Am. J. Hum. Genet. 64:346–353, 1999

HUMAN GENETICS '99: TRINUCLEOTIDE REPEATS
Biological Implications of the DNA Structures Associated with Disease-Causing Triplet Repeats

Richard R. Sinden

Center for Genome Research, Institute of Biosciences and Technology, Texas A&M University, Houston

Within the last 6–8 years, the remarkable revelation that many neurodegenerative diseases are caused by the expansion of triplet repeats has prompted a great deal of research regarding the structure and biology associated with triplet-repeat sequences in DNA (Wells and Warren 1998). To date, >12 human genetic diseases (table 1), including myotonic dystrophy, fragile X syndrome, Huntington disease, several spinocerebellar ataxias, and Friedreich ataxia, have been associated with the expansion of CTG, CGG, or GAA repeats (Pearson and Sinden 1998*a*). (For simplicity, references to CTG, CGG, and GAA repeats indicate duplex-repeat sequences—that is, $(CTG)_n$• $(CAG)_n$, $(CGG)_n$• $(CCG)_n$, or $(GAA)_n$• $(TTC)_n$. Single-strand tracts will be referred to as $(CTG)_n$, $(CGG)_n$, or $(GAA)_n$ tracts.) The expansion of triplet repeats represents a novel type of mutagenic event that has not been identified, to date, in genetically tractable organisms such as bacteriophage, *Escherichia coli,* yeast, or *Drosophila.*

Survival and adaptation of a species requires a balance between the repair of errors inherent in the replication, repair, and recombination of DNA and the introduction of occasional mutations as fodder for natural selection. This balance is complex because not all sequences of DNA undergo spontaneous mutagenesis at the same rate. Spontaneous mutation hot spots are frequently associated with repeated DNA sequences, including inverted repeats, mirror repeats, runs of a single nucleotide, and di-, tri-, tetra-, or higher-order–nucleotide repeats. Therefore, a high frequency of spontaneous mutagenesis associated with tracts of triplet repeats is not a completely surprising event. However, the massive expansion of triplet repeats associated with fragile X syndrome, myotonic dystrophy, and Friedreich ataxia was completely unexpected.

Genomic DNA should not be considered merely an informational repository in which structural RNAs and proteins are encoded. Particular sequence arrangements of DNA are also responsible for the shape of the DNA helix, including bending of DNA, and variations in the three-dimensional structure of the helix (Sinden 1994; Bianchi and Beltrame 1998). This sequence heterogeneity causes structural variations in DNA that are responsible for the specificity of interaction with different proteins. The particular sequence arrangements in DNA can also lead to specific mutational events.

This review will briefly summarize work from the last several years directed toward understanding the particular structural properties associated with CTG, CGG, or GAA repeats; the interaction of proteins with triplet-repeat DNA sequences; and several models to account for the expansion and instability of triplet-repeat sequences.

Structural Properties Associated with Duplex Triplet-Repeat Tracts

Linear DNA molecules containing triplet repeats migrate anomalously during polyacrylamide electrophoresis. This unexpected property immediately indicated that triplet-repeat–containing DNA sequences adopt unusual structures with novel biophysical properties. One aspect of these properties involves the inherent flexibility of both CTG- and CGG-repeat tracts. Cyclization kinetics measurement has shown that CTG and CGG repeats possess greater flexibility than the mixed-sequence DNA (Bacolla et al. 1997). Chastain and Sinden (1998) demonstrated that even very short tracts of triplet repeats can impart a high degree of torsional flexibility in small DNA molecules by showing that CTG repeats were just as effective in removing DNA bending resulting from phased polyadenosine tracts as were two consecutive T•T mismatches.

Received December 10, 1998; accepted for publication December 19, 1998; electronically published February 1, 1999.

Address for correspondence and reprints: Dr. Richard R. Sinden, Center for Genome Research, Institute of Biosciences and Technology, Texas A&M University, 2121 West Holcombe Boulevard, Houston, TX 77030-3303. E-mail: rsinden@ibtmail.tamu.edu

This article represents the opinion of the author and has not been peer reviewed.

0002-9297/99/6402-0004$02.00

Table 1

Trinucleotide Repeats in Human Genetic Disease

Disease	Repeat[a]	Normal Length[b]	Intermediate Length (premutation)[b,c]	Full Disease Length[b]	Possible Biological Effect of Expansion
Fragile XA (FRAXA)	$(CGG)_n$	6–52	59–230	230–2,000 (pure)	Repeat tract 5′ to coding region; promoter methylation, gene silencing
Fragile XE (FRAXE)	$(CCG)_n$	4–39	? (31–61)	200–900	Promoter methylation, gene silencing
Fragile XF (FRAXF)	$(CGG)_n$	7–40	?	306–1,008	Promoter methylation, gene silencing
FRA16A	$(CCG)_n$	16–49	?	1,000–1,900	Promoter methylation, gene silencing
Jacobsen syndrome (FRA11B)	$(CGG)_n$	11	80	100–1,000	Promoter methylation, gene silencing
Kennedy syndrome (SMBA)	$(CAG)_n$	14–32	?	40–55	Polyglutamine tract expansion
Myotonic dystrophy (DM)	$(CTG)_n$	5–37	50–80	80–1,000; congenital, 2,000–3,000	Repeat tract 3′ to coding region; altered mRNA processing and transport, altered gene expression due to chromatin changes
Huntington disease (HD)	$(CAG)_n$	10–34	36–39	40–121	Polyglutamine expansion
Spinocerebellar ataxia 1 (SCA1)	$(CAG)_n$	6–39	…	40–81 (pure)	Polyglutamine expansion
Spinocerebellar ataxia 2 (SCA2)	$(CAG)_n$	14–31	…	34–59 (pure)	Polyglutamine expansion
Spinocerebellar ataxia 3 (SCA3)/Machado Joseph disease (MJD)	$(CAG)_n$	13–44	?	60–84	Polyglutamine expansion
Spinocerebellar ataxia 6 (SCA6)	$(CAG)_n$	4–18	?	21–28	Polyglutamine expansion
Spinocerebellar ataxia 7 (SCA7)	$(CAG)_n$	7–17	?	38–130	Polyglutamine expansion
Haw River syndrome (HRS; also DRPLA)	$(CAG)_n$	7–25	?	49–75	Polyglutamine expansion
Friedreich ataxia (FRDA)	$(GAA)_n$	6–29	? (>34–40)	200–900	Repeat tract in an intron; altered mRNA production, altered replication

[a] Typically, repeat tracts contain sequence interruptions. See Pearson and Sinden (1998*a*) for a discussion of the sequence interruptions.

[b] No. of triplet repeats.

[c] A question mark (?) indicates potential mutagenic intermediate length, and an ellipsis (…) indicates none. Not all diseases are associated with a premutation length repeat tract or premutation disease condition.

Protein Interactions with Duplex Triplet Repeats

Because the duplex helix structure associated with triplet-repeat DNA sequences is so anomalous compared with the mixed-sequence DNA found in most coding regions, the interaction of proteins with certain triplet-repeat tracts may be quite distinctive. Interactions of nucleosomes with CTG and CGG repeats, for example, are both unusual but lead to opposite effects. CTG repeats become organized into nucleosomes more easily than mixed-sequence DNA, whereas CGG repeats are only poorly organized into nucleosomes (Wang and Griffith 1995; Wang et al. 1996). The ability of CGG triplet repeats to organize into nucleosomes is influenced by the state of methylation of CpG dinucleotides and the length of the repeat tract (Godde et al. 1996; Wang and Griffith 1996). This differential behavior is surprising, given the fact that both classes of repeats exhibit similar degrees of increased flexibility. If the organization into nucleosomes were purely a function of flexibility, then both repeats might be expected to show increased association with nucleosomes. Possibly the radius of curvature required to assemble DNA into a nucleosome is not favorably adopted by the CGG repeats. The differential organization into nucleosomes may have important biological consequences, because the organization of DNA into nucleosomes can be important for the regulation of gene expression. Hence, two DNA sequences with nearly identical biophysical properties may lead to dramatically different biological effects.

Long tracts of triplet repeats have the potential to influence replication and transcription. As described in more detail below, CTG-repeat sequences can block replication forks in *E. coli* (Samadashwily et al. 1997) and promote strand switching during replication in vitro (Ohshima and Wells 1997). However, once expansion to thousands of repeats has occurred in individuals with fragile X or myotonic dystrophy, the sequences appear to be replicated with a relatively high degree of fidelity and can even continue to expand in length (Wong et al. 1995). This difference between replication in human cells and replication in vitro or in bacteria must reflect a difference in the chromatin organization and/or replication machinery within these different systems.

Although DNA polymerase traverses triplet-repeat tracts in vitro with difficulty, the same is not true of human RNA polymerase II. RNA polymerase II efficiently transcribes long $(CTG)_n$, $(CAG)_n$, $(CGG)_n$, or $(CCG)_n$ tracts, albeit at altered rates relative to mixed-sequence DNA (Parsons et al. 1998). This result is consistent with the observation that mRNA is expressed from the expanded alleles of the myotonic dystrophy gene but that expanded CUG transcripts interact abnormally with specific RNA-binding proteins, so they may not be processed properly (Krahe et al. 1995; Timchenko 1999 [in this issue]).

Alternative DNA Structures Formed in Triplet Repeats

There are several alternative DNA conformations that can form within triplet-repeat DNA sequences (Mitas 1997; Pearson and Sinden 1998*b*). Several different hairpin structures and quadruplex structures can form within single-strand tracts of CXG triplet repeats (where X represents any nucleotide). Theoretically, slipped-strand DNA structures can form within duplex DNA in any triplet-repeat sequence. Intramolecular triplex DNA structures form within $(GAA)_n$• $(TTC)_n$ repeats.

Hairpins

Single-strand $(CTG)_n$, $(CAG)_n$, $(CGG)_n$, and $(CCG)_n$ repeat tracts can form intramolecular hairpin structures stabilized by base pairs formed within every two of the three bases of the triplet repeat (fig. 1*A*) (Mitas 1997; Darlow and Leach 1998; Pearson and Sinden 1998*b*). Both $(CGG)_n$ and $(CCG)_n$ strands can each fold into a hairpin structure in two ways, involving either a G•G or C•C mismatch. In contrast, $(CTG)_n$ and $(CAG)_n$ repeats can each form only one type of hairpin structure with a T•T or an A•A mismatch. As might be expected, the stability of these different hairpin structures varies considerably. For example, hairpins formed within $(CTG)_n$ repeats are more stable than hairpins formed within $(CAG)_n$ repeats. This effect may help explain the differential stability (discussed later) of CTG repeats that are cloned in opposite orientations into different plasmids.

Because hairpins formed from (CXG) repeats contain an X•X mismatch every 2 bp, DNA mismatch-repair proteins can recognize features of these mismatch hairpins. In *E. coli,* the methyl-directed mismatch-repair system recognizes mismatches and deletion/insertion loops of <3–4 nucleotides. The human mismatch-repair system functions similarly but involves a wide variety of proteins that may be specialized for different mismatches, and it can recognize mismatches and large loops as long as 5–16 nucleotides. Purified hMSH2 protein binds differentially to heteroduplexes containing $(CTG)_n$ and $(CAG)_n$ loops in slipped-strand DNA and in "slipped intermediate" structures (duplex DNA with only $(CTG)_n$ or $(CAG)_n$ hairpins) (Pearson et al. 1997). hMSL2 binds preferentially to the $(CAG)_n$ loop (or hairpin) relative to a similar-sized $(CTG)_n$ hairpin. This differential binding may be reflected in differential repair efficiencies of DNA intermediates in a process of mutagenesis.

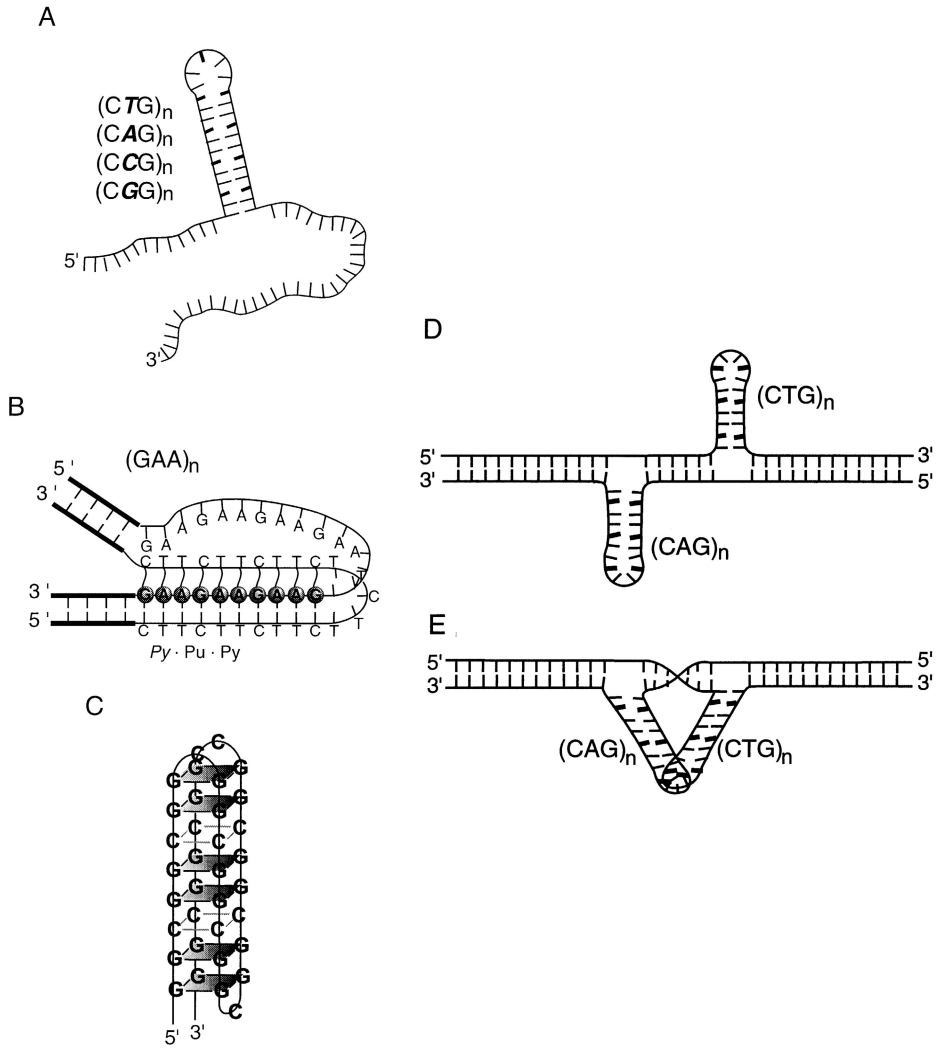

Figure 1 DNA structures associated with triplet repeats. *A,* Hairpin structures. Hairpin structures can form from $(CTG)_n$, $(CAG)_n$, $(CCG)_n$, and $(CGG)_n$ repeats. In all cases, 2 CG bp are followed by either a T•T, A•A, C•C, or G•G mispair. These hairpins have various degrees of thermal stability. *B,* Triplex DNA structures. Intramolecular triplex structures can form in $(GAA)_n$• $(TTC)_n$ repeat tracts. In the structure shown, the 3′ end of the pyrimidine-rich strand within the GAA tract unpairs and forms Hoogsteen base pairs with the purine bases of the Watson-Crick helix, and the 5′ end of the GAA tract remains single stranded. *C,* CGG quadruplex. Single-stranded DNA containing runs of G residues can form quadruplex DNA structures in which four Gs are held together by Hoogsteen base pairs. In the case of a CGG-triplet repeat, the single-strand DNA can fold into a quadruplex-type structure containing two G quartets and a quartet composed of four C residues that are weakly hydrogen bonded. *D,* Slipped-strand structure. Slipped-strand DNA structures can form after denaturation of the repeated DNA sequences, if the repeat tract is renatured out of register. In the case of a (CTG)•(CAG) repeat, the slipped-strand structure would involve the formation of a $(CTG)_n$, as well as a $(CAG)_n$ hairpin in opposite strands. *E,* Folded slipped structure. A slipped-strand DNA structure formed within a triplet repeat may organize into a folded-type structure in which the single-stranded DNA within the loop of the hairpins can engage in the Watson-Crick–type hydrogen bonding.

Triplex DNA

Intramolecular triplex DNA forms in polypurine • polypyrimidine sequences with mirror-repeat symmetry, a condition met by GAA repeats (fig. 1*B*) (Bidichandani et al. 1998). The formation of triplex DNA within GAA repeats is favored by conditions of low pH and unrestrained negative DNA supercoiling—conditions that can exist in living cells. DNA sequences that can form intramolecular triplex structures (Sinden 1994) can act

as effective blocks to DNA replication in vitro and in cells (Baran et al. 1987; Rao et al. 1988; Krasilnikov et al. 1997). Presumably, replication pausing is caused by dissociation of polymerase from the nascent strand and formation of intramolecular triplex structures. Non–B-DNA structures may form in long tracts of GAA repeats and cause reduced transcription and replication, as has been observed in vitro and in kidney cells from African green monkeys (COS-7) (Bidichandani et al. 1998; Ohshima et al. 1998). Thus, it is possible that long tracts of GAA repeats in the frataxin gene may participate in the formation of triplex DNA, an RNA:DNA triplex, or a triplex containing an RNA:DNA hybrid.

Quadruplex DNA

DNA sequences containing tracts of guanine nucleotides can form quadruplex DNA structures held together by Hoogsteen base pairs between guanine bases (fig. 1*C*) (Sinden 1994). CGG as well as AGG and TGG repeats (and pure poly G tracts) can form a variety of quadruplex DNA structures (Usdin 1998). The formation of these structures would require long tracts of single-strand repeats, as these quadruplex structures may not form readily from duplex DNA structures. The base pairings involved in the quadruplex structure preclude formation via a Watson-Crick base-paired hairpin intermediate. Like intramolecular triplex structures, quadruplex structures may act as efficient blocks to DNA replication or transcription.

Slipped-Strand DNA

Slipped-strand DNA structures can form within repeated DNA sequences (Sinden 1994; Pearson and Sinden 1998*a*). Slipped-strand DNA forms from an out-of-register alignment of complementary duplex strands within the repeat region (fig. 1*D*). Such out-of-register misalignment will result in the formation of loops in the complementary strands. After denaturation and renaturation of DNA molecules containing CTG or CGG repeats, a high proportion of the DNA population adopts an alternative conformation that can be deduced from the retarded mobility of these molecules in polyacrylamide gels. The amount of slipped-strand DNA structure formed is proportional to the repeat tract length and heterogeneity. Retarded mobility is consistent with the formation of branches, loops, or hairpin arms within the duplex DNA. Biochemical evidence and analysis by means of electron microscopy are consistent with the existence of slipped-strand DNA structures formed within the triplet repeats in otherwise duplex DNA molecules (Pearson and Sinden 1996; Pearson et al. 1998). Because the looped-out strands in this structure are complementary, it is very possible that folded slipped structures could form (fig. 1*E*). Slipped-strand DNA structures represent mutational intermediates in the expansion or deletion of triplet-repeat DNA sequences.

Genetic Instability of Triplet Repeats

A remarkable feature associated with triplet-repeat diseases is the unusual massive expansion associated with fragile X syndrome, myotonic dystrophy, and Friedreich ataxia. This type of mutagenic event had not been observed before the discovery of the molecular basis for these human genetic diseases. Unfortunately, this massive expansion has not been duplicated in simple model systems, and it remains a unique feature of human genetics. Nevertheless, studies in model organisms reveal some significant properties of triplet repeats.

In general, triplet repeats, especially long ones, are rather unstable in *E. coli* and yeast. Moreover, repeat-tract orientation can alter stability. Thus, CTG tracts are more prone to contraction (or deletion) when the $(CTG)_n$ tract is found on the lagging, rather than the leading, template strand. Conversely, expansions are favored when the $(CTG)_n$ tract occurs on the leading template strand (Kang et al. 1995). Deletions can result from hairpin formation, both because hairpins stabilize misaligned intermediates preferentially in the lagging strand (Trinh and Sinden 1991) and because they promote misalignment (Sinden et al. 1991). DNA polymerases pause at the base of hairpin structures and then slowly replicate through the hairpin, and this pausing can promote ectopic primer-template misalignment.

The known features of triplet-repeat stability in *E. coli* for CTG repeats can be explained by the differential stability of $(CTG)_n$ or $(CAG)_n$ hairpins and the different biochemical properties associated with replication of the leading and lagging strands (fig. 2). First, the probability of hairpin formation is greater in the lagging strand where a single-strand DNA exists during replication. The length of this single-strand region is always equal to that of an Okazaki fragment. If single-strand binding proteins are not bound to DNA, the single-strand triplet-repeat tract can fold into a hairpin structure. Since $(CTG)_n$ forms a more stable hairpin than does $(CAG)_n$, more hairpins are expected to form when the lagging template strand comprises $(CTG)_n$ (fig. 2*B*) After a hairpin forms, replication slippage involving primer template misalignment would lead to variable-sized deletions (fig. 2*B*). When $(CAG)_n$ is in the lagging strand template, there is a lower probability that a hairpin will form in the lagging strand and lead to deletions, although a $(CTG)_n$ hairpin could still form in leading template strand and cause a deletion (fig. 2*C, left*). Alternatively, a hairpin can form in the lagging nascent strand, resulting in a duplication (expansion) of the sequences in the hairpin (fig. 2*C, right*).

Deletions resulting from hairpin formation would nec-

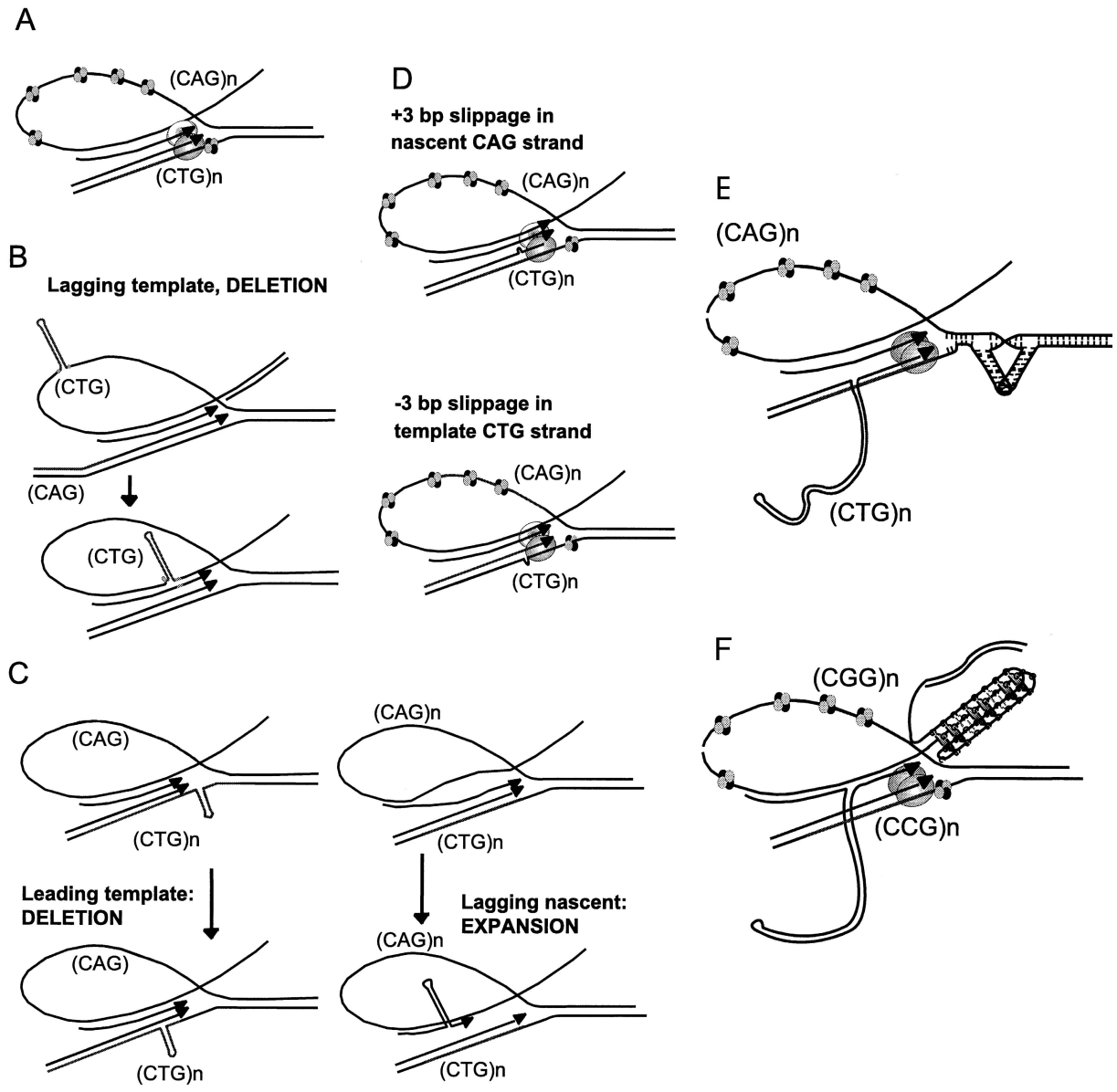

Figure 2 Models for genetic instability of triplet repeats. *A,* Error-free replication. The replication fork is shown in the "concurrent model" form, in which a dimeric complex of two holoenzyme polymerase complexes replicates both the leading and the lagging strands simultaneously. In this case, no misincorporations or primer template misalignment events leading to deletions or insertions of triplet repeats occur. Single-strand binding protein (SSB) is shown as tetrameric balls. *B,* $(CTG)_n$ hairpin formation in the lagging strand. There is a higher probability that $(CTG)_n$ can form a hairpin when the lagging template strand, rather than the leading template strand, is composed of $(CTG)_n$. Such a hairpin can form in the absence of binding of SSB on the lagging template strand and promote primer template misalignment. The CTG tract extruded into the hairpin is then deleted as DNA polymerase (larger shaded balls) bypasses the structure during replication. *C,* $(CTG)_n$ hairpin formation in the leading strand. As in the case of the lagging strand, primer template misalignment across a $(CTG)_n$ hairpin would lead to deletion during leading strand replication (*left side*). The probability of $(CAG)_n$ hairpin formation is much lower than that for $(CTG)_n$ hairpin formation, because of the difference in their thermal stabilities. However, the primer strand may unwind and form a $(CTG)_n$ hairpin in the lagging nascent strand, which would lead to the expansion or lengthening of a CTG tract (*right side*). *D*, 3-bp replication slippage. Slippage of the replication machinery by exactly 3 bp would cause a much smaller primer template misalignment than in the above models. When $(CTG)_n$ is found in the leading template strand, a +3 bp loop-out of CAG in the nascent strand would lead to an addition of 3 bp, if uncorrected (*top*), and a +3 bp loop-out in the CTG template strand would cause a 3-bp deletion (*bottom*). *E,* Reiterative synthesis at a folded slipped structure in the leading-strand template. In certain cases, DNA polymerase is known to idle on a short template, generating very long nascent DNA strands. A block to DNA polymerase movement may lead to long tracts of triplet repeats that become stabilized by hairpin (or other structure) formation. Such blocks to DNA replication could include slipped-strand DNA structures or a folded slipped-strand structure. *F,* Reiterative synthesis at a quadruplex in the lagging strand template. A quadruplex DNA structure could potentially form within the single-strand DNA in the lagging template strand. This structure may form a significant block to DNA replication, promoting reiterative synthesis during replication of the lagging strand.

essarily involve a minimum length required to form a hairpin (~5–10 repeats). However, small deletions and insertions of a single triplet repeat can also arise by primer template misalignment involving 3 bp (fig. 2*D*). In fact, recent genetic evidence has shown that such short slippage events occur frequently in *E. coli,* and they exhibit strand asymmetry (V. I. Hashem and R. R. Sinden, unpublished results). Primer template slippage events involving 3 bp are likely recognized by the *E. coli* mismatch-repair system, whereas slippage involving ⩾6 bp would not be recognized. Recent evidence has shown increased heterogeneity within triplet-repeat tracts in mismatch-repair–defective *E. coli* (Schumacher et al. 1998), consistent with the existence of short slippage events involving a single triplet repeat during DNA replication.

Triplet-Repeat Expansions Leading to Human Disease

Numerous models have been proposed to explain the expansion of triplet repeats leading to human disease (Pearson and Sinden 1998*a*). Expansions (and deletions) can occur through primer template misalignment during replication of a triplet-repeat tract (as discussed above and in fig. 2). Primer template misalignment (and slippage) at the 3′ end of an Okazaki fragment could also lead to the expansion of triplet repeats. Simple recombination or gene-conversion–type mechanisms might lead to doubling or possibly tripling of the repeat length, but these mechanisms cannot easily explain expansion events of the magnitude observed in fragile X syndrome or myotonic dystrophy. However, such mechanisms could be involved in shorter expansions observed in many other triplet-repeat diseases.

Repetitive misalignment events or reiterative DNA synthesis could lead to expansions longer than the length of the original repeat tract. Sinden and Wells (1992) suggested that the massive expansion could result from reiterative DNA synthesis occurring when DNA polymerase encounters a physical block to replication. This model alone can easily explain expansion from hundreds to thousands of triplet repeats in a single step (fig. 2*E* and *F*).

A general feature of triplet-repeat disorders is that instability and disease are often associated with the formation of a pure tract of triplet repeats from one that contains sequence interruptions. Occasionally, normal individuals are identified with lengths of triplet repeats beyond those associated with the disease threshold. In these cases, the long tracts of triplet repeats are usually interrupted with a nonrepeat trinucleotide. Conversely, relatively short repeat tracts may be associated with disease when interruptions in the sequence are lost. Therefore, the development of the disease state, as well as the propensity for subsequent expansion, is associated with the length of a pure repeat tract (Pearson and Sinden 1998*a*).

The probability of forming slip-stranded DNA structures is proportional to the length of the triplet-repeat tract. Moreover, sequence interruptions within pure repeat tracts dramatically decrease the propensity to form slipped-strand DNA structures as well as the heterogeneity of structures formed. Thus, the protective effect of the sequence interruptions in the formation of alternative DNA structures parallels that for the acquisition of the disease state. There are several possible explanations for this correlation. First, after primer template slippage, a mismatch will destabilize the duplex helix. Thus, the total number of possible stable structures formed will be much lower for a triplet-repeat tract containing sequence interruptions. Similarly, hairpin structures containing a sequence interruption are relatively unstable compared with those formed within pure repeat tracts. Finally, hairpins containing additional mismatches (or regions of slipped mispairing containing such mismatches) will be more prone to repair by mismatch-repair systems, thus reducing the propensity for expansions or deletions. Thus, if the formation of an alternative DNA structure is involved in the instability associated with repeat tracts, sequence interruptions certainly have a protective effect on genetic instability.

Conclusion

The discovery of expanding repeat tracts associated with human disease has revealed an unexpected characteristic of a simple DNA sequence in the context of the human genome. This instability and the propensity for massive expansion likely reflect a new mutational mechanism (or combination of processes) that is not a feature of simpler model genetic organisms. The molecular mechanisms responsible for the genetic instability observed in triplet-repeat–associated disorders likely involve the unique structural properties associated with simple triplet-repeat tracts. A greater understanding of both the triplet-repeat DNA structures and the molecular mechanisms responsible for spontaneous mutations will be required before we can understand which of the models discussed above are responsible for repeat instability leading to triplet-repeat–associated human diseases.

References

Bacolla A, Gellibolian R, Shimizu M, Amirhaeri S, Kang S, Ohshima K, Larson JE, et al (1997) Flexible DNA—genetically unstable CTG•CAG and CGG•CCG from human hereditary neuromuscular disease genes. J Biol Chem 272: 16783–16792

Baran N, Lapidot A, Manor H (1987) Unusual sequence element found at the end of an amplicon. Mol Cell Biol 7: 2636–2640

Bianchi ME, Beltrame M (1998) Flexing DNA: HMG-box proteins and their partners. Am J Hum Genet 63:1573–1577

Bidichandani SI, Ashizawa T, Patel PI (1998) The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am J Hum Genet 62:111–121

Chastain PD, Sinden RR (1998) CTG repeats associated with human genetic disease are inherently flexible. J Mol Biol 275:405–411

Darlow JM, Leach DRF (1998) Secondary structures in d(CGG) and d(CCG) repeat tracts. J Mol Biol 275:3–16

Godde JS, Kass SU, Hirst MC, Wolffe AP (1996) Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter. J Biol Chem 271: 24325–24328

Kang S, Jaworski A, Ohshima K, Wells RD (1995) Expansion and deletion of CTG triplet repeats from human disease genes are determined by the direction of replication in *E. coli*. Nat Genet 10:213–218

Krahe R, Ashizawa T, Abbruzzese C, Roeder E, Carango P, Giacanelli M, Funanage VL, et al (1995) Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. Genomics 28:1–14

Krasilnikov AS, Panyutin IG, Samadashwily GM, Cox R, Lazurkin YS, Mirkin SM (1997) Mechanisms of triplex-caused polymerization arrest. Nucleic Acids Res 25:1339–1346

Mitas M (1997) Trinucleotide repeats associated with human disease. Nucleic Acids Res 25:2245–2253

Ohshima K, Montermini L, Wells RD, Pandolfo M (1998) Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication in vivo. J Biol Chem 273:14588–14595

Ohshima K, Wells RD (1997) Hairpin formation during DNA synthesis primer realignment in vitro in triplet repeat sequences from human hereditary disease genes. J Biol Chem 272:16798–16806

Parsons MA, Sinden RR, Izban MG (1998) Transcriptional properties of RNA polymerase II within triplet repeat-containing DNA from the human myotonic dystrophy and fragile X loci. J Biol Chem 273:26998–27008

Pearson CE, Ewel A, Acharya S, Fishel RA, Sinden RR (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. Hum Mol Genet 6:1117–1123

Pearson CE, Sinden RR (1996) Alternative DNA structures within the trinucleotide repeats of the myotonic dystrophy and fragile X locus. Biochemistry 35:5041–5053

Pearson CE, Sinden RR (1998*a*) Slipped strand DNA, dynamic mutations, and human disease. In: Wells RD, Warren ST (eds) Genetic instabilities and hereditary neurological disorders. Academic Press, San Diego, pp 585–621

———(1998*b*) Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. Curr Opin Struct Biol 8: 321–330

Pearson CE, Wang Y-H, Griffith JD, Sinden RR (1998) Structural analysis of slipped strand DNA (S-DNA) formed in $(CTG)_n$• $(CAG)_n$ repeats from the myotonic dystrophy locus. Nucleic Acids Res 26:816–823

Rao BS, Manor H, Martin RG (1988) Pausing in simian virus 40 DNA replication by a sequence containing $(dG\text{-}dA)_{27}$• $(dT\text{-}dC)_{27}$. Nucleic Acids Res 16:8077–8094

Samadashwily GM, Raca G, Mirkin SM (1997) Trinucleotide repeats affect DNA replication in vivo. Nat Genet 17: 298–304

Schumacher S, Fuchs RPP, Bichara M (1998) Expansion of CTG repeats from human disease genesis dependent upon replication mechanisms in *Escherichia coli*—the effect of long patch mismatch repair revisited. J Mol Biol 279: 1101–1110

Sinden RR (1994) DNA structure and function. Academic Press, San Diego

Sinden RR, Wells RD (1992) DNA structure, mutations, and human genetic disease. Curr Opin Biotechnol 3:612–622

Sinden RR, Zheng G, Brankamp RG, Allen KN (1991) On the deletion of inverted repeated DNA in *Escherichia coli*: effects of length, thermal stability, and cruciform formation in vivo. Genetics 129:991–1005

Timchenko LT (1999) Myotonic dystrophy: the role of RNA CUG triplet repeats. Am J Hum Genet 64:360–364 (in this issue)

Trinh TQ, Sinden RR (1991) Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. Nature 352:544–547

Usdin K (1998) NGG-triplet repeats form similar intrastrand structures: implications for the triplet expansion diseases. Nucleic Acids Res 26:4078–4085

Wang Y-H, Gellibolian R, Shimizu M, Wells RD, Griffith J (1996) Long CCG triplet repeat blocks exclude nucleosomes—a possible mechanism for the nature of fragile sites in chromosomes. J Mol Biol 263:511–516

Wang Y-H, Griffith JD (1995) Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. Genomics 25:570-573

———(1996) Methylation of expanded CCG triplet repeat DNA from fragile X syndrome patients enhances nucleosome exclusion. J Biol Chem 271:22937–22940

Wells RD, Warren ST (1998) Genetic instabilities and hereditary neurological diseases. Academic Press, San Diego

Wong L-JC, Ashizawa T, Monckton DG, Caskey CT, Richards CS (1995) Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. Am J Hum Genet 56:114–122