Seven-Base-Pair Inverted Repeats in DNA Form Stable Hairpins *in Vivo* in *Saccharomyces cerevisiae*

Dilip K. Nag and Thomas D. Petes

Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280

Manuscript received June 10, 1991
Accepted for publication July 19, 1991

ABSTRACT

Palindromic sequences in single-stranded DNA and RNA have the potential for intrastrand base pairing, resulting in formation of "hairpin" structures. We previously reported a genetic method for detecting such structures *in vivo* in the yeast *Saccharomyces cerevisiae*. Below, we describe evidence indicating that a 14-base-pair palindrome (7 bp per inverted repeat) is sufficient for formation of a hairpin *in vivo*.

PALINDROMIC DNA sequences (inverted repeats) are common in both prokaryotic and eukaryotic genomes. Although it is clear that such sequences would allow intrastrand base pairing in single-stranded DNA, the *in vivo* structural configuration of these sequences in double-stranded DNA is less certain. *In vitro*, in supercoiled plasmids, inverted repeats sometimes pair in the cruciform configuration, in which both strands of the helix form intrastrand hairpins disrupting (locally) the interstrand pairing conformation (GELLERT *et al.* 1979; GIERER 1966; LILLEY 1980). Kinetic studies suggest, however, that cruciform formation is not likely to be common *in vivo* (GELLERT, O'DEA and MIZUUCHI 1983). Electron microscopic examination of psoralen-crosslinked DNA failed to reveal cruciforms in bacterial and Drosophila DNA (CECH and PARDUE 1976; SINDEN, BROYLES and PETITJOHN 1983).

There are, however, several indications that DNA secondary structures may be biologically important. First, long palindromes in plasmid and bacteriophage DNA are unstable, although these sequences can be stabilized in nuclease-deficient strains of *Escherichia coli* (COLLINS 1981; LEACH and STAHL 1983; CHALKER, LEACH and LLOYD 1988). Second, palindromes greater than 26 bp in length stimulate deletion formation (WESTON-HAFER and BERG 1991). Third, certain mutations in *E. coli* appear to reflect mismatch correction occurring between palindromic sequences (RIPLEY and GLICKMAN 1983). Fourth, efficient recognition of promoters by N4 virion RNA polymerase requires single-stranded DNA with a hairpin structure at a specific position (S. GLUCKSMANN, P. MARKIEWICZ and L. ROTHMAN-DENES personal communication). Finally, as described below, we found that heterozygous palindromic insertions in heteroduplex DNA are poorly recognized by the yeast mismatch repair system (NAG, WHITE and PETES 1989).

In these experiments, we constructed a diploid yeast strain in which one copy of the *HIS4* gene (a gene required for histidine biosynthesis) had an insertion of a 26-bp palindromic oligonucleotide and the other copy was wild type. The heterozygous diploid was sporulated, and tetrads were dissected onto a rich growth medium. The colonies that formed from the spores were replica-plated onto medium lacking histidine. Although the majority of the tetrads had two His^+ and two His^- spore colonies as expected from simple Mendelian segregation, about 25% of the tetrads had one or more colonies that sectored His^+/His^-. Such sectoring events are called "postmeiotic segregation" (PMS) events (FOGEL, MORTIMER and LUSNAK 1981).

For most heterozygous alleles in *Saccharomyces cerevisiae*, PMS events are very infrequent relative to other types of segregation. Typically, 95% of tetrads segregate 2:2, about 5% show gene conversion (3:1 or 1:3), and only 0.2% show PMS (FOGEL, MORTIMER and LUSNAK 1981). Thus, for most alleles, PMS events represent less than 10% of the aberrant (non-2:2) segregations. According to most models of recombination, both gene conversion and PMS events reflect heteroduplex formation between homologous chromosomes, an intermediate in recombination (HOLLIDAY 1964; MESSELSON and RADDING 1975). If the heteroduplex includes a region of the chromosome that is heterozygous, a mismatch occurs. If the mismatch is efficiently corrected, either a gene conversion event or a restoration of 2:2 segregation is observed; failure to correct the mismatch results in a PMS event. Since most mutant alleles in yeast result in less PMS than gene conversion, it follows that most types of mismatches in yeast are efficiently repaired. There are, however, two types of mutations that show very high levels of PMS: single base-pair changes resulting in a C/C mismatch (FOGEL, MORTIMER and

LUSNAK 1981; DETLOFF, SIEBER and PETES 1991) and palindromic insertions (NAG, WHITE and PETES 1989).

A heteroduplex formed between a wild-type strand and mutant strand with an insertion would be expected to contain a single-stranded loop, representing the sequences in the insertion that have no complement in the wild-type strand. If the insertion is palindromic, one would expect that the sequences within the insertion would form intrastrand base pairs, resulting in a stem-loop (hairpin) structure. We believe that this structure is inefficiently recognized by the yeast mismatch repair system, perhaps because a protein binds to the hairpin and protects it from repair. The conclusion that intrastrand secondary structure of the insertion is responsible for the high level of PMS (rather than the primary sequence) is based on the following observations: (1) changes within the insertion that interfere with intrastrand base pairing lower the frequency of PMS, (2) alteration of the changed sequence to restore base pairing elevates PMS, and (3) all palindromic sequences examined (3 of 3), regardless of their sequence, have high levels of PMS (NAG, WHITE and PETES 1989). In this report, using PMS as an assay system, we examined the minimal size necessary for a palindromic sequence to form a stable hairpin *in vivo*.

MATERIALS AND METHODS

Plasmids: Standard molecular techniques were used for all plasmid constructions (MANIATIS, FRITSCH and SAMBROOK 1982). All plasmids contained palindromic oligonucleotides inserted into the plasmid pDN9 (NAG, WHITE and PETES 1989). The plasmid pDN9 was constructed by cloning a *Xho*I-*Bgl*II fragment containing most of the *HIS4* coding region into *Sal*I-and *Bam*HI-treated YIp5. The first seven oligonucleotides shown in Figure 1 are perfect palindromes. Self-annealing of these oligonucleotides results in a double-stranded molecule with 4-bp cohesive ends (TCGA) compatible with *Sal*I. These double-stranded molecules were inserted into the *Sal*I restriction site of pDN9. The oligonucleotides IR7A and IR7A′ are complementary, and contain both palindromic and non-palindromic sequences. These oligonucleotides were annealed together and inserted into the *Sal*I site of pDN9. Similarly, IR7B and IR7B′ were annealed and inserted into the *Sal*I site of pDN9 to construct a non-palindromic insertion in *HIS4*. All constructions were confirmed by DNA sequence analysis.

Strains: The haploid strains used in these studies were derivatives of AS13 (**a** *leu2-Bs ura3-52 ade6*) or AS4 (*α trp1-1 arg4-17 tyr7-1 ura3 ade6*) (STAPLETON and PETES 1991; SYMINGTON and PETES 1988). Haploid strains with the wild-type chromosomal copy of *HIS4* gene replaced by a mutant copy were constructed using the two-step transplacement procedure (SCHERER and DAVIS 1979). The plasmids containing the *his4* genes with the oligonucleotide insertions were linearized with *SnaBI* and Ura$^+$ transformants were selected. For the second step of the transplacement, Ura$^-$ derivatives were selected with 5-fluoro-orotate (BOEKE, LACROUTE and FINK 1984). Southern analysis was done to confirm the structure of all transformants.

Name	Sequence
Lop	TCGAGTACTGTATGTACATACAGTAC
IR9	TCGAGTACTGTATATACAGTAC
IR7	TCGAGTACTGTACAGTAC
IR5	TCGAGTACTAGTAC
IR4	TCGAGTACGTAC
IR3	TCGAGTATAC
IR2	TCGAGTAC
IR7A	TCGAGTGCTTGGTGAGTACTAGTACTCGAGTTCGCC
IR7A'	TCGAGGCGAACTCGAGTACTAGTACTCACCAAGCAC
IR7B	TCGAGTGCTTGGTGAGTTCGCC
IR7B'	TCGAGGCGAACTCACCAAGCAC

FIGURE 1.—Oligonucleotides used to construct palindromic insertions. The sequence of each oligonucleotide is shown 5′ to 3′. For the first seven, self-annealing results in a double-stranded palindromic sequence with 4-bp cohesive ends (TCGA). These sequences were inserted into the *Sal*I restriction site in the plasmid pDN9 (NAG, WHITE and PETES 1989).

For all strains, except DNY67, the mutant *his4* gene was present in the AS13 background and the diploid strains were constructed by mating *his4* mutant derivatives of AS13 with AS4. In DNY67, the *his4* mutation with the in-frame palindromic insertion (*his4-IR7A*) was in AS13; the AS4 derivative contained an out-of-frame non-palindromic insertion (*his4-IR7B*) in *his4*.

Genetic techniques: Standard genetic methods and media were used (SHERMAN, FINK and HICKS 1983). Diploid strains heterozygous for various *his4* mutant alleles were sporulated (usually at 18°) and dissected onto solid rich growth medium (YPD; SHERMAN, FINK and HICKS 1983). After 3 days of growth at 30°, the spore colonies were replica-plated to medium lacking histidine. The segregation patterns were scored after 24 hr of growth at 30°.

The *HIS4* gene has three parts (*HIS4A*, *HIS4B*, *HIS4C*), each responsible for a different enzymatic activity (KEESEY, BEGELIS and FINK 1979). Since the *Sal*I restriction site, used to insert all the oligonucleotides, is in the *HIS4A* region (at the 5′ end of the gene), any out-of-frame insertion at this position causes loss of all three enzymatic activities, whereas any in-frame insertion at this position results in loss only of *HIS4A* activity. Therefore, to score sectored colonies for DNY67, we mated all the spore colonies derived from this strain to a tester strain containing the *his4C* mutation (which complements the *his4A* mutation). After 10 hr of incubation at 30°, the mated cells were replica-plated to medium lacking histidine to score the *his4* segregation pattern.

RESULTS AND DISCUSSION

To determine the minimum length of an inverted repeat required to form a hairpin structure *in vivo*,

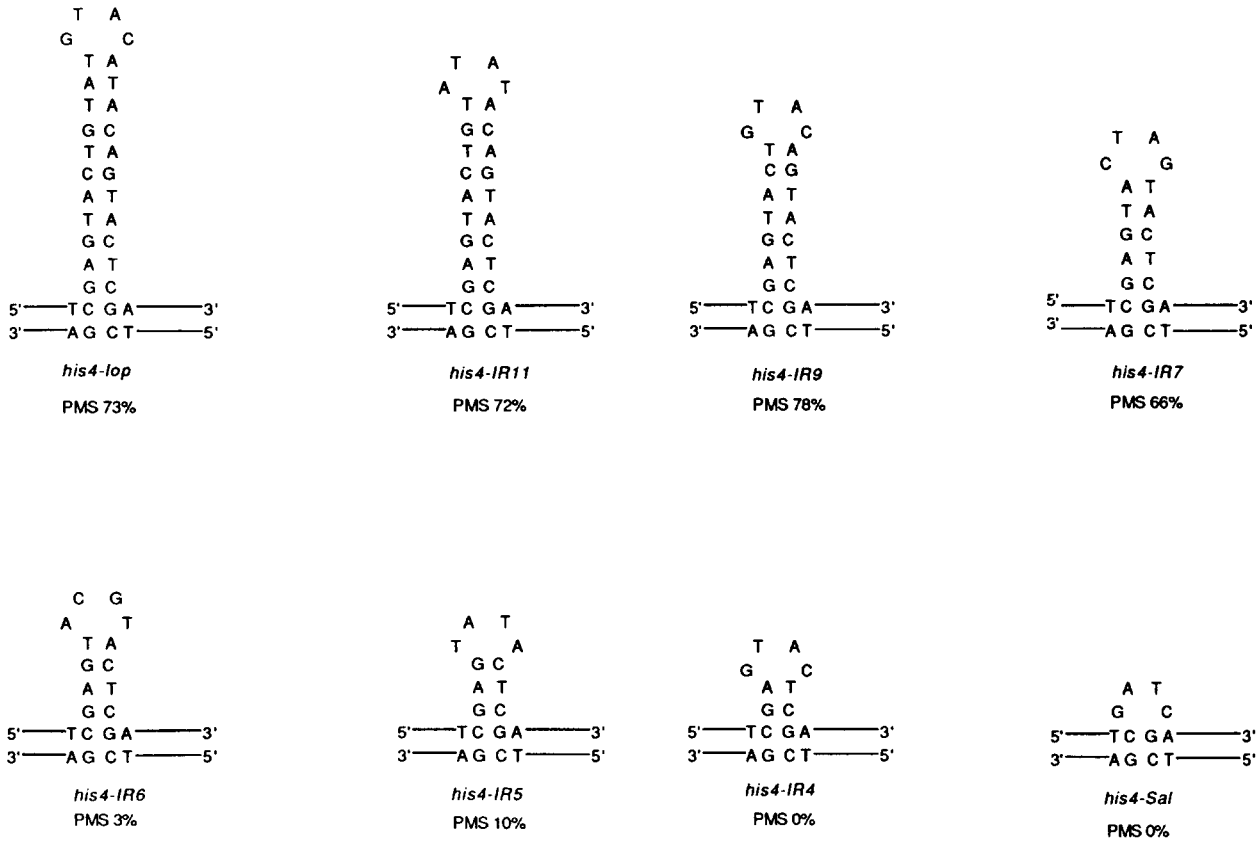

FIGURE 2.—Expected hairpin configuration in heteroduplex formed between a wild-type strand and mutant DNA strand with a palindromic insertion. The hairpin is drawn with a 4-bp loop, although some studies suggest the possibility of a 2-bp loop. Below each configuration, the level of PMS (as percentage of total aberrant segregations) is indicated. All palindromic insertions except *his4-Sal*, were constructed as described in MATERIALS AND METHODS. The *his4-Sal* mutation is the result of filling-in the cohesive ends of *Sal*I-cleaved DNA and ligating together the resulting ends (NAG, WHITE and PETES 1989). This procedure results in a 4-bp palindromic insertion.

we inserted several palindromic oligonucleotides of various lengths into a *Sal*I site within the *HIS4* gene in plasmid pDN9 (NAG, WHITE and PETES 1989; see MATERIALS AND METHODS). The resultant mutant alleles were transplaced into the *HIS4* chromosomal locus. Diploid strains heterozygous for these insertions were constructed and induced to undergo meiosis.

Previous work has indicated a high frequency of heteroduplex formation during meiotic recombination at *HIS4* (NAG, WHITE and PETES 1989); heteroduplexes containing heterozygous palindromic insertions would be expected to contain stem-loop structures (Figure 2). As shown in Table 1 and Figure 2, inverted repeats of lengths 7, 9, 11 or 13 bp had high levels of PMS; 66–78% of the aberrant segregations were PMS. Inverted repeats of 6, 5, 4, or 2 bp had much lower levels of PMS (0–10% of total aberrant segregations). Since the level of PMS is related to the secondary structure of the insertion (NAG, WHITE and PETES 1989), this result indicates that base-pairing for a 7-bp, but not a 6-bp, inverted repeat is stable *in vivo*. This stability, of course, may be influenced by proteins that bind to the hairpin structure.

The palindromic oligonucleotides shown in Figure 1 were inserted into the palindromic *Sal*I restriction site. It was possible, therefore, that base-pairing might occur within these sequences extending the length of the stem. In order to determine whether the bases flanking the insertion contributed to the stability of the hairpin, we constructed a diploid strain (DNY 67) in which a palindromic oligonucleotide with 7 bp repeats was flanked by nonpalindromic sequences (Figure 3). The frequency of PMS in this strain was similar to that observed previously (Table 1). This result indicates that the sequences in the *Sal*I site do not contribute to the hairpin, presumably because these sequences are base-paired between the two homologous chromosomes.

Several further points should be discussed here. First, although the insertions used here represent perfect palindromes, for stereochemical reasons, intrastrand base-pairing does not extend to the tip of the hairpin. Several studies indicate that the minimal size of the central loop is 2–4 bp (BLOMMERS *et al.* 1989; GOUGH, SULLIVAN and LILLEY 1986; TINOCO, UHLENBECK and LEVINE 1971). Our results, therefore, suggest that the number of paired bases necessary for formation of a stable *in vivo* hairpin is either

TABLE 1

Patterns of aberrant meiotic segregation of different *his4* alleles

Strain (genotype)	Total no. of tetrads	Aberrant segregation (% of total)	PMS (% of aberrants)[a]
DNY11 (*his4-lop/HIS4*)[b]	312	35	72
DNY62 (*his4-IR11/HIS4*)	357	34	72
DNY48 (*his4-IR9/HIS4*)	379	37	78
DNY50 (*his4-IR7/HIS4*)	351	36	66
DNY50 at 25°	347	24	54
DNY67 (*his4-IR7A/his4-IR7B*)	336	23	63
DNY60 (*his4-IR6/HIS4*)	354	20	3
DNY54 (*his4-IR5/HIS4*)	362	23	10
DNY56 (*his4-IR4/HIS4*)	373	24	0
MW3 (*his4-Sal/HIS4*)[b]	291	31	0

The genotypes at the *HIS4* locus are indicated in parentheses. Any tetrad with a departure from 2:2 segregation on these plates was scored as an aberrant segregant and any tetrad with a sectored colony was scored as a PMS tetrad.

[a] Includes any tetrads with one or more sectored colonies (5:3, 3:5, aberrant 4:4, etc.).

[b] Data from NAG, WHITE and PETES (1989).

5 or 6. Although it might be expected that this number of base pairs would be a function of the temperature, we found no significant difference in the level of PMS for the 7-bp inverted repeat at 18° (the sporulation temperature used routinely in our experiments) and 25° (Table 1).

It is not understood why stem and loop structures are not efficiently recognized by the yeast mismatch repair system. As mentioned previously, certain protein(s) might bind the hairpin structure and prevent the mismatch repair system from removing the mismatch. This binding may serve to protect the palindromic sequences present at functionally important sites when they are fortuitously extruded into the cruciform structure. In addition, many cellular proteins recognize secondary structures in RNA. It is conceivable that a sub-set of these or equivalent proteins also bind DNA hairpins.

The possibility of such a protein suggests an alternative interpretation of our observation that the 6 bp inverted repeat has a low level of PMS. If this protein can bind to the hairpin structure formed by the 7-bp repeat, but not that formed by the 6 bp repeat, our observed low level of PMS could reflect the size of the stem-loop structure rather than its stability. This possibility could be tested by comparing the effects of 7-bp AT-rich palindromes with 6 bp GC-rich inverted repeats. Although there is no published procedure for calculating the expected stability of hairpin structures in heteroduplexes, we estimate (using equations developed for conventional oligonucleotide duplexes) that a 5-bp GAAAA/CTTTT stem would be about 3 kcal/mol less stable than a 4-bp GGGG/CCCC stem (BRESLAUER *et al.* 1986). Thus, if stability of the hairpin was the principle factor in determining the level of PMS, the short GC-rich insertion should have

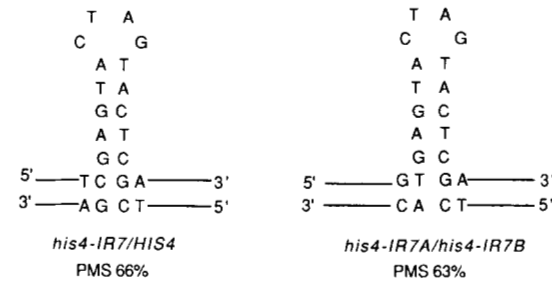

FIGURE 3.—Expected hairpin configuration for palindromes with different flanking sequences. Although the looped-out sequences in these two insertions are identical, the IR7 insertion is flanked by palindromic sequences, whereas the insertion in IR7A is not. Thus, the stem in the puncture involving IR7 could, in theory, be extended by two additional base pairs, disrupting interstrand base-pairing. The observation that the level of PMS is the same for the two configurations, however, indicates that this extended base-pairing probably does not occur.

a higher level of PMS than the longer AT-rich insertion.

Five- and six-base-pair inverted repeats showed 10 and 3% PMS, respectively, among the aberrant segregation events (Table 1 and Figure 2). This result suggests either that hairpin structures occasionally form for these smaller repeats or that the mismatch repair system repairs these loops inefficiently. The latter explanation seems more likely to us, since one nonpalindromic loop has been previously observed to be inefficiently corrected (WHITE *et al.* 1988).

In summary, we show that heterozygous 14-bp, but not 12-bp, palindromic insertions have high levels of PMS. We interpret this result as indicating that 7-bp, but not 6-bp, inverted repeats in single-stranded DNA are sufficient for formation of a hairpin secondary structure *in vivo*.

We thank T. DONAHUE and ANN STAPLETON for yeast strains, S. L. HODGE and J. SIEBER for technical assistance, and members of the Petes laboratory and Doug Berg for comments on the manuscript. This research was supported by the National Institutes of Health grant GM24110.

LITERATURE CITED

BLOMMERS, M. J. J., J. A. L. I. WALTERS, C. A. G. HAASNOOT, J. M. A. AELEN, G. A. VAN DER MAREL, J. H. VAN BLOOM and C. W. HILBERS, 1989 Effects of base sequence on the loop folding in DNA hairpins. Biochemistry **28:** 7491–7498.

BOEKE, J. D., F. LACROUTE and G. R. FINK, 1984 A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast. Mol. Gen. Genet. **197:** 345–346.

BRESLAUER, K. J., R. FRANK, H. BLOCKER and L. A. MARKEY, 1986 Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA **83:** 3746–3750.

CECH, T. R., and M. L. PARDUE, 1976 Electron microscopy of DNA crosslinked with trimethylpsoralen: test of the secondary structure of eukaryotic inverted repeat sequences. Proc. Natl. Acad. Sci. USA **73:** 2644–2648.

CHALKER, A. F., D. R. F. LEACH and R. G. LLOYD, 1988 *Escherichia coli sbcC* mutants permit stable propagation of DNA replicons containing a long palindrome. Gene **71:** 201–205.

COLLINS, J., 1981 Instability of palindromic DNA in *Escherichia coli*. Cold Spring Harbor Symp. Quant. Biol. **45:** 409–416.

DETLOFF, P., J. SIEBER and T. D. PETES, 1991 Repair of specific base pair mismatches formed during meiotic recombination of the yeast *Saccharomyces cerevisiae*. Mol. Cell. Biol. **11:** 737–745.

FOGEL, S., R. K. MORTIMER and K. LUSNAK, 1981 pp. 289–339 in *The Molecular Biology of the Yeast Scaccharomyces cerevisiae*, Vol. I, edited by J. STRATHERN, E. W. JONES and J. BROACH. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

GELLERT, M., M. H. O'DEA and K. MIZUUCHI, 1983 Slow cruciform transitions in palindromic DNA. Proc. Natl. Acad. Sci. USA **80:** 5545–5549.

GELLERT, M., K. MIZUUCHI, M. H. O'DEA, H. OHMORI and J. TOMIZAWA, 1979 DNA gyrase and DNA supercoiling. Cold Spring Harbor Symp. Quant. Biol. **49:** 35–40.

GIERER, A., 1966 Model for DNA and protein interactions and the function of the operator. Nature **212:** 1480–1481.

GOUGH, G. W., K. M. SULLIVAN and D. M. J. LILLEY, 1986 The structure of cruciforms in supercoiled DNA: probing the single-stranded character of nucleotide bases with bisulphite. EMBO J. **5:** 191–196.

HOLLIDAY, R., 1964 A mechanism of gene conversion in fungi. Genet. Res. **5:** 282–304.

KEESEY, J., R. BIGELIS and G. R. Fink, 1979 The product of the *HIS4* gene cluster in *Sacharomyces cerevisiae*. J. Biol. Chem. **254:** 7427–7233.

LEACH, D. R. F., and F. W. STAHL, 1983 Viability of phages carrying a perfect palindrome in the absence of recombination nucleases. Nature **305:** 448–451.

LILLEY, D. M. J., 1980 The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. Proc. Natl. Acad. Sci. USA **77:** 6468–6472.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MESELSON, M., and C. RADDING, 1975 A general model for recombination. Proc. Natl. Acad. Sci. USA **72:** 358–361.

NAG, D. K., M. W. WHITE and T. D. PETES, 1989 Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. Nature **340:** 318–320.

RIPLEY, L. S., and B. W. GLICKMAN, 1983 Unique self-complementarity of palindromic sequences provides DNA structural intermediates for mutation. Cold Spring Harbor Symp. Quant. Biol. **47:** 851–861.

SCHERER, S., and R. W. DAVIS, 1979 Replacement of chromosomes segments with altered DNA sequences constructed *in vitro*. Proc. Natl. Acad. Sci. USA **76:** 4951–4955.

SHERMAN, F., G. R. FINK and J. B. HICKS, 1983 *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SINDEN, R. R., S. S. BROYLES and D. E. PETITJOHN, 1983 Perfect palindromic *lac* operator DNA sequence exists as a stable cruciform structure in supercoiled DNA. Proc. Natl. Acad. Sci. USA **80:** 1797–1801.

STAPLETON, A., and T. D. PETES, 1991 The Tn*3* β-lactamase gene acts as a hotspot for meiotic recombination in yeast. Genetics **127:** 39–51.

SYMINGTON, L. S., and T. D. PETES, 1988 Expansions and contractions of the genetic map relative to the physical map of yeast chromosome III. Mol. Cell. Biol. **8:** 595–604.

TINOCO, I., C. C. UHLENBECK and M. D. LEVINE, 1971 Estimation of secondary structure in ribonucleic acids. Nature **230:** 362–370.

WESTON-HAFER, K., and D. E. BERG, 1991 Limits to the role of palindromy in deletion formation. J. Bacteriol. **173:** 315–318.

WHITE, J. H., F. DIMARTINO, R. W. ANDERSON, K. LUSNAK, D. HILBERT and S. FOGEL, 1988 A DNA sequence conferring high postmeiotic segregation frequency to heterozygous deletions in *Saccharomyces cerevisiae* is related to sequences associated with eucaryotic recombination hotspots. Mol. Cell. Biol. **8:** 1253–1258.

Communicating editor: D. E. BERG