C1 Inhibitor Gene Sequence Facilitates Frameshift Mutations

John J. Bissler, Qin-Shao Meng, and Theresa Emery

The Children's Hospital Research Foundation, Cincinnati, Ohio, U.S.A. Accepted October 23, 1998

Abstract

Mutations disrupting the function or production of C1 inhibitor cause the disease hereditary angioneurotic edema. Patient mutations identified an imperfect inverted repeat sequence that was postulated to play a mechanistic role in the mutations. To test this hypothesis, the inverted repeat was cloned into the chloramphenicol acetyltransferase gene in pBR325 and its mutation rate was studied in four bacterial strains. These strains were selected to assay the effects of recombination and superhelical tension on mutation frequency. Mutations that revert bacteria to chloramphenicol resistance (Cm^r) were scored. Both pairs of isogenic strains had reversion frequencies of approximately 10^{-8} . These rare reversion

Introduction

Hereditary angioneurotic edema is due to a decreased functional level of C1 inhibitor (C1INH). This disease process manifests as spontaneous nonpruritic edema of the skin or mucous membranes. The results of this swelling can be temporary disfigurement, or more serious outcomes such as bowel obstruction or asphyxiation. We have analyzed patients with hereditary angioneurotic edema and found a site of two overlapping frameshift mutations (Fig. 1A) (1,2). This region of the C1INH gene (nucleotides 17922– 17969, accession number X54486), located in exon 8, is an imperfect inverted repeat, or quasipalindrome, and can be configured into an imevents in bacteria were most often a frameshift that involved the imperfect inverted repeat with a deletion or a tandem duplication, an event very similar to the human mutations. Increased DNA superhelical tension, which would be expected to enhance cruciform extrusion, did not accentuate mutagenesis. This finding suggests that the imperfect inverted repeat may form a stem-loop structure in the singlestranded DNA created by the duplex DNA melting prior to replication. Models explaining the slippage can be drawn using the lagging strand of the replication fork. In this model, the formation of a stemloop structure is responsible for bringing the end of the deletion or duplication into close proximity.

perfect hairpin structure (Fig. 1B). A great deal of work has been done with perfect inverted repeats in bacterial models (3-8), and imperfections in the inverted repeat sequence reduce the frequency of mutations (9). Quasipalindromes can also extend the mutation spectrum through intra- or interstrand switch mechanisms (10,11). We hypothesized that the imperfect inverted repeat in the C1INH gene contributed to the frameshift mutations observed in our patients. To test this hypothesis, we used an Escherichia coli model and found that larger frameshift mutations, both tandem duplications and deletions similar to those seen in our patients, occurred at a low frequency independent of recombination and superhelical tension. There are at least 10 different diseases that are caused by mutations associated with imperfect inverted repeats (12). The similar mutation spectrum in the bacterial model and the human C1INH gene may indicate that such a

Address correspondence and reprint requests to: Dr. John J. Bissler, Children's Hospital Research Foundation #5, 3333 Burnet Avenue, Cincinnati, OH 45229-3039, U.S.A. Phone: 513-559-4536; Fax: 513-559-7407; E-mail: john.bissler@chmcc.org





bacterial system can be used to more fully characterize molecular events leading to these inverted repeat-associated diseases.

Materials and Methods

Bacterial Strains and Media

The following K12-derived bacterial strains were used: HB101 [F-hsd20 mcrB mrr recA13 ara14 proA2 galK2 rpsL20 (Sm1) xyl-5 mtl-1 supE44], RR1

[identical to HB101 except it is recA(+)], JTT1 [*pyrF287 nirA iclR7 gal25 rspL195 reA36*], RS2 [identical to JTT1 except it is *topo10A*(-)]. Luria broth contains 10 g NaCl, 5 g yeast extract, and 10 g Bacto Tryptone per liter, pH 7.0.

Oligodeoxyribonucleotide Synthesis, Fragment Purification, and Cloning

Complementary oligonucleotides with EcoRI overhanging ends were synthesized using β -phosphoramidite chemistry, purified on an oligonucleotide purification cartridge (Applied Biosystems, Foster City, CA), and phosphorylated on their 5' end using T4 polynucleotide kinase (GIBCO BRL, Gaithersburg, MD). Equal molar concentrations of each oligonucleotide (AATTCACAGAGACTGGGGGTGGAGGCGGCT-GCAGCCTCCGCCATCTCTGTGG and AATTC-CACAGAGATGGCGGAGGCTGCAGCCGCCTC-CACCCCAGTCTCTGTG) were combined in 10 mM Tris, 50 mM NaCl, and 1 mM EDTA (pH 7.6) and heated in an aluminum temperature block at 95°C. The temperature block was slowly cooled to room temperature over 4 hr. The reaction mixture was resolved on a 15% polyacrylamide gel (2% bis) and the duplex fragment was excised, purified, and ligated into EcoRI cut, calf intestinal alkaline phosphatase (GIBCO BRL, Gaithersburg, MD) treated pBR325 (Fig. 1B). This ligation reaction was transformed into the bacterial strain HB101. Tetracycline-resistant colonies were selected and plasmid DNA was screened by PstI digestion to identify plasmid harboring the cloned sequence. Plasmid with the correct PstI cut sites were then sequenced. The correct clone, pBRCINR, was grown to stationary phase in 3 liters Luria broth, the cleared lysate prepared, and plasmid DNA purified by isopycnic banding to be used for subsequent transformations (13).

Reversion Frequencies and Analysis of Revertants

The vector pBR325 contains three genes that render the host resistant to ampicillin, tetracycline, and chloramphenicol (Cm^r) (14,15). The *Eco*RI site in the chloramphenicol gene has been used extensively for frameshift mutation frequency quantitation (7,8,16,17) as inserts at this site that alter the reading frame render hosts sensitive to chloramphenicol (14). Frameshift mutations that restore the reading frame restore Cm^r. The reversion rate is dependent on the characteristics of the study sequence. In preliminary experiments with a different inverted repeat, the randomly generated control sequence had a slightly higher reversion rate than the inverted repeat sequence. The mutation spectrum was guite different in that the control sequence was limited to a + 1 or a - 2 bp frameshift in a homopolymeric portion of the control sequence, while the inverted repeat had much larger deletions. We concluded that comparison of reversion rates from different sequences is at best crude, and critical information about intrinsic instability is also derived from the mutation spectrum. Early work by Streisinger et al. identified that small insertions and deletions dominate the mutation spectrum in homopolymeric sequence regions (18). An inverse relationship between the size of the frameshift mutation and the frequency of occurrence is also seen in human gene mutations (19). There are multiple sites in the C1 inhibitor inverted repeat that could undergo the necessary 1 or 2 bp frameshift mutations to restore the reading frame and result in reversion. This feature of the sequence allows it to act as an internal control. Analysis of the mutation spectrum determines whether the expected small slipped mispairing events occur, or whether the larger deletions predominate the mutation spectrum. Such large deletions and insertions would support the hypothesis that the inverted repeat formed a secondary structure that participated in the mutagenic process.

For reversion experiments, cultures (30 ml) were grown in LB medium at 37°C with shaking at 225 rpm overnight from a single colony grown on a LB plate under selective pressure from carbenicillin (50 μ g/ml) and tetracycline (25 μ g/ml) (Sigma Chemical Co., St. Louis, MO). Optical density measurements confirmed that the culture had reached stationary phase. The bacteria were concentrated, a 100- μ l aliquot was used to make dilutions to determine a viable cell count, and the remainder was plated on LB plates with selection by carbenicillin (50 μ g/ml), tetracycline (25 μ g/ml), and chloramphenicol (25 μ g/ml). The mutant frequency is determined as Cm^r revertants/*n*, where *n* is the total number of viable cells. The values given are the mean of six independent experiments. "Jackpot" frequenciesrare frequencies that are several orders of magnitude higher and represent a mutation in the starting culture-were not included in the data analysis (20).

To analyze revertants, each revertant was grown overnight in 10 ml cultures of Luria broth with carbenicillin (50 μ g/ml), tetracycline (25

 μ g/ml), and chloramphenicol (25 μ g/ml). The plasmid was harvested by alkaline lysis and used to transform bacteria strain DH5 α , made competent by the RbC1 Hanahan method (21). Retransformation was followed by chloramphenicol selective pressure to isolate purely Cm^r plasmid. (Because of copy number, the plasmid population prior to retransformation contains mostly nonmutant plasmid, making analysis difficult.)

Colonies were then subjected to polymerase chain reaction (PCR) analysis with primers identical to pBR325 nucleotides 4731-4760 (TATCCGGCCTTTATTCACATTCTTGCCCGC) and complementary to nucleotides 4860-4890 (AT-TCACTCCAGAGCGATGAAAACGTTTCAG). The conditions for PCR were as follows: melting at 94°C for 1 min, annealing at 55°C for 2 min, and extension at 72°C for 1 min. The reaction mixture contained 5% formamide (22,23), 50 mM KCl, 10 mM Tris (pH 9), 0.1% Triton X-100, and 1.5 mM MgCl₂. These conditions maximize product yield and eliminate polymerase pausing on the imperfect inverted repeat template (data not shown). PCR products were resolved on a 5% polyacryamide gel (2% bis). Because inverted repeat sequences can cause PCR artifacts, the results were also confirmed by analysis of an AluI digest of CsCl purified plasmid DNA. The AluI fragment of pBR325 containing the EcoRI site is 124 bp and the size of the insert is readily determined from analysis on a 5% polyacryamide gel (2% bis). Because the cloned sequence contains a PstI site, the integrity of the insert sequence was assayed by PstI digestion and resolution on a 2% agarose gel. Sequencing was carried out by labeling the primer complementary to pBR325 nucleotides 4860-4890 using $[\gamma^{-32}P]$ ATP (Dupont-NEN, Boston, MA) and T4 polynucleotide kinase (GIBCO BRL). Plasmid was then used for chain termination sequencing (GIBCO BRL) followed by resolution on a 7% polyacrylamide gel containing 8 M urea (Promega, Madison, WI). The gel was dried and exposed to Kodak XAR film.

To analyze growth rates of bacteria harboring different mutations, purified pBR325 and mutant plasmid bearing bacteria strain DH5 α were grown in 5 ml LB medium with Ampicillin (50 µg/ml), tetracycline (25 µg/ml), and chloramphenicol (25 µg/ml). Aliquots of these cultures were placed in fresh, warmed media under identical selective pressure. Aliquots of bacteria were removed at timed intervals and OD₆₅₀ was measured.

Bacteria	Genotype	Reversion Rate (×10 ⁻⁹)	Deletion	Insertion	Other	Total	
[RR1	F-,gyrA96	16.3 ± 7.9	7	9	14	30	
НВ101	F-, recA13, gyrA96	2.5 ± 3.7	4	21	3	28	
[JTT1	recA56	21.0 ± 27.7	22	23	5	50	
RS2	recA56, topA10	40.8 ± 28.1	29	1	9	39	

Table 1.	Reversion	rate and	mutation	type b	y bacterial	strain ^a
----------	-----------	----------	----------	--------	-------------	---------------------

^aBracketed strains are isogenic; the top strain of the pair is the parental strain and the bottom one is the mutant. A partial genotype is included and the boldface lettering is used to accentuate difference from the parental strain. Each experiment was performed six times. The reversion rate is the average of the quotient of revertant divided by the number of viable bacteria. Standard deviation is included.

Results

Assay of Mutagenic Potential

To assay the mutagenic potential of the imperfect inverted repeat sequence from C1INH, a 44-bp inverted repeat sequence was inserted into the chloramphenicol acetvltransferase gene in the vector pBR325. This resulted in sensitivity to chloramphenicol and the mutant frequency was then quantitated by reversion to Cm^r. Reversion to Cm^r occurs when a frameshift mutation restores the reading frame within, or downstream of, the insert. Mutations that perpetuated a frameshift could not be detected. The mutant frequency was examined in several bacterial strains (Table 1). The RS2 strain and its parental strain, JTT1, were used to determine whether enhanced negative superhelical tension in the plasmid (resulting from a mutation in topoisomerase I) enhance the mutant frequency by increased inverted repeat extrusion into a cruciform structure. Increased deletion of perfect inverted repeats in vivo with increased superhelical tension has been previously reported (7,17). The RS2 strain is identical to the JTT1 strain except for the mutation in the topoisomerase I gene. The reversion frequency of the imperfect inverted repeat was not significantly different between the more negatively supercoiled strain RS2 (40.8 \times 10⁻⁹) and JTT1 (21 \times 10⁻⁹) as assessed by Student's t-test.

The mutant frequencies were also examined in strains HB101 and RR1. These strains are isogenic except the RR1 strain is $recA^+$ and thus recombination proficient. The reversion rate for bacterial strain HB101 was 2.5×10^{-9} and 16×10^{-9} for RR1 (Table 1). These differences also failed to achieve statistical significance by Student's *t*-test.

RecA protein and the human homolog RAD51 are crucial for cellular proliferation because of their pivotal role in recombinatorial repair of doublestrand breaks. RecA polymerizes on singlestranded DNA and facilitates an ATP-dependant strand exchange reaction (24). Our results do not detect a significant change in reversion rate or mutation spectrum between cell lines that can or cannot express RecA. This insensitivity to RecA is also seen in perfect inverted repeats (25) and supports the hypothesis that inverted repeat facilitated mutagenesis does not require strand invasion.

Analysis of Frameshift Mutations

Mutant plasmid was purified by retransformation into the bacterial strain DH5 α . The transformed bacteria, harboring only the Cm^r gene, were grown and analyzed by PCR. Figure 2A illustrates an example of the data obtained for plasmid in the bacterial strain JTT1. Frameshift mutations >5 bp were detected by a change in migration pattern from the expected 203 bp. All large deletions lost the PstI site in the insert, and all small deletions or insertions maintained it (data not shown). Mutation types identified by PCR were subjected to plasmid isolation and restriction digestion analysis to confirm that the mutations occurred in the revertant plasmid and not as a result of an artifact of the PCR caused by transient intra- or interstrand template switching. Revertant plasmid was digested with AluI and resolved on a 5% polyacrylamide gel. The fragment containing the EcoRI cloning site migrates at 124 bp and frameshift mutations can be identified by a change in migration pattern. Figure 2B is an example of such an analysis. The PCR and restriction analyses were always in



Fig. 2. A PCR-based strategy to detect deletions and insertions. (A) This representative gel identifies frameshift mutations in JTT1 strain revertant clones. PCR products were resolved on a 5% polyacrylamide gel. Complete deletion (44 bp) results in a 160-bp fragment. Lanes 4, 5, 6, 9, 13, 14, and 16 have small frameshift mutations that are undetectable by this assay, while lanes 2, 10, 11, and 12 have large deletions and lanes 1, 7, 8, and 18 have insertions. Lanes 3, 15, and 17 illustrate short deletions. (B) Selected AluI digestions of revertant plasmid. Lane 1 contains the 100 bp marker, lane 2 contains the original cloning vector pBR325, lane 3 contains pBRCINR, the original plasmid used for transformation. Lane 4 contains plasmid with an insertion corresponding in size to lane 1 in Figure 2a. Lane 5 contains a deletion corresponding to lane 2 in Figure 2a. Lane 6 contains a frameshift mutation below the limit of detection of this assay, and lane 7 contains a small deletion. The restriction digestion analysis agreed with the PCR results in every case.

agreement. Table 1 lists the total number of deletions, insertions, and short frameshift mutations for the four bacterial strains and correlates the integrity of the cloned sequence by restriction analysis. According to PCR product and restriction endonuclease fragment size, large deletions and insertions were more abundant than undetectable frameshift mutations.

Sequence Analysis

To understand the precise frameshift mutations responsible for reversion of Cm^r in the bacterial model, approximately half of the revertant clones were sequenced (n = 61). We analyzed the mutations from the recA⁻ bacteria to avoid any RecA-mediated recombination events. We analyzed 14 revertant clones from the HB101, 12 from RS2, and 35 from JTT1 strains. The selection criterion used to select plasmids for sequencing was based on the detection of a frameshift mutation by electrophoresis. Several plasmids from each migration pattern were chosen and the rest were chosen from the undetectable frameshift mutant groups. The smallest deletion that can restore the reading frame is 2 bp and the inverted repeat sequence contains multiple potential sites in the sequence compatible with this sort of mutation. There were 58 detected deletions from revertant bacteria (n = 117). Of these, 51 (88%) were approximately 40-bp deletions, compatible with a complete deletion of the insert. Sequence analysis of these large deletions revealed two overlapping deletion patterns with slightly different boundaries, as illustrated in Figure 3A. The first mutation in the figure (labeled a) uses a 4-bp direct repeat (CATC) on which slip mispairing appears to occur. The second mutation (labeled b) appears to have utilized a single T. These mutation patterns represent all of the large deletions. The vector pBR325 has a unidirectional origin of replication; therefore, these frameshifts can be modeled as occurring on the lagging strand of the replication fork (see below). Because the selection criterion favored sequencing plasmids from the group of undetected frameshift mutations, the number of deletion mutations per bacterial strain in Figure 3A appears to be low. However, the plasmids that appeared to have a 40-bp deletion (labeled a and b), 4 in the RS2 strain and 6 in the JTT1 strain, all had the same deletion, as depicted in the first mutation below the plasmid sequence in Figure 3A. Likewise, of the three selected HB101 clones that appeared to have a 40-bp deletion (labeled a and b), two had a mutation similar to the JTT1 and RS2 mutations, and one had a similar deletion that was 2 bp larger. The larger



Deletions:	Deat	arial C	t main.
	RS2	HB101	JTT1
CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGCGGGGGGGG	С		
a.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCA TCTCTGTGAATTCC b.) CTTTATTCACATTCTTGCCCGCCCTGATGAATGCTCA TCTCTGTGAATTCC c.) CTTTATTCACATTCTTGCCCGCCTGATGAAGCT GACTGGGGTGGAGCGCGCGCGCGCCCCCGCCATCCTGTGAATTCC c.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCT GACTGGGGTGGAGCGCGC GCCATCTCTGTGAATTCC c.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGGGTGGGAGGCGGC CCCATCTCTGTGAATTCC c.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGGAGTGGGGGGGG	4	2 1	6 1 1 1 1 1 1 1 1
m.) CTTTATTCACATTCTTGCCCGCCTGATGAAT CT GACTGGGGTGGAGGCGGGTGCAGCCTCCGCCATCTCTGTGAATTC n.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATC GGAATTC CAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCGCCATCTCTGTGAATTC o.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCCGCCATCTCTGTGAATTC p.) CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCGCCATCTCTGTGAATTCC AGGCGGCTGCAGCCTCCCGCCTGATGAATGCTCATCCGGAATTC	C C C C		1 1
D Insertions:			
	Bact	erial S	train
CTTTATTCACATTCTTGCCCGCCTGATGATGCTCATCCGGAATTCACAGAGACTGCGGGGGGGG	RS2	HB101	JTT1
CTTTATTCACATTCTTGCCC GCCTGAGTAATGCTCATCCGGAATTCACAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCGCCATCTCTGTGAATTCC		2	
ctttattcacattcttgcccgcctgatgaatgctcatccggaattcacagagactgggggggg	2		
CTCATCC.			
CTTTATTCACATTCTTGCCCGCCTGATGATGCTCATCGGAATTCACAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCGCCATCTCTGTGAATTCC		1	1
CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGGGGGGGG	1		
CTTTATTCACATTCTTGCCCCCCTCATGAATGCTCATCCGGAATTCACAGAGACTGGGGGGGG			1
C			
Combination Insertion/Deletion:			
	Bacte:	rial St	rain:
CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGCGGGGGGGG	RS2	HB101	JTT
ACATTCTTGCCCGCCTGATGAATGCTCA		1	
		-	
TTCTTGCCCGCCTGATGAA			
CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATC TCTGTGAATTCC			1
CTCC			
CTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCACAGAGACTGGGGTGGAGGCGGCTGCAGCCTCCGCC TCTCTGTGAATTCC			1
СФФФ 3 ФФС 3 С 3 ФФСФФСССССССФС3 ФС3 3 ФССФС 3 ФФС С3 С3 С9 С9 С2 С2 С2 С2 С0 С9 СССССС 3 ФСФСФССС 3 ФСФСФСФС3 3 ФФС			1

Fig. 3. Sequence alignment of mutations. The top line is the sequence from plasmid pBRCINR including the 5' sequence and both *Eco*RI sites. (A) Deletions. The sequence from 38 bp 5' to the EcoR1 site, the inserted sequence, and the 3' EcoR1 site are illustrated. The restriction sites are identified by a line over the sequence. The inverted repeat is identified by the opposing arrows above the sequence. Unpaired bases in the proposed stem-loop structure are white with a black background, and bases across from these unpaired bases are black with a grav background. The hairpin loop is boxed. Arrows are above the inverted repeat sequence. Deletions are aligned beneath by gaps in the sequence. The sequences are referred to by the lowercase letters to the left of the sequences. To the right are the number of times this was identified in the three bacterial strains. Because the clones selected for sequence

were representative of mutation types and not every mutation event scored by reversion, this number is not indicative of the total number of times the event occurred. (B) Insertions. The sequence from 38 bp 5' to the EcoR1 site, the inserted sequence, and the 3'EcoR1 site are illustrated. The restriction sites are identified by a line over the sequence. The opposing arrows over the sequence depicts the inverted repeat. Unpaired bases in the proposed stem-loop structure are white with a black background, and bases across from these unpaired bases are black with a gray background, while the hairpin loop is boxed. The insertions are depicted in black and their template is illustrated by gray background shading. (C) Complex mutations. The coding is the same as in A and B. The inserts are in black and the templated sequences have gray background shading, while the gaps represent deletions.

deletion endpoints as well as the smaller deletions map to sites that are, or are across from, mismatched base pairs or at the 5' *Eco*RI site.

All of the insertion mutants larger than a single base pair involved the 5' EcoRI cloning site; these are illustrated in Figure 3B. Two of the inserts, a 19-bp and a 7-bp tandem duplication, overlap. There is one complex insertion that templates twice from the sequence 5' to the EcoRI site (bottom of Fig. 3B). Taken together, these insertions are presumed to be due to the DNA polymerase pausing at the imperfect stemloop structure and ratcheting on the template to produce the duplications. One mutation at the 3' EcoRI site is a +1 frameshift derived from slippage on thymidine. The G insertion at position 4822 occurs at a site that would be mismatched in the extruded cruciform. Four complex mutations that involved both deletion and tandem duplication were also identified. The insertions occurred either in the sequence that would immediately precede the extruded stem-loop structure or, in the case of the CTCC insertion, at a site corresponding to mismatched bases in the proposed stem-loop structure.

Discussion

Association of Mutation with the Stem-Loop Structure from C1 Inhibitor

Inverted repeats have been shown to be mutagenic in bacterial systems and imperfect inverted repeat sequences are associated with human disease (12). Mutations in the C1 inhibitor gene lead to an identifiable phenotype (hereditary angioneurotic edema). Approximately 120 different mutations leading to this disease have been identified in separate kindred (2). Two of these mutations involve the inverted repeat studied here and exhibit a similar duplication and deletion size. The frequency of these large frameshift mutations in humans is similar to that in the bacteria. The 5' and 3' restriction recognition sequence (EcoRI) at the ends of the imperfect palindrome, which would be excellent direct repeats for slipped mispairing, are involved less often in mutagenesis by slipped mispairing than by mutations involving the sequence of the stem-loop structure at sites of base mismatch in the extruded cruciform. This supports the hypothesis that the secondary structure is involved in mutagenesis and that the imperfect nature of the inverted repeat affects the mutation spectrum. In a quantitative assay system using a human cell line (26), the reversion rate was similar to a bacterial system with a similar-sized inverted repeat (27). A nonpalindromic sequence in this human system failed to show reversion. These results are consistent with the hypothesis that inverted repeat sequences facilitate mutagenesis both in prokaryotic models and in humans.

The clustering of these tandem duplications just prior to, and a deletion involving, the imperfect inverted repeat in these experiments suggest that this sequence mechanistically contributes to mutagenesis. Some revertants revealed complex mutations that included both insertions and deletions (Fig. 3C). The insertion site remained 5' to the imperfect inverted repeat, and the deletion segment involves the quasipalindrome at sites corresponding to stem-loop mismatched base pairs. These large deletions are similar in that they slip mispair on the CATC sequence found most frequently in the larger deletions. The insertions are 6 bp apart and again are tandem duplications. Both appear to slip mispair on an adenine residue. These complex mutations could be explained if the insertion mutations were followed by subsequent deletion mutations.

Role of Stem-Loop Structures in Mutagenesis

There are several possible ways cruciforms can be eliminated from the genome. Cruciforms could be involved in intra- or intermolecular recombination (28), or they could be removed by Holliday structure resolvase (3,29). A model can be drawn so that the cruciform formed in the lagging strand brings the 5' and 3' ends of the deletion in close proximity, enabling slipped mispairing to occur, ultimately leading to deletion (Fig. 4A). Slippage at the EcoRI restriction site during DNA replication could be predicted to mediate complete deletion; however, this was not identified in our studies. The partial deletion can be explained by slippage involving the "unpaired bubble" illustrated in Figure 4A. The first part of the cruciform would be melted down to the cluster of unpaired bases. If forward slippage occurred, followed by further replication, the deletion could be explained. The mutation could then be fixed in the plasmid by a further round of replication that would result in one deletion and one unchanged plasmid. Precisely how the unpaired bases are involved is not yet clear. They may contribute simply by the helical distortion (30,31) or flexibility they induce, causing an unfavorable milieu for the DNA replication machinery. These sites may also be tagged by protein



Fig. 4. Continued on next page.





the bottom illustrate the same sort of mechanism for the other large deletion slipping on the single T. (B) As in A, the bottom nucleotides are the template strand and the top nucleotides represent the newly synthesized strand. The polymerase pauses at the junction with the stem-loop structure, allowing for the duplex to melt and reanneal using a single guanine base pair. Polymerization then continues and either successfully negotiates the cruciform the second time or, as seen in Figure 3C, may undergo further mutagenesis.

binding, and the binding in some way induces mutation.

Duplications can also arise as a result of cruciform formation in the lagging strand (Fig. 4B). The extruded cruciform, forming by intrastrand base-pairing before single-strand binding proteins (SSB) can coat the single-stranded DNA, may pose an impediment to polymerase translocation along the DNA template. This pause may provide time for partial melting of the template and copied strands, followed by slipped mispairing and reannealing. The second pass of the polymerase would then successfully negotiate the cruciform and the mutation would be fixed in the plasmid population by another round of replication. Because this is such a rare event, as judged by the reversion rate, we were unable to identify successive repeats where the inverted repeat caused cycled tandem duplications, which may be important in triplex repeat expansion (32); however, the combination of tandem insertion and deletion mutants in this study may indicate that after the insertion, the cruciform caused a second pausing followed by a forward instead of a backward slippage, resulting in a deletion. Only two of these mutations were identified.

Cellular Features Affecting Quasipalindrome-Associated Mutagenesis

The bacterial reversion seen in these studies may be explained as a manifestation of adaptive mutation (33-35). Perhaps some special mutagenic quality can be attributed to chloramphenicol; perhaps as protein synthesis wanes in the presence of chloramphenicol the nuclear protein kinetic alterations in the cell activates the mutagenic potential (36). Proteins involved in preventing alternative DNA structures from forming, such as SSB, may be reduced (37). In this setting, intrastrand duplex structures would increase. These mutations then occur during replication attempts because of a reduction in a critical factor(s). The precise cause of the delay in identifying mutants is unclear and could potentially involve the interaction of the replication fork with transcription machinery (38,39). Studies with different mutations identified in this work do not support the notion that bacteria harboring different types of mutations grow at significantly different rates (data not shown).

The quasipalindrome in these experiments appeared to direct the type of slipped mispairing in-

termediate and mutation. However, this effect did not change with an increase in superhelical tension. This suggests that this quasipalindrome, because of its imperfect nature, does not have sufficient energy to increase mutagenesis resulting in increased numbers of the same sort of deletion mutations. The unpaired bases may inhibit the intrastrand annealing to form the complete duplex structure. We reason that the imperfect inverted repeat sequences affect mutations when it is in a single-stranded state, and superhelical density does not catalyze this type of mutation. (Alternatively, the increased superhelical tension may promote other mutations that do not restore Cm^r and cannot be identified in this assay system.) The bacterial strain variability in insertions and deletions is not clearly understood; it may be a result of the low number of mutations.

A model explaining this mutagenesis involves the lagging strand during DNA synthesis. The delay between duplex DNA melting and replication is longer on this strand. The fact that perfect palindromes exhibit an increased mutation rate under increased superhelical tension is perhaps because mutagenesis is also occurring in the leading strand. The increase in perfect palindrome mutagenesis in the RS2 strain may be due to the fact that the polymerase holoenzyme actually encounters an extruded cruciform instead of formation during DNA replication. This hypothesis would suggest that the balance between DNA replication fidelity and mutagenesis is in part determined by the kinetic balance between intrastrand secondary structure formation and loss of this potential by protein-DNA interaction. Imperfect palindromes, because of the impediment of intrastrand DNA duplex zippering by base mismatches, are predicted to be less of a mutagenic threat.

In summary, the results of these studies support the hypothesis that the imperfect inverted repeat sequence from the C1 inhibitor gene reactive center promotes mutations, as seen in the reported kindreds with hereditary angioneurotic edema (1). The size of deletion and insertion mutations in this study using a bacterial model was quite similar to those reported in patients with hereditary angioneurotic edema. The mutation similarities in the bacterial model to those in patients suggests that this model may be useful in further defining the precise mechanism of inverted repeatmediated mutagenesis.

Acknowledgments

This work was supported by grant DK0241801 from the National Institute of Diabetes and Digestive and Kidney Diseases. The authors are deeply in debt to Drs. Kathleen Dixon, Richard R. Sinden, William Rosche, and Alvin E. Davis III for their discussions, encouragement, and support. We thank Judy Howard for her secretarial talents.

References

- 1. Bissler JJ, Donaldson VH, Davis AR. (1994) Contiguous deletion and duplication mutations resulting in type 1 hereditary angioneurotic edema. *Hum. Genet.* **93:** 265–269.
- 2. Bissler JJ, Aulak KS, Donaldson VH, et al. (1997) Molecular defects in hereditary angioneurotic edema. *Proc. Assoc. Am. Phys.* **109:** 164–173.
- 3. Drake JW, Glickman BW, Ripley L. (1983) Updating the theory of mutation. *Am. Sci.* **71:** 621–630.
- 4. Ripley LS, Drake JW. (1984) Bacteriophage T4 particles are refractory to bisulfite mutagenesis. *Mutat. Res.* **129:** 149–152.
- 5. DasGupta U, Weston-Hafer K, Berg DE. (1987) Local DNA sequence control of deletion formation in *Escherichia coli* plasmid pBR322. *Genetics* **115**: 41–49.
- Glickman BW, Ripley LS. (1984) Structural intermediates of deletion mutagenesis: a role for palindromic DNA. *Proc. Natl. Acad. Sci. U.S.A.* 81: 512– 516.
- 7. Trinh TQ, Sinden RR. (1991) Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli. Nature* **352:** 544–547.
- 8. Trinh TQ, Sinden RR. (1993) The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli*. *Genetics* **134**: 409–422.
- Papanicolaou C, Ripley LS. (1991) An in vitro approach to identifying specificity determinants of mutagenesis mediated by DNA misalignments. J. Mol. Biol. 221: 805–821.
- Ohshima A, Inouye S, Inouye M. (1992) In vivo duplication of genetic elements by the formation of stem-loop DNA without an RNA intermediate. *Proc. Natl. Acad. Sci. U.S.A.* 89: 1016–1020.
- 11. Ripley LS. (1982) Model for the participation of quasipalindromic DNA sequences in frameshift mutation. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 4128–4132.
- 12. Bissler JJ. (1998) DNA inverted repeats and human disease. *Front Biosci.* **3:** D408-D418.
- Sambrook J, Fritsch EF, Manianius T. (1989) Molecular Cloning Laboratory Manual. Cold Spring Harbor Wyatt Press, Plainfield, NY.

- 14. Bolivar F. (1978) Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique *Eco*RI sites for selection of *Eco*RI generated recombinant DNA molecules. *Gene* **4**: 121–136.
- 15. Prentki P, Karch F, Iida S, Meyer J. (1981) The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. *Gene* 14: 289-299.
- Rosche WA, Trinh TQ, Sinden RR. (1995) Differential DNA secondary structure-mediated deletion mutation in the leading and lagging strands. *J. Bacteriol.* 177: 4385–4391.
- 17. Sinden RR, Zheng GX, 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.
- 18. Streisinger G, Okada Y, Emrich J, et al. (1966) Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* **31:** 77–84.
- 19. Cooper DN, Krawczak M. (1993) Human Gene Mutation BIOS Scientific Publishers, Oxford, U.K.
- 20. Luria SE, Delbruck M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491–511.
- 21. Hanahan D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-580.
- 22. Varadaraj K, Skinner DM. (1994) Denaturants or cosolvents improve the specificity of PCR amplification of a G+C-rich DNA using genetically engineered DNA polymerases. *Gene* **140**: 1–5.
- 23. Rychlik W. (1995) Priming efficiency in PCR. *Biotechniques* 18: 84–90.
- 24. Baumann P, West SC. (1998) Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem. Sci.* 23: 247–251.
- 25. Rosche WA, Trinh TQ, Sinden RR. (1995) Differential DNA secondary structure-mediated deletion mutation in the leading and lagging strands. *J. Bacteriol.* **177:** 4385–4391.
- 26. Kramer PR, Stringer JR, Sinden RR. (1996) Stability of an inverted repeat in a human fibrosarcoma cell. *Nucl. Acids Res.* **24**: 4234–4241.
- 27. Kazic T, Berg DE. (1990) Context effects in the formation of deletions in *Escherichia coli*. *Genetics* **126:** 17–24.
- Bi X, Liu LF. (1996) DNA rearrangement mediated by inverted repeats. *Proc. Natl. Acad. Sci.* U.S.A. 93: 819-823.
- 29. Lilley DM, Kemper B. (1984) Cruciform-resolvase interactions in supercoiled DNA. *Cell* **36**: 413–422.
- 30. Pulyaeva H, Zakharov SF, Garner MM, Chrambach A. (1994) Detection of a single base mismatch in double-stranded DNA by electrophoresis on uncrosslinked polyacrylamide gel. *Electrophoresis* **15**: 1095–1100.
- 31. Boulard Y, Cognet JA, Fazakerley GV. (1997) So-

lution structure as a function of pH of two central mismatches, C.T and C.C, in the 29 to 39 K-*ras* gene sequence, by nuclear magnetic resonance and molecular dynamics. *J. Mol. Biol.* **268**: 331–347.

- 32. Jaworski A, Rosche WA, Gellibolian R, et al. (1995) Mismatch repair in *Escherichia coli* enhances instability of (CTG)n triplet repeats from human hereditary diseases. *Proc. Natl. Acad. Sci.* U.S.A. **92:** 11019–110123.
- 33. Rosenberg SM, Thulin C, Harris RS. (1998) Transient and heritable mutators in adaptive evolution in the lab and in nature. *Genetics* **148**: 1559–1566.
- 34. Cairns J, Foster PL. (1991) Adaptive reversion of a frameshift mutation in *Escherichia coli. Genetics* **128:** 695–701.
- 35. Foster PL, Cairns J. (1992) Mechanisms of directed mutation. *Genetics* 131: 783–789.

- Ninio J. (1991) Transient mutators: a semiquantitative analysis of the influence of translation and transcription errors on mutation rates. *Genetics* 129: 957–962.
- Carraway M, Rewinski C, Marinus MG. (1990) Mutations produced by DNA polymerase III holoenzyme of *Escherichia coli* after in vitro synthesis in the absence of single-strand binding protein. *Mol. Microbiol.* 4: 1645–1652.
- Liu B, Alberts BM. (1995) Head-on collision between a DNA replication apparatus and RNA polymerase transcription complex. *Science* 267: 1131– 1137.
- Deshpande AM, Newlon CS. (1996) DNA replication fork pause sites dependent on transcription. *Science* 272: 1030–1033.