM1. In the congenital forms of DM1, as many as several thousand copies of repeats have been observed. Such a large repeating sequence can create significant unusual chromatin structure and amplify any deleterious effect. The work summarized below shows that expanded CTG/CAG triplet repeats can greatly alter chromatin structure.

3.1. Nucleosome positioning by CTG/CAG repeats

The knowledge that the threshold for triplet diseases frequently approximates the amount of DNA in a nucleosome, and the general understanding that any short repeating sequence may amplify sequence signals for nucleosome positioning or exclusion, led Wang *et al.* (5) to examine the ability of long repeating tracts of CTG/CAG to assemble into chromatin. In this study, direct electron microscopy (EM) was utilized to measure the position of several hundred individual nucleosomes along single DNAs, and the data was summed to create a nucleosome positioning map. Numerous DNA structural transitions are facilitated by negative supercoiling, or influenced by heating in the presence of magnesium. Therefore, the effects of magnesium and supercoiling of the template DNA were considered in this assay. The resulting nucleosome map revealed that 48% of all nucleosomes were present in the region containing 130 CTG/CAG repeats derived from a DM1 patient. In contrast, the map for the plasmid vector lacking the CTG/CAG insert showed a uniform distribution of nucleosomes over the length of the DNA. The use of linear or supercoiled DNA that had not been treated with heat or magnesium still resulted in 29% of all nucleosomes localized to the triplet repeat. Further analysis using a set of six plasmids containing 26 to 250 contiguous CTG/CAG triplet repeats showed that as the size of the triplet tracts increased, the efficiency of nucleosome formation at the repeats also increased.

This data provided the first direct evidence that repeating CTG/CAG triplets can serve as a strong nucleosome positioning element. Indeed, this increase in strength of nucleosome formation occurs over the range where the triplet tracts change from normal, to the intermediate zone, to disease-causing. The observations therefore suggest that expanded tracts may create unusually stable nucleosomes that could alter local chromatin structure. However, microscopy was unable to provide a precise measurement of the strength of the CTG/CAG repeats as nucleosome positioning elements, a value which can be measured by competitive nucleosome reconstitution. Using this method, DNAs containing 75 and 130 CTG/CAG repeats were compared to the *Xenopus borealis* somatic *5S RNA* gene (6). The *5S* element from several species was previously the strongest known natural nucleosome positioning element. The results revealed that the DNA containing the $(CTG/CAG)_{75}$ repeat is 5.6 ± 0.4-fold higher in nucleosome assembly compared to the *5S RNA* gene, and for the $(CTG/CAG)_{130}$ DNA, the difference is 8.7 ± 0.7-fold stronger than the *5S RNA* gene. These assays allow the calculation of the free energy difference for nucleosome

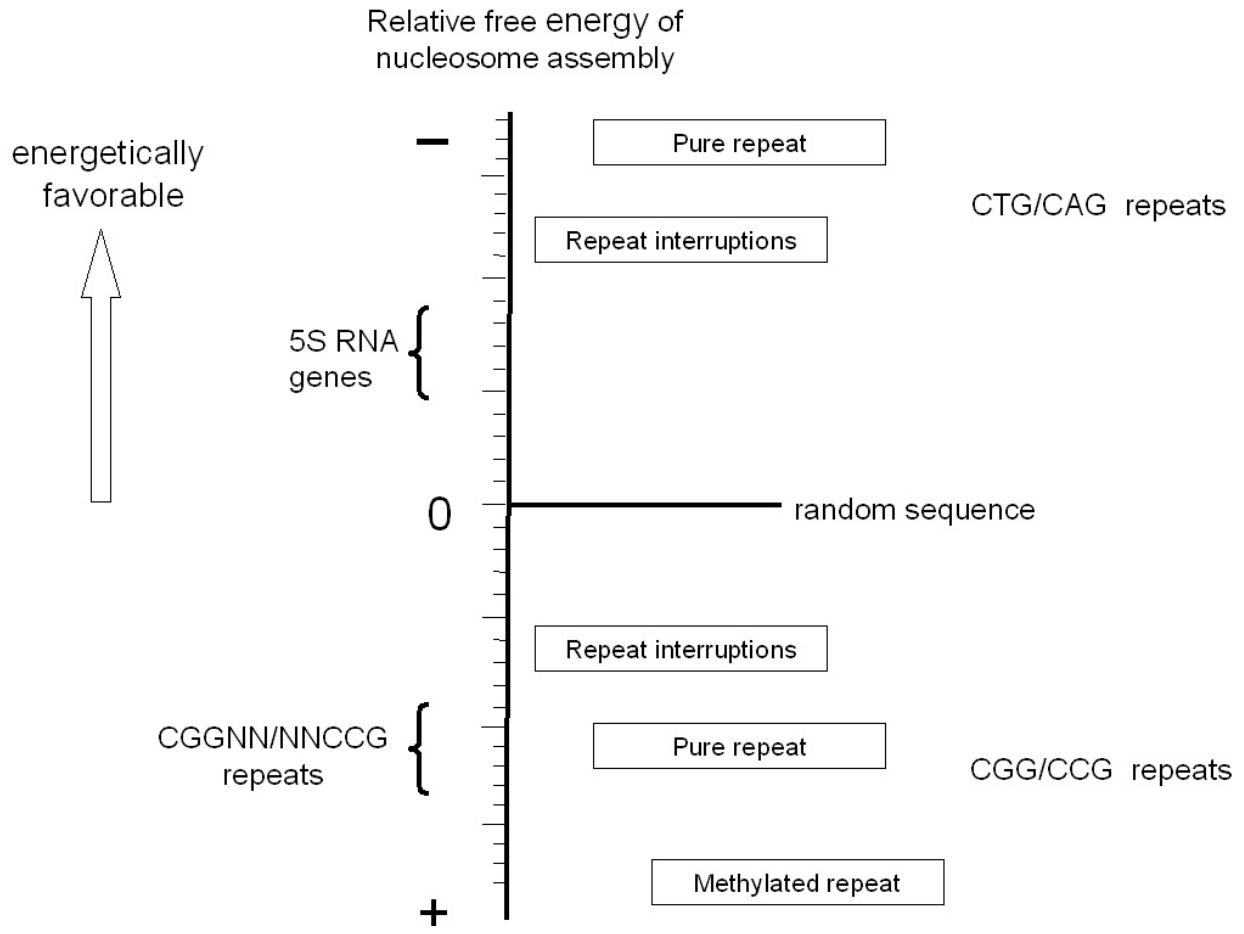

Figure 1. Comparison of relative free energies of nucleosome assembly from CTG/CAG and CGG/CCG repeating DNAs. The CTG/CAG repeats show energetically favorable binding of histones, while the CGG/CCG repeats exclude nucleosome formation. The repeat interruptions diminish the strongly enhanced- (CTG/CAG repeats) or excluded- (CGG/CCG repeats) nucleosome formations, thereby modulating the repeats to be more like random sequences in their ability to assemble into chromatin. The methylated CGG/CCG repeats further strengthen the nucleosome exclusion property.

formation. The difference in free energy between the $(CTG/CAG)_{130}$ DNA and the *5S RNA* gene is 1276±45 cal/mol (Figure 1).

These studies demonstrate that when CTG/CAG repeat tracts grow to sizes of n=75 or greater, they become the strongest known natural nucleosome positioning elements, and (at n=130) are nearly 10 times stronger than the *5S* element. The tendency in the ability of CTG/CAG repeat tracts to position nucleosomes is likely to increase in strength as the size of the repeat tract lengthens. Further, larger expansions could affect higher-order chromatin organization. Indeed, the strong correlation between the observation of the heterochromatin-like structure of the CTG/CAG sequence *in vivo* (see below) and the favorable energies of nucleosome formation by CTG/CAG repeats measured *in vitro*, strongly argues that the propensity of nucleosome assembly detected *in vitro* can provide insight into chromatin stability *in vivo*.

Godde and Wolffe (7) also examined nucleosome assembly on CTG/CAG repeats and confirmed the observation that CTG/CAG repeats create sites for preferential nucleosome assembly. Using nuclease and hydroxyl radical cleavage, they further mapped the translational and rotational positions of nucleosome formation over the CTG/CAG repeats. Probing of rotational position demonstrated a clear 10-11 bp repeating pattern over the DNA containing 6 copies of CTG/CAG repeat, and the pattern remains similar in $(CTG/CAG)_{55}$ DNA. Moreover, the $(CTG/CAG)_6$-containing DNA favors a single translational position with the CTG/CAG sequence bordering the nucleosomal dyad, whereas the $(CTG/CAG)_{55}$ DNA showed several nucleosome positions.

Also, twelve copies of CTG/CAG repeats have been introduced into a yeast minichromosome to examine the effect on chromatin structure and gene expression (30, 31). The (CTG/CAG)12 was shown to promote

nucleosome formation, but has no effect on gene expression in the assay. It is interesting that as short as 12 repeats of CTG/CAG has a measurable effect on direct nucleosome assembly, however, this length is probably not sufficient to have an impact on gene expression.

Although the precise physical reason why such repeating triplets generate such highly stable chromatin remains unclear, the appearance of such abnormally stable chromatin segments *in vivo* could exert serious biological impacts.

3.2. Chromatin structure of the *DM1* locus

The *in vitro* nucleosome reconstitution studies described above clearly show that repeating CTG/CAG triplets can generate very unusual chromatin structure. When the repeat tracts grow into the size range in which DM1 disease symptoms appear, these repeat tracts can possibly generate chromatin segments with stabilities that are ten times greater than the previously strongest known natural nucleosome positioning element, the *5S* DNA. Therefore, do expanded CTG/CAG repeats affect *in vivo* chromatin structure at the *DM1* locus? Does any unusual chromatin structure at the *DM1* locus contribute to the pathogenesis of DM1?

DM1 is an autosomal dominant disease, and the defect has been mapped to chromosome 19. Three genes at the *DM1* locus (*DMPK*, *SIX5* and *DMWD*) have been shown to be affected by the expanded CTG/CAG repeat, which is located in the 3' untranslated region of the *DMPK* gene, and is about 1 kb from the major transcription start site of the *SIX5* gene. A block of 5 to 37 CTG/CAGs is frequently found in normal individuals, while DM1 patients with mild symptoms have triplet blocks of 50-80 repeats, and expansions from 100 to several thousand repeats can be found in individuals with full-blown symptoms. DM1 patients are heterozygous with only one of the two alleles containing an expanded triplet block, since homozygotes are rare and lethal. Otten and Tapscott (25) have carried out *in vivo* mapping of nuclease hypersensitive sites with DNA samples of fibroblasts and myoblasts from three unrelated myotonic dystrophy patients, in which the *DM1* alleles are heterozygous and the expansion is ~6kb. Compared with the observations from an unaffected individual and the wild-type allele, the expanded allele shows loss of the DNaseI hypersensitive site located just 3' of the CTG/CAG repeat, and resistance to restriction endonuclease cleavage adjacent to the DNase I hypersensitive site. Further, this hypersensitive site has been mapped to an enhancer region which regulates the expression of the *SIX5* gene (26, 27). In cells of DM1 patients with a loss of the hypersensitive site, the amount of *SIX5* transcript was reduced compared to the controls, and the amount of transcript from the expanded allele was also greatly lowered compared to that from the wild- type allele.

The ability of the expanded CTG/CAG to repress adjacent gene expression by directing the formation of heterochromatin is further demonstrated outside the context of the *DM1* locus by using transgenic mice (32). The 192 copies of CTG/CAG repeat was introduced into a heterochromatin-sensitive human *CD2* transgene, and in all ten mouse lines, repression of the *CD2* gene was observed regardless of the location of the transgenes in chromosomes. This repression occurs at the level of transcription, and is correlated with a decrease in promoter accessibility. Over-expression of heterochromatin protein 1 (HP1) protein, a key component of heterchromatin structure, increases the percentage of repressed transgenes only in the lines carrying expanded CTG/CAG repeats, suggesting that induction of heterchromatin by the expanded CTG/CAG repeat contributes to the gene repression.

These results elegantly demonstrated that the expanded CTG/CAG repeat produces an altered chromatin structure, possibly resulting from the strong nucleosome positioning by the CTG/CAG repeat. The altered chromatin structure could suppress the transcription of the *SIX5* gene, which leads to some pathogenic features of DM1.

Another interesting question, from a chromatin organization point of view, is whether the heterochromatin structure induced by the CTG/CAG repeat influences higher-order organization at the *DM1* locus, and what other determinants are involved. The CTG/CAG repeat of the *DM1* locus has been shown to be located at the center of a 120 kb chromatin loop generated by two flanking MARs (11, Figure 2). The CTG/CAG repeat-containing loop encompasses the three DM1-related genes (*DMWD, DMPK* and *SIX5*). The arrangement is very intriguing because a chromatin loop domain is thought to act potentially as an independent regulatory unit, and cis elements such as enhancers or silencers can exert their effects over a long range on providing they are within the same domain. The regulatory function of each chromatin loop domain can be further fine-tuned by placing additional boundary elements such as insulators within the domain (33). Studies from Tapscott and colleagues have demonstrated that within the DM1 loop domain, the chromosome insulator protein CTCF binds to two CTCF-binding sites, which flank each side of the CTG/CAG repeat, and are about 176 bp apart (11). CTCF, an eleven-zinc finger DNA-binding protein, can bind to insulators of a conserved sequence (~50 bp) to block communication between enhancer and promoter (34). In the wild-type allele, the CTG/CAG repeats position a single nucleosome, and situate the CTCF-binding sites in the inter-nucleosomal region, allowing CTCF to bind, and preventing the spread of any altered chromatin structure. In contrast, in the large, expanded allele of congenital DM, the CTCF-binding sites are highly methylated, which inhibits the binding of CTCF (11), resulting in propagation of the heterochromatin structure. The heterochromatin structure is characterized by methylation of histone H3 at lysine 9 (H3K9) and enrichment of HP1gamma (35). In addition, the production of antisense transcripts from the enhancer region of the *SIX5* gene to the CTCF binding site has been demonstrated, and 21-nt RNA fragments have been detected and associated with H3K9 methylation and HP1gamma protein (35). These studies suggest, in addition to the ability of CTG/CAG repeat to position nucleosomes, that the RNA-directed heterochromatin could serve as a

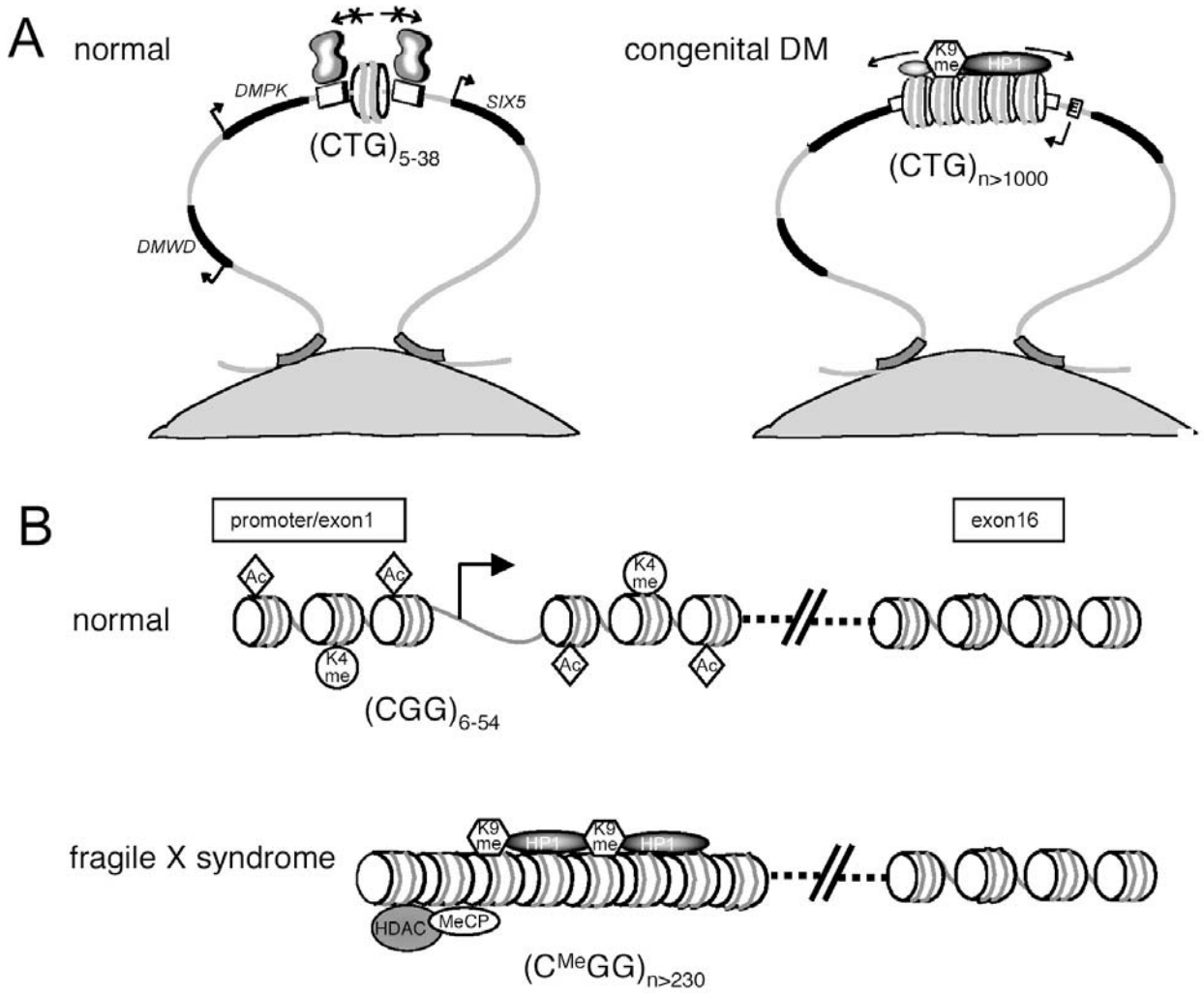

Figure 2. Chromatin structures at the *DM1* and fragile X loci. (A) The *DM1* chromatin domain is flanked and attached by MARs (grey curved rectangular shapes) to nuclear matrix (dome shape), and it encloses the *DMWD*, *DMPK* and *SIX5* genes (11). In wild type, the CTG/CAG repeat strongly positions one single nucleosome, but the binding of CTCF (dumbbell shapes) to the insulators (white rectangular boxes) prevents the spread of the heterochromatin structure. All three genes are actively transcribed (right angle arrows). As the repeat grows to thousands in congenital DM1, methylation of the insulator regions inhibits the binding of CTCF, allowing the propagation of the heterochromatin structure (11). The presence of antisense transcription (indicated by an angular arrow) from the enhancer region of the *SIX5* gene may contribute to maintainance and spread of the heterochromatin structure (35). (B) The promoter and exon 1 of the *FMR1* gene in normal individuals are associated with histones containing acethylated H3 and H4 (diamond shapes) and methylated H3K4 (round shapes), and the *FMR1* gene is actively transcribed (51-56). In fully mutated fragile X patients, CpG methylation of the promoter and the expanded CGG/CCG repeat, leads to changes in histone epigenetic modifications, including an increase of methylated H3K9 (hexagon shapes). Subsequently, possible binding of heterochromatin protein 1 (HP1), histone deacetylase (HDAC), and methyl-CpG binding protein (MeCP), may propagate the heterochromatin structure, resulting in transcriptional repression of the *FMR1* gene.

mechanism underlying gene regulation at the *DM1* locus. However, more studies are needed to address: 1) whether the Dicer pathway is used to generate the siRNA in this case; 2) whether the dsRNA transcripts, often involved in this process to produce siRNA, are produced from transcription of both strands, or more interestingly, from the folding of CUG or CAG transcripts.

4. EXPANDED CGG/CCG REPEATS AND FRAGILE X SYNDROME

The CGG/CCG expanded repeats have been found in all six folate-sensitive fragile sites sequenced to date. Fragile sites are chromosomal abnormalities in human linked to the incidence of certain cancers and other disorders. Cytologically, they are defined as sites of poor

staining, gaps, or DNA strand breakage, when chromosomes are treated under specific culture conditions. More than one hundred separate fragile sites have been identified in the human genome, classified as common or rare, and further divided according to the agents used to identify them (36). Fragile sites are highly conserved in chromosome evolution, and more importantly, most of them are inherited in families. Extensive studies have shown that the locations of many fragile sites are correlated to those of deletion and translocation breakpoints found in cancer cells (36). In addition, fragile sites have been mapped at the molecular level to the integration sites of oncogenic viruses and are major sites of drug-induced chromosomal amplification.

A number of fragile sites have been cloned and sequenced. It is of interest that sequence analysis of six rare, folate-sensitive sites: FRA10A, FRA11B, FRA16A, FRAXA, FRAXE, and FRAXF, all reveal long tracts of repeating CGG/CCG triplets together with frequent methylation of nearby CpG islands and the CGG/CCG repeats themselves. FRAXA has been studied extensively due to direct association with fragile X syndrome, one of the most common forms of inherited mental retardation. In FRAXA, the block of repeating CGG/CCG triplets maps in the 5' untranslated region of the *FMR1* gene. The instability of the CGG/CCG repeat is directly related to the severity of the disease, and as the repeat number increases, the symptoms of the disease intensify (37, 38). However, fragile X syndrome, unlike DM1, which may involve a complex interaction of several genes, appears to result from the inactivation of only the *FMR1* gene. This is illustrated by rare types of fragile X syndrome, shown to be due to point mutations or deletions in the *FMR1* gene as contrasted with the more common triplet expansion. Also, there is a strong correlation between the severity of the disease and the purity of the repeat tract (39, 40). Therefore, moderate sized pure CGG/CCG repeat tracts are more deleterious than longer repeat tracts containing interruptions. Finally, there is also a strong correlation between the severity of the disease symptoms and the level of DNA methylation at the CGG/CCG repeats and the promoter (41, 42). Unusual chromatin structure at FRAXA sites is considered the cause of transcriptional repression of the *FMR1* gene in fragile X syndrome. The CGG/CCG repeats, sequence interruptions, DNA methylation, and other epigenetic factors, have been shown to affect the formation of chromatin structure at the CGG/CCG repeat. Studies of *in vitro* nucleosome assembly on the CGG/CCG repeat, and characterization of chromatin structure at the *FMR1* locus are summarized below.

4.1. Nucleosome exclusion by CGG/CCG repeats, sequence interruptions, and CpG methylation

The CGG/CCG repeating sequence derived from all six folate-sensitive fragile sites sequenced to date, was shown to inhibit nucleosome formation as compared to non-CGG/CCG sequence, by reconstituting nucleosomes *in vitro* and mapping nucleosome positions by electron microscopy (43, Figure 1). Using DNA fragments containing various numbers of CGG/CCG repeats, competitive nucleosome reconstitution experiments have further demonstrated that as the size of the CGG/CCG repeat increases, the ability of the DNA to exclude nucleosomes increases, in proportion to the repeat block size (43, 44).

Although the reason for the strong nucleosome-positioning ability of the CTG/CAG repeats is unclear, the explanation for why repeating CGG/CGG DNA excludes nucleosomes has been explored. It is suggested that A/T sequences are preferred at the site of minor-groove compression, and G/C sequences in the major-groove compression region. Therefore, with a repeating sequence such as (A/T)3NN(G/C)3NN, by spacing these flexible hinges and alternating them every 5 bp along a DNA helix, the wrapping of DNA around the histone core is energetically favored (45, 46). In contrast, DNA containing long repeats in the form of $(G/C)_3NN(G/C)_3NN$ would exclude nucleosomes, because at each position where minor groove compression is required, the DNA presents the histone octamer with a G/C wedge that favors bending into the major groove. Model studies with a DNA containing a nucleosome-sized tract of repeating CGGNN pentanucleotides demonstrate that the pentanucleotide repeats resist nucleosome assembly in a manner comparable to the expanded CGG/CCG repeats (47). Both repeating sequences are members of the $(G/C)_3NN(G/C)_3NN$ motif family. This provides an explanation for nucleosome exclusion by repeating CGG/CCG triplets, and suggests that the CGG/CCG repeating sequence would only favor the major-groove compression, and therefore, its wrapping around a histone core would be energetically unfavorable.

The CGG/CCG repeats of *FMR-1* are normally interrupted by 1-3 AGG/CCT triplets in non-affected individuals (39, 40). Loss of the interruptions, resulting in a longer pure tract, correlates with disease symptoms and repeat instability (39, 40). How does AGG/CCT interruption influence the formation of nucleosomes on CGG/CCG repeats? Nucleosome assembly studies with pure and AGG/CCT-interrupted CGG/CCG repeating DNAs have shown a significant effect of interruptions on nucleosome formation by hyper-acetylated histones, but not by hypo-acetylated histones (8). Both pure and interrupted CGG/CCG DNAs that were examined showed nucleosome exclusion, and pure CGG/CCG DNAs showed further weakening of nucleosome assembly. Thus, AGG/CCT interruptions can modulate the ability of CGG/CCG repeats to assemble into chromatin, and this effect is dependent upon the acetylation status of the histones. These results suggest an association between sequence interruptions within the CGG/CCG repeats and *FMR1* gene expression. The observation that AGG/CCT-interrupted CGG/CCG tracts assemble hyper-acetylated nucleosomes more efficiently compared to pure repeats, may translate to increased transcription at the *FMR1* gene.

Interestingly, while AGG/CCT interruptions in CGG/CCG repeats increase nucleosome assembly relative to pure repeats, the ATG/CAT interruptions found in the CTG/CAG repeats decrease the strong nucleosome-

positioning ability of the pure repeats (8). Thus, AGG/CCT or ATG/CAT interruptions seem to modulate the repeats to be more like random sequences in their ability to assemble into chromatin (Figure 1). Such effects on chromatin structure may permit the interrupted repeats to be metabolized (with regard to transcription or mutation) more like random sequences rather than like genetically unstable sequences.

In fragile X syndrome, the degree of methylation of CpG dinucleotides in the regions of the expanded CGG/CCG repeats and the adjacent promoter is strongly correlated with the severity of the disease (41, 42), and the expression of the six CGG/CCG-containing fragile sites also depends on the degree of methylation. Therefore, the effect of DNA methylation on the ability of CGG/CCG repeats to assemble into chromatin was also investigated. Studies of CGG/CCG repeats with various degrees of methylation show that methylation of the CGG/CCG repeat enhances nucleosome exclusion (48). Furthermore, this nucleosome exclusion effect increases in proportion to the degree of methylation. DNA methylation has no measurable influence on nucleosome formation for the non-fragile DNA control.

Alterations in nucleosome assembly mediated by CGG/CCG expansions, AGG/CCT interruptions, and CpG methylation, could modulate chromatin accessibility. The observation that chromatin structure significantly affects *FMR1* gene expression is discussed below.

4.2. The *FMR1* locus: chromatin structure and transcription

Transcriptional repression of the *FMR1* gene is the major cause of fragile X syndrome. In normal individuals, the number of CGG/CCG repeats located in the 5' untranslated region of the *FMR1* gene ranges from 5 to 50. Patients with a larger expansion [(CGG/CCG)n, n>230] and highly methylated CpG islands, have no detectable amount of the *FMR1* transcript in lymphocytes, and show typical fragile X symptoms. Extensive studies have demonstrated that a compact chromatin structure contributes to the inactivation of the *FMR1* gene. Nuclease sensitivity studies in fibroblast and lymphoblast cells from fragile X males gave early support for a compact chromatin structure at the 5'end of the *FMR-1* gene (49, 50). In fragile X patients, the 5'end of the *FMR1* gene including the CGG/CCG repeating sequence, shows resistance to nuclease digestion similar to the same region on the inactive X chromosome of normal females. But in contrast, the loci of normal males and the active X chromosome of normal females display nuclease-hypersensitivity. This study indicates a compact chromatin conformation over the promoter and the CGG/CCG repeats of the *FMR1* gene, which severely limits the accessibility of nucleases (Figure 2).

The difference between compact chromatin structure observed in nuclei and *in vitro* nucleosome exclusion suggests that epigenetic marks (such as histone post-translational modifications) and/or trans-acting factors which recognize epigenetic marks, could play a role in the formation of unusual chromatin structure over the *FMR1* locus. Indeed, using chromatin immunoprecipitation to characterize the chromatin structure of the mutant expanded *FMR1* locus has shown the region lacking acetylated histones H3 and H4, decreasing methylation of histone H3 at lysine 4 (H3K4), and the presence of methylated H3K9 (51, 52). These post-translational epigenetic modifications are typical characteristics of a heterochromatin-like structure. In addition, treatments which reactivate the transcription of the mutant *FMR1* gene (51-54), also recruit acetylated histones back to the *FMR1* promoter region of the mutant locus (51, 52), indicating that the heterochromatin-like structure results in transcriptional repression of the expanded *FMR1* gene. Three regions of the *FMR1* gene: the promoter, exon 1 (including the CCG/CCG repeat) and exon 16, were examined for these epigenetic changes. The promoter and exon 1, but not exon 16, display the changes in histione modifications (as described above), when comparing wild-type and fragile X cells, and treatments to reactivate transcription were made (55, 56). These results suggest that alterations in chromatin structure affecting transcription of the *FMR1* gene, are exclusively placed at the promoter and the CGG/CCG-containing exon 1 regions, and that any sequence variations at these regions could influence the chromatin structure.

Indeed, DNA methylation of the expanded CGG/CCG repeat and the adjacent promoter region has been shown to significantly contribute to the chromatin structure and gene expression of the *FMR1* locus. De-methylation of DNA by 5-azadeoxycytidine can reactivate suppressed *FMR1* transcription in fragile X patient cells (51-56), and is coincident with re-association of acetylated histones and methylated H3K4 at the *FMR1* locus (52, 55, 56). A rare case of a normal individual carrying about 400 CGG/CCG repeats, but unmethylated at the *FMR1* locus, displays a normal level of *FMR1* transcription (57). This indicates that long CGG/CCG repeats alone cannot repress transcription of the *FMR1* gene, whereas DNA methylation is required for this repression. Characterization of the cell line derived from this rare case reveals that, although to a lesser degree compared to highly methylated full mutation, histone deacetylation and H3K9 methylation are present at the promoter and exon1 regions, despite active transcription (56). Therefore, the absence of DNA methylation and the presence of methylated H3K4 in this cell line are major epigenetic markers resulting in active transcription.

Interestingly, both the CGG/CCG repeats and the CGGNN/NNCCG repeating sequence which inhibit nucleosome formation *in vitro,* have been shown to exclude nucleosomes and enhance transcription of reporter genes in yeast (30, 31, 58), in which no CpG methylation is detected. Based on all of the information, it is speculated that long CGG/CCG triplet repeats inhibit chromatinization of DNA, which would increase the accessibility of these regions to various protein machinery. The transcription factors USF1, USF2, and alpha-Pal/Nrf-1 have been shown to interact with the *FMR1* promoter (59, 60). Binding of such proteins to the CGG/CCG repeat and the adjacent

promoter can influence directly and indirectly the transcriptional levels and/or transcription start sites of the *FMR1* gene. Recent studies show that the intermediate lengths of (CGG/CCG)n, where n = 40 to 230 and are unmethylated, are associated with premature ovarian failure (61) and the fragile X tremor ataxia syndrome (62), that are clinically distinct from FRAXA mental retardation. These "intermediate" lengths of CGG/CCG repeat display increased levels of transcription compared to shorter non-disease-associated lengths of 30 repeats (63-66). These expanded lengths are also associated with variant sites of transcription initiation (67).

In contrast, for full CGG/CCG expansions (>200 repeats) in fragile X syndrome patients, the long "unchromatinized" region in combination with the high density of CpG dinucleotides could attract the binding of DNA methyltransferase, which directs DNA methylation. Subsequently, the binding of CpG-methyl binding proteins, such as MeCP2, has been shown to associate with the histone deacetylase complex, which can prepare lysine residues of histones for further H3K9 methylation by H3K9 methyltransferase (68). Interestingly, methylation of H3K9 has been demonstrated to direct DNA methylation as well (69). This feedback process allows the propagation and maintenance of the heterochromatin-like structure. The net effect of these events would be the transcriptional repression of the *FMR1* gene.

5. SUMMARY AND PROSPECTIVE

Studies summarized here describe the ability of two trinucleotide repeating sequences, CTG/CAG and CGG/CCG, to form nucleosomes – the basic structural unit of chromatin. The expansion of CAG/CTG repeats is the cause of 12 human genetic diseases, while six rare chromosomal fragile sites, three of which are disease-associated, are caused by unstable CGG/CCG repeats. In particular, at DM1 and fragile X syndrome loci, respectively, the effects of expanded CTG/CAG and CGG/CCG repeats on chromatin packaging and regulation of transcription were examined.

The expanded CTG/CAG repeat is identified, by nucleosome assembly assay, as the strongest known natural nucleosome positioning element. This nucleosome positioning ability is translated inside the cells to form heterochromatin-like structure, characterized by the presence of HP1gamma protein and dimethylation at histone H3K9, and repression of the transcription of adjacent genes. At the *DM1* locus, two CTCF-bound insulators were found to flank either side of the CTG/CAG repeat, serving as an additional element to regulate the influence of heterochromatin structure. The detection of antisense transcription at this region raises the possibility that an RNA-directed mechanism could be involved in maintenance and propagation of the heterochromatin structure. It is of interest to discover the *DM1* chromatin domain flanked by MARs and enclosing three DM1-related genes. This provides an attractive model for how the expanded CTG/CAG repeat could exert its biological effect specifically on DM1-related genes. The clinical observation that the presentation of DM1 differs depending on the size of the CTG/CAG repeat may reflect a progressive involvement of additional genes and regulatory elements within the *DM1* chromatin domain, as the repeat increases in size.

The CGG/CCG repeat exhibits a low affinity to bind histone proteins *in vitro*, and in combination with epigenetic modifications of the DNA sequence and histones, it forms a nuclease-resistant heterochromatin-like structure at the *FMR1* locus. CpG methylation of the expanded repeat and the adjacent promoter region is the major determinant of the chromatin compaction and transcriptional repression of the *FMR1* gene, leading to fragile X syndrome. Also, this compact chromatin structure may play a role in the generation of fragile sites by causing failure of further chromatin condensation. Visualization of the FRAXA site on a fragile X chromosome by scanning electron microscopy has found that the fragile site appears as an isochromatid gap connecting two portions of the chromosome by individual chromatin fibers (70). The diameter of these fibers is about 25 nm, in agreement with the measurement of a single chromatin fiber. This result suggests that fragile sites are those at which chromatin fibers fail to fold into a higher-order organization.

In addition to repression of transcription, the appearance of a segment of heterochromatin-like structure may in itself promote further expansion of the triplet repeat. During the replication of eukaryotic DNA, the replication machinery must remove the DNA template from the histone octamers prior to, or possibly in concert with, the helicase action that separates the two DNA strands. Providing that the replication machinery would be severely hindered in this action by the presentation of an array of hyperstable nucleosomes, the likelihood of polymerase slippage as the replication fork transits the triplet repeat would be expected to be much greater.

Finally, studies of the effect of these simple repeats in forming chromatin, and the ability of such chromatinized templates to be replicated or transcribed have been and will contribute greatly to our knowledge of the role of chromatin structure in fundamental biological processes.

6. ACKNOWLEDGEMENTS

This work was supported by Public Health Service grants CA085826 and CA113836 from the National Cancer Institute. Graphic illustrations by Michael J. Yablonski are much appreciated.

7. REFERENCES

1. K. Luger: Structure and dynamic behavior of nucleosomes. *Curr. Opin. Genet. Dev.* 13, 127-135 (2003)
2. C. L. Woodcock and S. Dimitrov: Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev.* 11, 130-135 (2001)

3. J. R. Swedlow and T. Hirano: The making of the mitotic chromosome: modern insights into classical questions. *Mol. Cell.* 11, 557-569 (2003)
4. A. E. Ehrenhofer-Murray: Chromatin dynamics at DNA replication, transcription and repair. *Eur. J. Biochem.* 271, 2335-2349 (2004)
5. Y.-H. Wang, S. Amirhaeri, S. Kang, R. D. Wells and J. D. Griffith: Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* 265, 669-671 (1994)
6. Y.-H. Wang and J. D. Griffith: Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. *Genomics* 25, 570-573 (1995)
7. J. S. Godde and A. P. Wolffe: Nucleosome assembly on CTG triplet repeats. *J. Biol. Chem.* 271, 15222-15229 (1996)
8. D. J. Mulvihill, K. Nichol, K. A. Hagerman, C. E. Pearson and Y.-H. Wang: Effect of CAT or AGG Interruptions and CpG methylation on nucleosome assembly on triplet repeats of SCA1 and FRAXA. *J. Biol. Chem.* 280, 4498-4503 (2005)
9. I. K. Hornstra, D. L. Nelson, S. T. Warren and T. P. Yang: High resolution methylation analysis of the FMR1 gene trinucleotide repeat region in fragile X syndrome. *Hum. Mol. Genet.* 2, 1659–1665 (1993)
10. G. J. Brock, N. H. Anderson and D. G. Monckton: Cis-acting modifiers of expanded CAG/CTG triplet repeat expandability: associations with flanking GC content and proximity to CpG islands. *Hum Mol Genet.* 8, 1061-1067 (1999)
11. G. N. Filippova, C. P. Thienes, B. H. Penn, D. H. Cho, Y. J. Hu, J. M. Moore, T. R. Klesert, V. V. Lobanenkov and S. J. Tapscott: CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet.* 28, 335-343 (2001)
12. E. Bernstein and C. D. Allis: RNA meets chromatin. *Genes Dev.* 19, 1635-1655 (2005)
13. R. Almeida and R. C. Allshire: RNA silencing and genome regulation. *Trends Cell Biol.* 15, 251-258 (2005)
14. S. Michalowski, J. W. Miller, C. R. Urbinati, M. Paliouras, M. S. Swanson and J. D. Griffith: Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res.* 27, 3534-3542 (1999)
15. M. Broda, E. Kierzek, Z. Gdaniec, T. Kulinski and R. Kierzek: Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats- Implication for prediction of RNA structure. *Biochemistry* 44, 10873-10882 (2005)
16. B. H. Mooers, J. S. Logue and J. A. Berglund: The structural basis of myotonic dystrophy from the crystal structure of CUG repeats. *Proc Natl Acad Sci U S A.* 102, 16626-16631 (2005)
17. J. R. Gatchel and H. Y. Zoghbi: Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet.* 6, 743-755 (2005)
18. M. L. Moseley, T. Zu, Y. Ikeda, W. Gao, A. K. Mosemiller, R. S. Daughters, G. Chen, M. R. Weatherspoon, H. B. Clark, T. J. Ebner, J. W. Day and L. P. Ranum: Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet.* 38, 758-769 (2006)
19. R. N. Kanadia, K. A. Johnstone, A. Mankodi, C. Lungu, C. A. Thornton, D. Esson, A. M. Timmers, W. W. Hauswirth and M. S. Swanson: A muscleblind knockout model for myotonic dystrophy. *Science* 302, 1978-1980 (2003)
20. G. Novelli, M. Gennarelli, G. Zelano, A. Pizzuti, C. Fattorini, C. T. Caskey and B. Dallapiccola: Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. *Biochem Mol Biol Int.* 29, 291-297 (1993)
21. Y. H. Fu, D. L. Friedman, S. Richards, J. A. Pearlman, R. A. Gibbs, A. Pizzuti, T. Ashizawa, M. B. Perryman, G. Scarlato, R. G. Fenwick Jr. and C. T. Caskey: Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235-238 (1993)
22. P. Carango, J. E. Noble, H. G. Marks and V. L. Funanage: Absence of myotonic dystrophy protein kinase (DMPK) mRNA as a result of a triplet repeat expansion in myotonic dystrophy. *Genomics* 18, 340-348 (1993)
23. R. Krahe, T. Ashizawa, C. Abbruzzese, E. Roeder, P. Carango, M. Giacanelli, V. L. Funanage and M. J. Siciliano: Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. *Genomics* 28, 1-14 (1995)
24. R. Frisch, K. R. Singleton, P. A. Moses, I. L. Gonzalez, P. Carango, H. G. Marks and V. L. Funanage: Effect of triplet repeat expansion on chromatin structure and expression of DMPK and neighboring genes, SIX5 and DMWD, in myotonic dystrophy. *Mol Genet Metab.* 74, 281-291 (2001)
25. A. D. Otten and S. J. Tapscott: Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc Natl Acad Sci U S A.* 92, 5465-5469 (1995)
26. T. R. Klesert, A. D. Otten, T. D. Bird and S. J. Tapscott: Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat Genet.* 16, 402-406 (1997)
27. C. A. Thornton, J. P. Wymer, Z. Simmons, C. McClain and R. T. Moxley 3rd: Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat Genet.* 16, 407-409 (1997)
28. M. Eriksson, T. Ansved, L. Edstrom, M. Anvret and N. Carey: Simultaneous analysis of expression of the three myotonic dystrophy locus genes in adult skeletal muscle samples: the CTG expansion correlates inversely with DMPK and 59 expression levels, but not DMAHP levels. *Hum Mol Genet.* 8, 1053-1060 (1999)
29. M. Alwazzan, E. Newman, M. G. Hamshere and J. D. Brook: Myotonic dystrophy is associated with a reduced level of RNA from the DMWD allele adjacent to the expanded repeat. *Hum Mol Genet.* 8, 1491-1497 (1999)
30. M. Shimizu, R. Fujita, N. Tomita, H. Shindo and R. D. Wells: Chromatin structure of yeast minichromosomes containing triplet repeat sequences associated with human hereditary neurological diseases. *Nucleic Acids Res Suppl.* (1), 71-72 (2001)
31. N. Tomita, R. Fujita, D. Kurihara, H. Shindo, R. D. Wells and M. Shimizu: Effects of triplet repeat sequences on nucleosome positioning and gene expression in yeast

minichromosomes. *Nucleic Acids Res Suppl.* (2), 231-232 (2002)
32. A. Saveliev, C. Everett, T. Sharpe, Z. Webster and R. Festenstein: DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature* 422, 909-913 (2003)
33. A. C. Bell, A. G. West and G. Felsenfeld: Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* 291, 447-450 (2001)
34. A. C. Bell, A. G. West and G. Felsenfeld: The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387-396 (1999)
35. D. H. Cho, C. P. Thienes, S. E. Mahoney, E. Analau, G. N. Filippova and S. J. Tapscott: Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol Cell*. 20, 483-489 (2005)
36. R. I. Richards: Fragile and unstable chromosomes in cancers: cause and consequences. *Trends Genet.* 17, 339-345 (2001)
37. D. L. Nelson:The fragile X syndromes. *Seminars in Cell Biol.* 6, 5-11 (1995)
38. G. R. Sutherland and R. I. Richards: Simple tandem DNA repeats and human genetic disease. *Proc. Natl. Acad. Sci. USA* 92, 3636-3641 (1995)
39.C. B. Kunst and S. T. Warren: Cryptic and polar variation of the Fragile X repeat could result in predisposing normal alleles. *Cell* 77, 853-861 (1994)
40. C. B. Kunst, E. P. Leeflang, J. C. Iber, N. Arnheim and S. T. Warren: The effect of FMR1 CGG repeat interruptions on mutation frequency as measured by sperm typing. *J. Med. Genet.* 34, 627-631 (1997)
41. M. Pieretti, F. Zhang, Y.-H. Fu, S. T. Warren, B. A. Oostra, C. T. Caskey and D. L. Nelson: Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66, 817-822 (1991)
42. A. McConkle-Rosell, A. Lachiexwicz, G. Spiridigliozzi, J. Tarleton, S. Schoenwald, C. Phelan, P. Goonewardena, X. Ding and W. Brown: Evidence that methylation of the FMR-I locus is responsible for variable phenotypic expression of the fragile X syndrome. *Am . J. Genet*. 53, 800-809 (1993)
43. Y.-H. Wang, R. Gellibolian, M. Shimizu, R. D. Wells and J. D. Griffith: Long CCG triplet repeat blocks exclude nucleosomes: a possible mechanism for the nature of fragile sites in chromosomes. *J. Mol. Biol.* 263, 511-516 (1996)
44. J. S. Godde, S. U. Kass, M. C. Hirst and A. P. Wolffe: Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter. *J. Biol. Chem.* 271, 24325-24328 (1996)
45. T. E. Shrader and D. M. Crothers: Artificial nucleosome positioning sequences. *Proc. Natl. Acad. Sci. USA* 86, 7418-7422 (1989)
46. T. E. Shrader and D. M. Crothers: Effects of DNA sequence and histone-histone interactions on nucleosome placement. *J. Mol. Biol.* 216, 69-84 (1990)
47. Y.-H. Wang and J. D. Griffith: The $[(G/C)_3NN]_n$ motif: a common DNA repeat that excludes nucleosomes. *Proc. Natl. Acad. Sci. USA* 93, 8863-8867 (1996)
48. Y.-H. Wang and J. D. Griffith: Methylation of expanded CCG triplet repeat DNA from fragile X syndrome patients enhances nucleosome exclusion. *J. Biol. Chem.* 271, 22937-22940 (1996)
49. S. Luo, J. C. Robinson, A. L. Reiss and B. R. Migeon: DNA methylation of the fragile X locus in somatic and germ cells during fetal development: relevance to the fragile X syndrome and X inactivation. *Somat. Cell Mol. Genet.* 19, 393-404 (1993)
50. D. E. Eberhart and S. T. Warren: Nuclease sensitivity of permeabilized cells confirms altered chromatin formation at the fragile X locus. *Somat. Cell Mol. Genet.* 22, 435-441 (1996)
51. B. Coffee, F. Zhang, S. T. Warren and D. Reines: Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nat. Genet.* 22, 98-101 (1999)
52. B. Coffee, F. Zhang, S. Ceman, S. T. Warren and D. Reines: Histone modifications depict an aberrantly heterochromatinized *FMR1* gene in fragile x syndrome. *Am. J. Hum. Genet.* 71, 923-932 (2002)
53. P. Chiurazzi, M. G. Pomponi, R. Willemsen, B. A. Oostra and G. Neri: *In vitro* reaction of the FMR1 gene involved in fragile X syndrome. *Hum. Mol. Genet*. 7, 109-113 (1998)
54. P. Chiurazzi, M. G. Pomponi, R. Pietrobono, C. E. Bakker, G. Neri and B. A. Oostra: Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. *Hum. Mol. Genet*. 8, 2317-2323 (1999)
55. E. Tabolacci, R. Pietrobono, U. Moscato, B. Oostra, P. Chiurazzi and G. Neri: Differentail epigenetic modifications in the FMR1 gene of the fragile X syndrome after reactivating pharmacological treatments. *Euro. J. Hum Genet.* 13, 641-648 (2005)
56. R. Pietrobono, E. Tabolacci, F. Zalfa, I. Zito, A. Terracciano, U. Moscato, C. Bagni, B. Oostra, P. Chiurazzi and G. Neri: Molecular dissection of the events leading to inactivation of the FMR1 gene. *Hum. Mol. Genet.* 14, 267-277 (2005)
57. H. J. Smeets, A.P. Smits, C. E. Verheij, J. P. Theelen, R. Willemsen, I. Van de Burgt, A. T. Hoogeveen, J. C. Oosterwijk, and B. A. Oostra: Normal phenotype in two brothers with a full FMR1 mutation. *Hum Mol. Genet*. 4, 2103-2108 (1995)
58. D. T. Kirkpatrick, Y.-H. Wang, M. Dominska, J. D. Griffith and T. Petes: Control of meiotic recombination and gene expression in yeast by a simple repetitive DNA sequence that excludes nucleosomes. *Mol. Cell. Biol.* 19, 7661-7671 (1999)
59. D. Kumari and K. Usdin: Interaction of the transcription factors USF1, USF2, and alpha -Pal/Nrf-1 with the FMR1 promoter. Implications for Fragile X mental retardation syndrome. *J. Biol. Chem.* 276, 4357-4364 (2001)
60. K. T. Smith, B. Coffee and D. Reines: Occupancy and synergistic activation of the FMR1 promoter by Nrf-1 and Sp1 *in vivo*. *Hum. Mol. Genet.* 13, 1611-1621(2004)
61. A. Murray: Premature ovarian failure and the FMR1 gene. *Semin. Reprod. Med.* 18, 59-66 (2000)
62. P. J. Hagerman and R. J. Hagerman: The fragile-X premutation: a maturing perspective. *Am. J. Hum. Genet.* 74, 805-816 (2004)
63. F. Tassone, R. J. Hagerman, W. D. Chamberlain and P. J. Hagerman:Transcription of the FMR1 gene in

individuals with fragile X syndrome. *Am. J. Med. Genet.* 97, 195-203 (2000)
64. A. Kenneson, F. Zhang, C. H. Hagedorn and S. T. Warren: Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum. Mol. Genet.* 10, 1449-1454 (2001)
65. B. A. Oostra and R. Willemsen: A fragile balance: FMR1 expression levels. *Hum. Mol. Genet.* 12 Spec No 2:R249-R257 (2003)
66. E. G. Allen, W. He, M. Yadav-Shah and S. L. Sherman: A study of the distributional characteristics of FMR1 transcript levels in 238 individuals. *Hum. Genet.* 114, 439-447 (2004)
67. A. Beilina, F. Tassone, H. Schwartz, P. Sahota and P. J. Hagerman: Redistribution of transcription start sites within the FMR1 promoter region with expansion of the downstream CGG-repeat element. *Hum. Mol. Genet.* 13, 543-549 (2004)
68. M. Lachner and T. Jenuwein: The many faces of histone lysine methylation. *Curr. Opin. Cell Biol.* 14, 286-298 (2002)
69. R. J. Sims 3rd, K. Nishioka and D. Reinberg: Histone lysine methylation: a signature for chromatin function. *Trends Genet.* 19, 629-639 (2003)
70. C. J. Harrison, E. M. Jack, T. D. Allen and R. Harris: The fragile X: a scanning electron microscope study. *J. Med. Genet.* 20, 280-285 (1983)

Abbreviations: MARs: matrix attachment regions, SARs: scaffold attachment regions DM1: myotonic dystrophy 1, SCA8: spinocerebellar ataxia, EM: electron microscopy, HP1: heterochromatin protein 1, H3K4: histone H3 at lysine 4, H3K9: histone H3 at lysine 9

Key Words: Chromatin structure, Nucleosome assembly, Triplet expansion disease, Myotonic dystrophy, Fragile X syndrome, Review

Send Correspondence to: Dr Yuh-Hwa Wang, Department of Biochemistry, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157-1016, USA, Tel: 336-716-6186, Fax: 336-716-7671, E-mail: ywang@wfubmc.edu