J Mol Evol (1987) 25:241–247

Journal of Molecular Evolution

Sequence Instability in the Long Terminal Repeats of Avian Spleen Necrosis Virus and Reticuloendotheliosis Virus

Gur Notani[1] and Walter Sauerbier

Institute of Human Genetics and Department of Microbiology, 230 Biological Sciences Center, University of Minnesota, 1445 Gortner Avenue, St. Paul, Minnesota 55108, USA

Summary. Sequence divergence between the 3′ long terminal repeats (LTR) of avian reticuloendotheliosis virus (REV), deletion variant proviral clone 2-20-4, and spleen necrosis virus (SNV)—proviral clones 14-44, 60, and 70—was found to involve two classes of base substitutions: low-frequency interspersed and high-frequency clustered substitutions. Clones 2-20-4 and 14-44 have diverged 4.4% owing to low-frequency substitutions. In contrast, two high-frequency substitution segments have diverged by 30% and 29%, respectively. Clustered substitutions appear to be located either within or next to tandem repeats, suggesting their introduction concomitant with sequence deletions and duplications commonly associated with such repeats. A new 19-bp tandem repeat is found in clone 2-20-4. Its sequence could have evolved from the 26-bp repeats found in the SNV clones.

Key words: LTR — Sequence divergence — Interspersed and clustered base substitutions

Introduction

The U3 region of long terminal repeats (LTR) of avian retrovirus can tolerate extensive sequence variations without noticeable effects on viral infectivity (Shimotohno et al. 1980; Shimotohno and Temin 1982). In order to deduce mechanisms and rules that govern sequence rearrangements, we compare LTRs of reticuloendotheliosis virus (REV) and spleen necrosis virus (SNV), which are two homologous viruses, as determined by RNA fingerprint (Kang and Temin 1973) and heteroduplex mapping analyses (Rice et al. 1982). SNV has been shown previously to undergo size variations in the LTR involving DNA duplications and deletions (Shimotohno and Temin 1982). These size variations were localized in the U3 region of the LTR to the 5′ side of the two transcription control sequences CCAAT and TATAAG.

Molecular clone 2-20-4 is a deletion variant of REV-T that was shown to be autonomously replicating (Chen et al. 1981). We have sequenced the U3 and adjoining regions of this clone and compared it with the homologous sequences of SNV clones 14-44, 60, and 70 (Shimotohno et al. 1980). A new 19-bp tandem repeat was found in REV clone 2-20-4 that appears to be derived from a 26-bp tandem repeat present in SNV proviral clone 70. Adjacent to the 19-bp tandem repeat, a 44-bp sequence has accumulated a high number of base substitution mutations. Mechanisms for generating sequence variations in the U3 region and ancestral relationships between SNV and REV will be discussed.

Materials and Methods

Provirus Clone. Recombinant lambda bacteriophage Charon 4A clone 2-20-4, containing the deletion variant of REV-T, was provided by Howard Temin. The map of restriction enzyme cleavage sites (Chen et al. 1981) and the sequencing strategy are outlined in Fig. 1.

Offprint requests to: W. Sauerbier

[1] *Current address:* Department of Biochemistry, University of New Mexico Medical School, Albuquerque, New Mexico 87131, USA

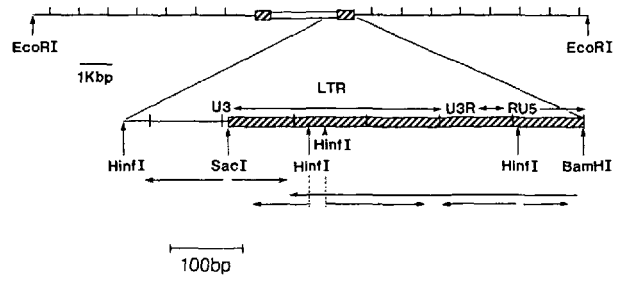


Fig. 1. Restriction cleavage map of clone 2-20-4. The map of restriction cleavage sites is according to Chen et al. (1981). Viral sequences are indicated by double horizontal lines. The LTRs are indicated by hatched boxes. The 3′ LTR and the adjoining viral DNA from the BamHI site to the HinfI site are shown expanded in the lower frame. The U3, R, and U5 regions in the LTR are indicated by outward arrows (U3 is the sequence from the 3′ end of the viral RNA, U5 is the sequence from the 5′ end of the viral RNA between R and the tRNA primer binding site, which is also present at the 3′ end of the viral DNA). Restriction sites utilized for DNA sequencing and sequencing directions are shown by arrows

Subcloning of the 3′ LTR and the Adjoining Viral DNA of Clone 2-20-4. Clone 2-20-4 was digested with EcoRI. The 1065-bp fragment containing the 3′ LTR was purified and inserted at the EcoRI site of plasmid pUC9. The ligated DNA was used to transform *Escherichia coli* JM103 (Messing et al. 1981). Transformants were identified on nutrient agar plates containing dibromodichloro-indolyl-D-galactoside and isopropylthiogalactoside. Plasmids were isolated according to Holmes and Quigley (1981) and purified by CsCl–ethidium bromide density gradient centrifugation (Currier and Nester 1976).

DNA Sequencing. DNA sequencing was carried out with 3′ and 5′ ^{32}P-end-labeled fragments (Maxam and Gilbert 1980) from the HinfI, SacI, and BamHI sites as indicated in Fig. 1.

Enzymes and Radioisotopes. T4 polynucleotide kinase, bacterial alkaline phosphatase, *E. coli* DNA polymerase Klenow fragment, T4 DNA ligase, and restriction endonucleases BamHI, HinfI, and EcoRI were purchased from Bethesda Research Laboratories, Maryland, and used according to vendor specifications. α-^{32}P-dATP was purchased from New England Nuclear, Boston, Massachusetts, and γ-^{32}P-ATP from ICN Pharmaceuticals, Irvine, California.

Results

DNA Sequence Comparisons of the LTR and the Adjoining Viral DNA of Clones 2-20-4 and 14-44

The nucleotide sequence of a proviral clone 2-20-4 segment, extending from the BamHI site in U5 of the 3′ LTR through R and U3 into the adjoining viral sequence up to the HinfI site (see Fig. 1), was determined and compared with the published sequence (Shimotohno et al. 1980) of clone 14-44 (Fig. 2). Comparison of the two segments shows a size variability between nucleotides (nt) 76 and 182 and primary sequence variability between nt 183 and 226. Although the regions between nt 183 and 226 appear to be structurally homologous in clones 2-20-4 and 14-44 (Fig. 2), 13 of the 45 bases (29%) differ in the two clones. A second cluster of base substitutions is located at nt −17 to −39 (30% divergence). The remainder of the sequences has diverged only 3.7% with respect to base substitution mutations. Putative transcription and replication control sequences also display a high degree of homology with regard to sequence and position. By sequencing the two ends of the SNV viral RNA, Shimotohno et al. (1980) identified the transcription promotion and termination regions, respectively. They observed the promoter elements CCAAT (Fig. 2, nt 320 to 324) and TATAAG (nt 368 to 373), and poly-A addition signal AATAAA, and the putative transcription termination signal TTTT (nt 448 to 453 and 494 to 497, respectively). We found identical sequences in clone 2-20-4. They were displaced by only one or two nucleotides due to short deletions (Fig. 2). The polypurine sequence GGGGAA (nt −5 to 0) flanking the 3′ LTR has been suggested to play a role in the initiation of (+) strand DNA synthesis during retroviral reverse transcription (Swanstrom et al. 1981); the same sequence is present in clone 2-20-4. Shimotohno et al. (1980) observed another set of sequences that could have been considered for some regulatory role in viral expression because of their repeat and symmetrical arrangements: an imperfect 7-bp tandem direct repeat at position 349, a 31-bp sequence with a dyad symmetry at position 354, and two symmetrical sequences of 7 bp at position 410 (see Fig. 2). Interestingly, clone 2-20-4 had partial deletions in two of these sequences (Fig. 2).

DNA Sequence Comparisons of the U3 Regions of SNV and REV Clones

Segments of the U3 regions of SNV clones 60, 70, and REV-T deletion variant clone 2-20-4 are compared in Fig. 3. A new 19-bp tandem repeat is observed in clone 2-20-4 in place of the 26-bp tandem repeat in clones 60 and 70. This 19-bp tandem repeat has probably been derived from the 26-bp tandem repeat by the steps outlined in Fig. 4. These involve an A:T to T:A base-pair substitution at nt 170, deletion of a 7-bp sequence from the 26-bp tandem repeat (nt 171 to 177), and duplication of the rearranged 19-bp sequence to obtain the 19-bp tandem repeat of clone 2-20-4. The intervening sequences in the 26- and 19-bp tandem repeats are flanked by TTCTCG and ATCA repeats, respectively. Sequence ATTTC (nt 146 to 150, clone 2-20-4) appears to be the remnant of the 5′ end of the 26-bp tandem repeat. Between nt 150 and 5′ ATCA flank, deletions and insertions have further altered the 2-20-4 clone compared to its progenitor

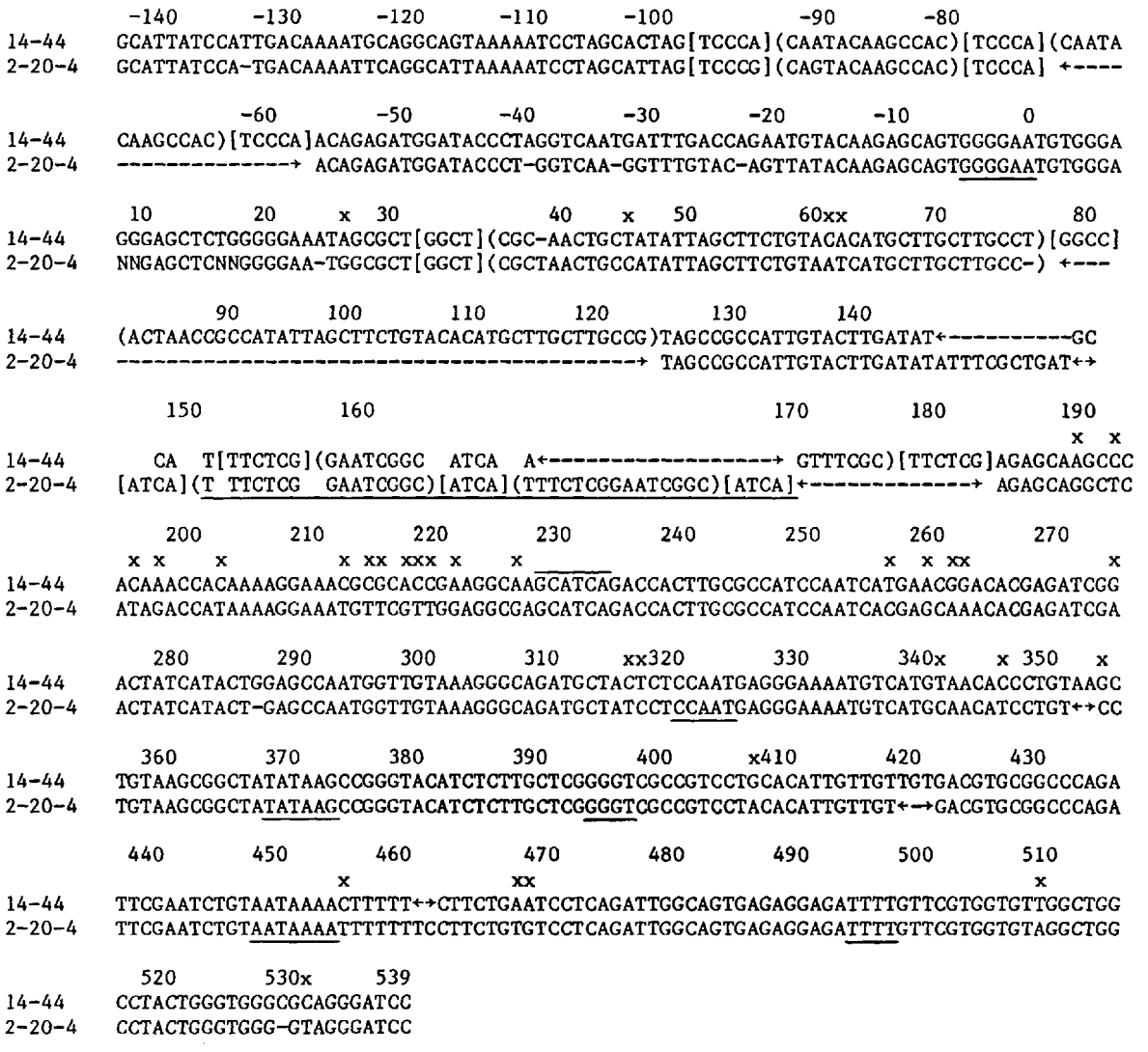


Fig. 2. Nucleotide sequences of the 3′ LTR and the adjoining viral DNAs of clones 14-44 and 2-20-4. Sequences of the 3′ LTR and the adjoining viral DNA of clone 2-20-4 were determined and compared with the published (Shimotohno et al. 1980) homologous sequence of SNV clone 14-44. Numbering is from the beginning of the terminal repeats published for clone 14-44. Substituted bases are marked by X. Flanking and intervening sequences of repeats are indicated by brackets and parentheses, respectively. Possible regulatory sequences are underlined. A 44-bp sequence from nt 183 to 226 has 13 base substitutions between clones 14-44 and 2-20-4. Regions of size variations extend from nt 76 to 182

clone 70. Tandem repeat and flanking sequences can either be exact duplications or show divergence by deletions and substitutions (Figs. 2–4). For instance, the 18-bp tandem repeat of clone 14-44 has identical sequences, whereas the 46-bp tandem repeat has seven base-pair substitutions (Shimotohno and Temin 1982). Sequences flanking the 5′ intervening sequences in the 46-bp tandem repeat have also changed by a single base-pair substitution. The 26-bp tandem repeat of clone 70 has a single base-pair mismatch (Shimotohno and Temin 1982) and the 19-bp tandem repeat of clone 2-20-4 has identical duplication sequences.

Only the 5′ repeat element of the 46-bp tandem repeat of clone 60 is represented in clone 70 (Shimotohno and Temin 1982); the 3′ repeat unit is missing (see Fig. 3, nt 32 to 123). Clone 2-20-4 also contains only the 5′ repeat element (Fig. 3).

It is noteworthy that the flanking sequences of all four tandem repeats (Figs. 2 and 3) have different base compositions.

Discussion

Comparison between the 3′ LTR and the adjoining DNA of proviral clone 2-20-4 and the homologous regions of infectious SNV proviral clones 60, 70, and 14-44 (Shimotohno and Temin 1982) shows a striking sequence divergence due to DNA rearrangements, but little divergence due to base substitutions. Of 684 bp, 43 (6.3%) are substituted. Inter-

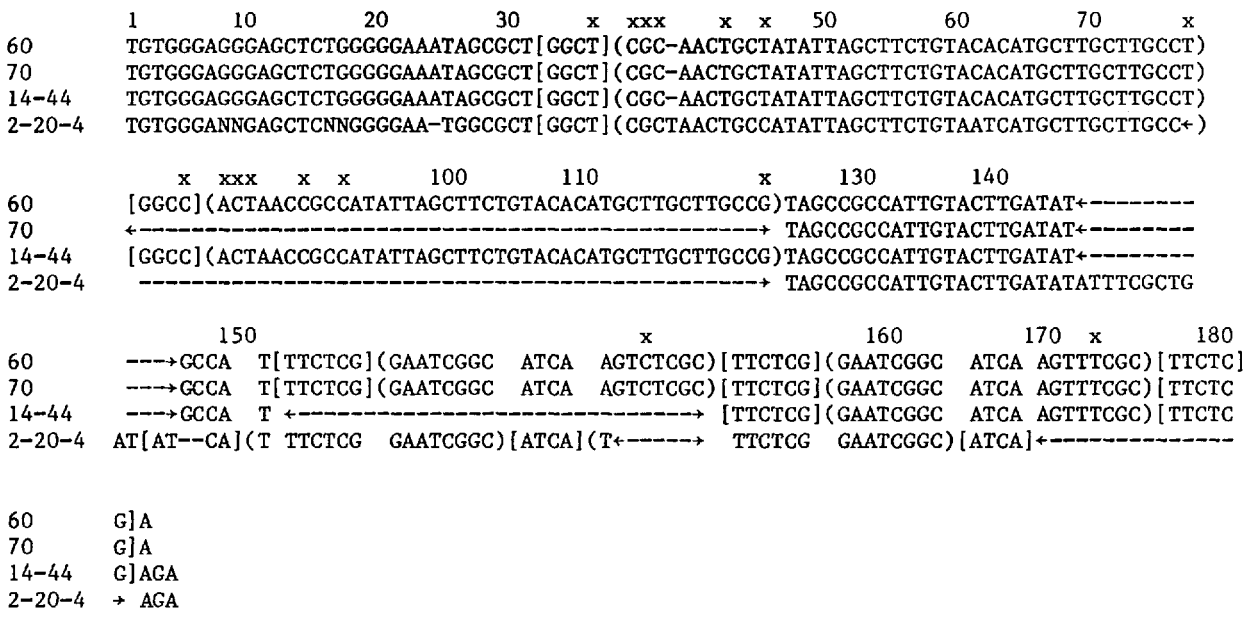


Fig. 3. Comparison of U3 sequences. The DNA sequences of clones 60, 70, and 14-44 are according to Shimotohno and Temin (1981). Repeat sequences are indicated as in Fig. 2. Sequence changes between elements of the same repeat are marked by X

estingly, 13 of the substitutions are clustered in a 45-bp region 3′-adjacent to the 19-bp direct repeat of clone 2-20-4 (29% divergence). A second cluster of seven base substitutions is located 3′-adjacent to the 18-bp direct repeat of clone 14-44 (nt −17 to −39; 30% divergence). Thus, only 23 substitutions are interspersed among 619 bp (3.7%).

Base Substitution Mutations

The rate of base substitutions in retroviral genomes is several orders of magnitude higher than in comparable segments of eukaryotic genomes (Holland et al. 1980; Gojobori and Yokoyama 1984). Nonetheless, base substitution frequencies are only slightly higher in noncoding regions of retroviral genomes than in coding regions (Graur 1985), suggesting only weak purifying selection at the amino acid level. Furthermore, no preference for base substitutions at CG dinucleotides, the sites where methylation can lead to cytosine deamination and mispairing (Coulondre et al. 1978; Razin and Riggs 1980), over CG base pairs was noted (Graur 1985). From these observations and from the low fidelity of reverse transcriptase (Battula and Loeb 1974), it can be argued that most base substitution mutations arise during reverse transcription (Graur 1985). It is surprising, therefore, when base substitution mutations are found clustered in short segments of the LTR U3 regions (Figs. 2 and 3).

Mutations could accumulate in clusters either from a lack of selective pressure on nonessential DNA regions, or by site-dependent high rates of mutation. While we cannot rule out the former, we suggest that several considerations favor the latter interpretation. We hypothesize that the events leading to clustered base substitutions and DNA rearrangements in the U3 region could be by-products of *cis*-acting mechanisms of transcription modulation (vide infra) and that recurring repair replication and not reverse transcription causes the accumulation of base substitutions within and adjacent to repeats.

DNA Rearrangements

The 19-bp repeat of clone 2-20-4 illustrates both the plasticity imposed on sequences upstream of LTR promoters and DNA structure rules that govern the rearrangement processes. In conjunction with the identical arrangement of a 46-bp repeat, it allows one to trace an ancestral relationship between clones 70 and 2-20-4 through the specific steps that led to the derivation of the 19-bp repeat. The sequences to be discussed are aligned in Fig. 3. They are the U3 regions of: SNV clone 60, which has been suggested to be the parent to clones 70 and 14-44 (Shimotohno and Temin 1982) and which has 46- and 26-bp direct repeats; clone 14-44, which also has the 46-bp repeat but only one element of the 26-bp repeat; clone 70, which in contrast to 14-44 has only one element of the 46-bp repeat but the complete 26-bp direct repeat; and REV clone 2-20-4, which has only one element of the 46-bp repeat and which varies in addition by having a new 19-bp direct repeat sequence in place of the 26-bp repeat. From this comparison it is obvious that U3 is a region of unique sequence instability (see also O'Rear and Temin 1981; van Beveren et al. 1982; Hughes 1982).

The sequence organization of the 26-bp repeat of

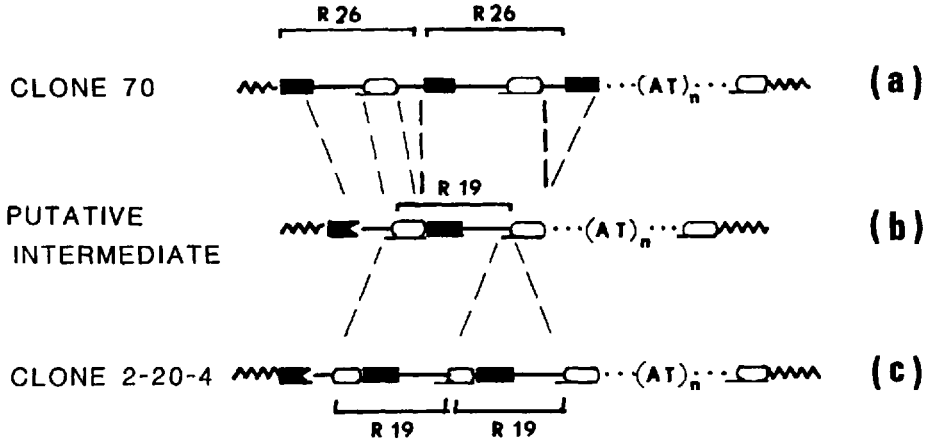
CLONE 70

```
140       150         160         170         180         190         200
   GATATGCCAT[TTCTCG]GAATCGGC(ATCA)AGTCTCGC[TTCTCG]GAATCGGC(ATCA)AGTTT
                                      -t-
           210         220         230         240         250         260
  CGC[TTCTCG]Agagcaagcccacaaaccacaaaaggaaacgcgcaccgaaggcaagc(atca)gacc
                    g--t--t-g----t---------t-tt-gtt-g----g(Clone 2-20-4)
```

CLONE 2-20-4

```
140       150         160         170         180         190         200
  TGATATatttcGCTGAT(ATCA)TTTCTCGGAATCGGC(ATCA)TTTCTCGGAATCGGC(ATCA)AG

           210         220         230         240         250         260
  AGCAGGCTCATAGACCATAAAAGGAAATGTTCGTTGGAGGCGAGC(ATCA)GACCACTTGCGC
```



KEY:
- sequences outside repeat units;
- short flanking sequences of the 26 bp repeat of clone 70;
- short flanking sequences of the 19 bp repeats of clone 2-20-4;
- remnant of short flanking sequence of the 26 bp repeat;
- ..... region of clustered base substitutions;

Fig. 4. Sequence changes proposed for generating the 19-bp tandem repeat in clone 2-20-4 from the 26-bp tandem repeat in clone 70. The sequence printed in lowercase letters is from clone 14-44. Lowercase letters spaced by hyphens represent clustered sequence differences between clones 14-44 and 2-20-4. The solid line indicates AT-rich stretches. Short flanking sequences [TTCTCG] of clone 70 repeats, and the flanking sequences (ATCA) of the new 19-bp repeat, are boxed in. The sequence atttc, printed in lowercase letters in clone 2-20-4, is likely to be the remnant of clone 70 nt 150 to 154. Diagram lines a through c show the DNA organization around the clone 70 parental repeats, R26, the putative intermediate with a single R19 repeat element, and the duplicated R19 of clone 2-20-4. R19 is derived from clone 70 by an A:T to T:A base-pair substitution at clone 70 nt 170 and a partial deletion involving GTCTCGC (nt 171 to 177). The new GC[ATCA] flanking sequences were apparently instrumental in demarcating the boundaries for the duplication of R19. Further deletions of GCC and TCGGAATCGGC and insertion of GCTGAT 5′ to R19 have modified the sequences upstream from R19 of clone 2-20-4. A fourth GC[ATCA] is located 3′ to the AT-rich sequences and, interestingly, marks the end of a DNA segment with clustered base substitutions

clone 70 and the 19-bp repeat of clone 2-20-4 is shown in Fig. 4. Nucleotide numbering is that of clone 14-44 in Fig. 2. Sequences printed in lowercase letters are from clone 14-44. Base substitutions separating sequences of clones 70/14-44 and clone 2-20-4 are printed in lowercase letters and spaced by hyphens. Line a of the diagram (Fig. 4) indicates sequences we suggest to be instrumental in triggering deletions, duplications, and clustering of base substitutions involved in the genesis of the U3 region of clone 2-20-4. They are the short flanking sequences of the 26-bp repeat, R26 (filled rectangles);

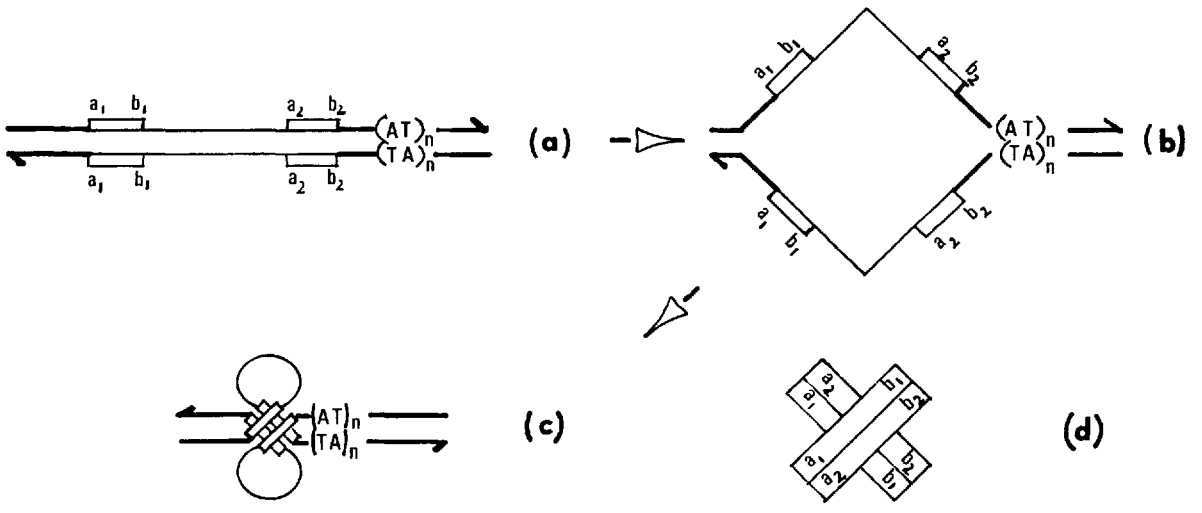


Fig. 5. Model of staggered base pairing and loop formation involving two spaced short repeat sequences. Boxes indicate short repeat sequences; letters a and b mark their direction; (AT)n represents an AT-rich sequence. **a** Configuration of native DNA with two short repeat sequences and an adjoining AT-rich segment. **b** Unstable, denatured intermediate, formed under negative torsional stress. The AT-rich region, which has served as a melt-out nucleus, has renatured while torsional stress has been released. **c** Torsionally relaxed, stable form with staggered base pairing involving the short repeats and unpaired loops. **d** Enlarged view of staggered base pairing at the center of form **c**

the sequence *GC*(ATCA), which is to become the new flanking sequence of the 19-bp repeat, R19 (open ellipsoid); and the AT-rich sequence around nucleotide 230 of Fig. 4. The following steps, which may have occurred during one or more rearrangement events, convert R26 to R19: a base substitution from A to T at position 170 and the deletions indicated by broken-lined wedges between lines **a** and **b** of the diagram. The resulting intermediate would have only one element of the 19-bp repeat (line **b**). A duplication of R19 would generate the sequence organization found in clone 2-20-4 (line **c**). Adjoining 5′ to the R19 is the sequence ATTTC (printed in lowercase letters in Fig. 4), which appears to be the remnant of the 5′ flanking sequence of R26, AT[TTCTCG]. Between this remnant and the ATCA flanking sequence, position 156 to 165, sequences including the *GC* of the flanking sequence have been deleted and replaced by an unrelated stretch of 6 bp, GCTGAT. Adjoining R19 at the 3′ side is an AT-rich segment and another copy of the short flanking sequence, *GC*(ATCA) (at position 243). Within and around this AT-rich segment and, interestingly, bounded by *GC*(ATCA), we find a high incidence of base substitutions (13 out of 45, or 29%) when comparing clones 14-44 and 2-20-4 (see Fig. 2). In contrast, the 19-bp repeat sequences that are also bracketed by the same short repeat *GC*(ATCA) have not accumulated any base substitutions.

Inspection of the −17 to −39 base substitution cluster shows a similar motif: 3′-adjacent to a direct repeat; 3′-bounded by a hexanucleotide (TACAAG; −12 to −17) that is repeated within the 18-bp direct repeats of clone 14-44; and an AT-rich segment (66% AT; −110 to −131) 11 bp 5′ to the direct repeat.

Sequence Instability at U3 and Mechanisms of DNA Rearrangement

In considering causes for the localized sequence instability at U3, we are guided by our studies on localized denaturation of AT-rich sequences and the formation of hairpin loops in negatively supercoiled plasmids. We observe that the immediate response to heat and torsional stress is a localized melt-out of AT-rich sequences. Upon further incubation at 37°C, most of these plasmids with AT-region melt-outs convert to a form with a stem loop at an inverted repeat adjoining the AT segment and a renatured AT region (Wang and Sauerbier, unpublished). The latter appears to be an energetically more favorable form. Chromatin, like plasmids, is torsionally constrained and negatively supercoiled in its transcriptionally active form (Ryoji and Worcel 1984, 1985; Rodi and Sauerbier, unpublished). We therefore suggest that during transcription activation, which is likely to involve an underwinding of DNA, AT-rich sequences located upstream from promoters melt out under torsional stress with or without the assistance of DNA-binding regulatory proteins. In analogy to the "migration" of the localized melt-out from the AT-rich loop to the actually GC-rich stem loop observed with a plasmid DNA (Wang and Sauerbier, unpublished), the AT melt-out in the U3 region would expand or migrate into the adjoining direct repeat segment. Once a significant length of DNA has become single stranded, the molecule is no longer under rigid torsional constraint and is enabled to accommodate some change in writhing number. Short sequence duplications within the single-strand loops, if sufficiently

spaced to accommodate necessary bends, can engage in "staggered" base pairings as indicated in Fig. 5c. This form would be analogous to a hairpin loop and energetically more favorable than the open loop configuration (Fig. 5b). In this context, it is important that the staggered pairings of the short repeats will reinforce each other and stabilize the loops against hydrodynamic and other forces acting on the hydrogen bonds of the short repeats (Fig. 5c). With three or more short flanking sequences [e.g., *GC*(ATCA), as we find in clone 2-20-4, or TACAAG in clone 14-44], more complex forms of staggered base pairing can be envisaged. Departing from structures depicted in Fig. 5, incisions, ligations, and repair DNA synthesis could lead to deletions, duplications, and base substitution errors in ways that are topologically equivalent to those invoked in past "slippage mispairing" models (Farabaugh et al. 1978; Efstratiadis et al. 1980; Yu and Manley 1986).

Acknowledgment. This study was supported by NIH grant 5R01-GM30319 and an American Cancer Society Institutional Grant.

References

Battula N, Loeb LA (1974) The infidelity of avian myeloblastoma virus deoxyribonucleic acid polymerase in polynucleotide replication. J Biol Chem 249:4086–4093

Beveren C van, Rands E, Chattopadhyay SK, Lowy DR, Verma I (1982) Long terminal repeat of murine retroviral DNAs: sequence analysis, host–proviral junctions, and preintegration site. J Virol 41:542–556

Chen ISY, Mak TW, O'Rear TJ, Temin HM (1981) Characterization of reticuloendotheliosis virus strain T DNA and isolation of a novel variant of reticuloendotheliosis virus strain T by molecular cloning. J Virol 40:800–811

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in *Escherichia coli.* Nature 274:775–780

Currier TC, Nester EW (1976) Isolation of covalently closed circular DNA of high molecular weight from bacteria. Anal Biochem 76:431–441

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, Deriel JK, Forget BG, Weisman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human β-globin gene family. Cell 21:653–666

Farabaugh PJ, Schmeissner U, Hofer M, Miller J (1978) Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots in the lac I gene of *Escherichia coli.* J Mol Biol 126:847–863

Gojobori T, Yokoyama S (1984) Rates of nucleotide substitution for cellular and viral oncogenes. Genetics 107:s39

Graur D (1985) Pattern of nucleotide substitution and the extent of purifying selection in retroviruses. J Mol Evol 21:221–231

Holland J, Spindler K, Horodyski F, Grabau E, Nicol S, VandePol S (1982) Rapid evolution of RNA genomes. Science 215: 1577–1585

Holmes DS, Quigley M (1981) A rapid boiling method for the preparation of bacterial plasmids. Anal Biochem 114:193–197

Hughes SH (1982) Sequence of the long terminal repeat and adjacent segments of the endogenous avian virus Rous-associated virus O. J Virol 43:191–200

Kang CY, Temin HM (1973) Lack of sequence homology among RNAs of avian leukosis-sarcoma viruses, reticuloendotheliosis viruses, and chicken endogenous RNA directed DNA polymerase activity. J Virol 12:1314–1324

Maxam A, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. In: Grossman L, Moldave K (eds.) Methods of enzymology, 65. Academic Press, New York, pp 499–560

Messing J, Crea R, Seeburg PH (1981) A system for shotgun DNA sequencing. Nucleic Acids Res 9:309–321

O'Rear JJ, Temin HM (1981) Mapping of alterations in noninfectious proviruses of spleen necrosis virus. J Virol 39:138–149

Razin A, Riggs AD (1980) DNA methylation and gene function. Science 210:604–610

Rice NR, Hiebsch RR, Gonda MA, Bose HR, Gilden RV (1982) Genome of reticuloendotheliosis virus: characterization by use of cloned proviral DNA. J Virol 42:237–252

Ryoji M, Worcel A (1984) Chromatin assembly in *Xenopus* oocytes: in vivo studies. Cell 37:21–32

Ryoji M, Worcel A (1985) Structure of the two distinct types of mini-chromosomes that are assembled on DNA injected in *Xenopus* oocytes. Cell 40:923–932

Shimotohno K, Temin HM (1982) Spontaneous variation and synthesis in the U3 region of the long terminal repeat of an avian retrovirus. J Virol 41:163–171

Shimotohno K, Mizutani S, Temin H (1980) Sequence of retrovirus provirus resembles that of bacterial transposable element. Nature 285:550–554

Swanstrom R, DeLorbe WJ, Bishop JM, Varmus HE (1981) Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc Natl Acad Sci USA 78:124–128

Yu YT, Manley L (1986) Structure and function of the S1 nuclease-sensitive site in the adenovirus late promoter. Cell 45:743–751

Received October 13, 1986/Revised December 30, 1986