Slippage synthesis of simple sequence DNA

Christian Schlötterer* and Diethard Tautz*
Institut für Genetik und Mikrobiologie der Universität München, Maria-Wardstraße 1a, 8000 München 19, FRG

Received November 8, 1991; Revised and Accepted December 23, 1991

ABSTRACT
The analysis of slippage synthesis of simple sequence DNA in vitro sheds some light on the question of how simple sequences arise in vivo. We show that it is possible to synthesize all types of repetitious di- and trinucleotide motifs starting from short primers and a polymerase in vitro. The rate of this synthesis depends on a sequence specific slippage rate, but is independent of the length of the fragments being synthesized. This indicates that only the ends of the DNA fragments are involved in determining this rate and that slippage is accordingly a short range effect. Slippage synthesis occurs also on a fixed template where only one strand is free to move, a situation which resembles chromosome replication in vivo. It seems therefore likely that slippage during replication is the cause of the observed length polymorphism of simple sequence stretches between individuals of a population.

INTRODUCTION
Simple sequences are monotonous repetitions of very short (1–5nt) nucleotide motifs, which occur as interspersed repetitive elements in all eukaryotic genomes (1–4). They show extensive length polymorphisms when analysed between different individuals of the same species (4–9), a fact that can be utilized for DNA-fingerprinting purposes (4–10). It seems likely that this length polymorphism is due to slippage mutations which are produced during replication of DNA (3,11). The frequent occurrence of slippage events at random duplications of short nucleotide motifs may in fact be the mechanism of how simple sequences are generated in the first place (1–3).

It has long been known that simple sequences can be synthesized in vitro using only short primers, DNA-nucleotides and a DNA polymerase (12–15). This synthesis is probably due to slippage of the internally repetitive motifs. We have therefore reinvestigated this phenomenon at a higher resolution then in previous experiments, in order to understand how the observed polymorphisms may arise in vivo.

MATERIALS AND METHODS
Primer
Simple sequence primers were synthesized on a Cyclone DNA-synthesizer. The following primer pairs were used (all in 5½-3½ direction; the total length of the primer is written as a subscript behind the repeated motif): GT15/CA9, AG15/TC9, ATT15/AAT9, GTT15/ACA9, TCT15/AGA9, TCA15/ATG9, TAC15/TA9, TCC15/GGA9, CCA15/GGT9, TCG15/ACG9, TGC15/GCA9 and GCC15/GGC9. When referring to double-stranded sequences in the text, we write the complementary motifs from both strands.

Slippage synthesis
The typical reaction conditions are: 50mM Potassium-phosphate buffer pH 7.5, 15 mM MgCl2, 4mM DTT, 2mM of each dNTP, 5µM of each of the complementary primers and 0.1 U/µl DNA Polymerase I (Boehringer Mannheim). Incubation was at 37°C for up to two hours. These reaction conditions are optimized with respect to the nucleotide (15) and polymerase concentration. The Klenow enzyme (Boehringer Mannheim) as well as the T4 and T7 DNA polymerases (USB) were used under the same reaction conditions. The synthesis products were partly purified by centrifugation dialysis (Amicon 30) against TE (10mM Tris pH 8.0, 1mM EDTA). For the extension reaction on a fixed template, the primer CA9 was 5½-endlabeled with γ-32P-ATP and hybridized in a molar 1:2 ratio to single stranded M13 template for 15min at 65°C. The extension reaction was done with the Klenow enzyme. The buffer was the same as described above, but contained only 20µM each of dATP and dCTP and no dGTP and dTTP.

RESULTS
For the synthesis of simple sequences in vitro, we used the respective oligo-nucleotides with a length of 15nt and complementary primers with a length of 9nt. Most experiments were done with DNA-polymerase I from E.coli under optimized reaction conditions, though other polymerases were also tested (see below). Figure 1 shows an agarose gel of samples from a time course of such a synthesis reaction for the dinucleotide motif

* Present address: Zoologisches Institut der Universität München, Luisenstraße 14, 8000 München 2, FRG
CA/TG and the tri-nucleotidet motif GGA/TCC. It is evident that the fragments grow continuously with a constant rate until the reaction components are exhausted. The length variability of the fragments generated in the reaction is very small at any given time point, indicating that they all grow at a similar rate.

The final reaction product can be purified and used as a template for a further synthesis. The fragments start to grow again at a similar rate as in the first reaction (Figure 2). This isolation-synthesis cycle can be repeated once before apparent secondary reaction products (15) start to accumulate (Figure 2). This experiment shows that the synthesis is independent of the length of the primer fragments and depends therefore only on slippage occurring at the ends of these fragments.

Simple sequences with a tri-nucleotide repeat unit behave in a similar way as the dinucleotide repeats (Figure 1). It is, however, evident that the growth of the tri-nucleotide repeat products is slower which suggests that the slippage rate is lower for these repeats. We have tested all possible combinations of tri-nucleotide repeat motifs in a similar assay and found that they all grow at different rates (Figure 3; Table 1). AAT/ATT repeat units grow fastest and GCC/GGC slowest, indicating that slippage is potentially dependent on the AT-content of the sequences involved. However, specific sequence properties may also influence the rate of slippage, since the repeat units with 66% AT content (ACA/TGT, AGA/TCT, ATG/CAT and TAG/C- TA) have not a consistently higher slippage rate than those with only 33% AT-content (TCC/GGA, ACC/GGT, ACG/GCT and AGC/GCT), though the latter are on average slower (Table 1).

To investigate the slippage-synthesis properties in more detail, we have analysed the early synthesis products on a high resolution denaturing gel (Figure 4). If slippage would be the only rate limiting step, one would have expected that the length distribution of the products reflects the repeat unit length of the synthesized DNA. However, we found that this could be observed only for repeat units which show a low rate of slippage, such as GGA/TCC (Figure 4c), while the others show a somewhat different pattern. GAA/TTC shows marked double bands, while CA/TG shows only bands with a single nucleotide spacing (Figure 4a,b).

This observation and those from the above experiments suggest the following model for the slippage synthesis reaction (Figure 5). The original starter primers anneal and the overlapping ends are filled up by the polymerase. The polymerase is then released and a slippage event occurs in the double stranded DNA, creating new overlapping ends, which can again be filled up by the polymerase. Thus, the reaction is determined by the rate at which slippage occurs and by the rate at which the DNA-polymerase complex is reestablished. The latter effect is responsible for the continuous growth of all fragments at the same time and at the same rate. The pattern of the early reaction products shows that

![Figure 1](image1.png)

**Figure 1.** Time course of a simple sequence synthesis reaction for the dinucleotide repeat unit GT/AC and the trinucleotide repeat unit TCC/GGA. A standard reaction was set up and samples were taken in 15 min intervals. The reaction products were resolved on an 2% agarose gel. The marker bands (length in kb) are indicated to the right.

![Figure 2](image2.png)

**Figure 2.** Simple sequence synthesis with an elongated polymer as template. A standard reaction was set up for the GT/AC polymer and the final product was purified after 3 hours of synthesis by centrifugation dialysis. This reaction product was then used as a template in a new standard reaction. The whole cycle was repeated once. The figure shows samples of time points during the three reaction cycles, resolved on an 2% agarose gel. The marker bands (length in kb) are indicated to the right.

![Figure 3](image3.png)

**Figure 3.** Slippage synthesis products of the different simple sequence polymers after two hours of synthesis under standard reaction conditions, resolved on an 2% agarose gel. The marker bands (length in kb) are indicated to the right.

![Table 1](image4.png)

**Table 1.** Synthesis rates for the different simple sequence polymers. The rates were determined for standard reaction conditions. The length of the final product after 120 min of synthesis is shown, as well as the rate of synthesis, as determined by the analysis of intermediate time points during the linear synthesis phase.
this process is somewhat more complex. It is best explained, if one assumes that DNA-polymerase I does not form very stable complexes with DNA-ends which overlap by only one base-pair. This is inferred from the pattern of the GAA/TTC reaction which can be interpreted as follows: after slippage has occurred by one register unit, the polymerase will form a synthesis complex with this template. Two nucleotides are added before the complex becomes somewhat unstable and the polymerase falls off at a certain rate. The polymerase can of course reestablish the complex again to add the final nucleotide, but it is possible that another slippage event has occurred during this time which would

Figure 4. Early phase of slippage synthesis for three different simple sequence polymers (GT/AC, GAA/TTC and TCC/GGA). Standard synthesis reactions were set up, but with one primer 32P-labeled at its 5½-end. Aliquots were taken in 15 min intervals and resolved on an 8% denaturing acrylamide gel. The visible bands represent single nucleotide spacings, the lowest bands have a length of about 20nt. See text for further details.

Figure 5. Model for the slippage synthesis. The figure depicts the assumed events during the synthesis of an GAA/TTC polymer at one of the two ends. After slippage has occurred, a transient bulge will form, which probably moves through the whole DNA strand. The overhanging 5½-end can then be filled up (nucleotides in italics). As explained in the text, it is assumed that only two nucleotides are filled up, before the polymerase/DNA complex becomes unstable. This allows either that a second slippage event occurs before a new fill-up reaction starts, or that the remaining nucleotide is filled up after the polymerase/DNA complex has been reestablished.

Figure 6. Slippage synthesis with different DNA-polymerases. Standard reactions were set up for the GT/AC and the TCC/GGA polymers with three different polymerases. The reaction proceeded for 2 hours. Aliquots were taken every 15min and resolved on an 2% agarose gel. The marker bands (length in kb) are indicated to the right. It is evident that the T7 polymerase shows no defined size classes of the reaction products, indicating a different mode of synthesis (see text for further details).

Figure 7. Slippage synthesis on a fixed template. The sequence of the template is shown at the bottom of the figure. Aliquots of the reaction were taken in 2 min intervals and resolved on a 8% denaturing acrylamide gel. See text for further details.
again favour the synthesis of the other nucleotides, apart of the last one. This effect would lead to the observed double hands for the GAA/TTC motifs and the single nucleotide spacing for the CA/TG motifs. In the case of the GGA/TCC motifs, the polymerase would have enough time to fill up also the last nucleotide, since the slippage rate is apparently lower than the rate at which the complex is reestablished.

These experiments suggest that the details of the simple sequence DNA synthesis may be influenced by the particular properties of the DNA polymerase used. We have therefore tested other polymerases in the same assay. The Klenow enzyme which is derived from DNA Polymerase I (16), but also the T4 polymerase behave indeed similarly as DNA polymerase I (Figure 6a,b). However, T7 DNA polymerase shows an interesting difference. The fragments grow unevenly when this polymerase is used and the reaction products form a smear on the gel (Figure 6c). Such a behavior is expected, if the DNA/polymerase complex remains more stable after the fill up reaction. In this case, fragments which are already loaded with a polymerase molecule would grow by several slippage cycles, before the polymerase is released and can start on another fragment. This would result in the observed length heterogeneity of the fragments. A similar result was also obtained with Taq-polymerase (not shown).

Slippage synthesis can also be observed, when only one strand is free to slip. We have used a single-stranded M13 template for this experiment which contains a cloned GT simple sequence stretch with a length of 53nt (Figure 7). A complementary CA oligo-nucleotide was end-labeled and was used as a primer. The primer was allowed to anneal to the template and only dCTP and dATP were included as nucleotides for the synthesis. Thus, the extension should proceed only to the end of the simple sequence stretch, since the nucleotide following the stretch is an A residue. The primer can anneal at any point within the GT-stretch and one would expect therefore a random distribution of fragment length after the synthesis, with a peak in the region of half of the size of the simple sequence stretch. This can indeed be observed after a short synthesis time (Figure 7). However, as the reaction proceeds, the fragments become longer and longer, eventually with a main fragment length of 53nt. We interpret this observation as a slippage towards the 5½-end of the already synthesized strands with a concomittant elongation at the 3½-end. Interestingly, there are also some products longer than 53nt, suggesting that part of the 5½-end of the fragments remain unpaired, while there is still an elongation at the 3½-end. This observation is in line with the assumption that transient bulges of unpaired DNA may exist within simple sequence stretches (see Discussion).

**DISCUSSION**

Our results support the notion that the in vitro synthesis of simple sequences proceeds via slippage reactions. These produce contiously unpaired free ends which can be filled up by a DNA-polymerase. Our experiments show that the rate of growth of the DNA fragments is largely independent of the length of the DNA-fragments being synthesized. A length dependence would be expected, if a slippage event would have to be transmitted through the whole stretch of DNA before new nucleotides are added to the ends. The observation of a length independence indicates therefore that transiently unpaired bulges must exist in simple sequence stretches, which move through the DNA strand while it is elongated at the ends. Such bulges can be structurally accommodated in a DNA helix (17).

These observations bear directly on the understanding of the in vivo situation. Simple sequence stretches occur very frequently in genomic DNA and are usually flanked by non-repetitious sequences. Thus, both ends of a simple sequence stretch would normally be fixed and would not be free to slip. However, the above postulated bulges could nonetheless be formed and would create transiently single stranded regions. These regions appear to be a substrate for single strand endonucleases in vitro (18) and could therefore also be a substrate for repair enzymes in vivo. Frequent initiation of a repair process in simple sequence regions would thus contribute to the observed length variability. On the other hand, the loci analysed so far appear to be somatically stable which would argue against a contribution by the repair processes. Furthermore, the analysis of the mutation rate in simple sequence stretches in E.coli showed that the presence of repair enzymes reduces in fact the mutation rate (11).

The creation and elongation of simple sequence stretches occurs therefore probably predominantly during replication. It is unclear though whether slippage synthesis occurs during the replication itself (‘polymerase-stuttering’), or whether free DNA ends are necessary. Our in vitro observations suggest that the latter may be more likely, since the reaction is best explained, if one assumes that the polymerase always falls off before a new slippage event occurs. This may however depend on the characteristics of the polymerase. It seems that the T7 DNA polymerase does not follow this scheme, since it remains apparently bound to the DNA ends while slippage proceeds. On the other hand, our knowledge of the replication process suggests that transiently free DNA strands (or primer RNA strands) may occur frequently, at least on the lagging strand (19). If such a free end would lie within a simple sequence region, it would be possible that slippage can occur before the polymerase complex would be reestablished. The newly synthesized strand would then contain a bulge which could need to be repaired. This repair process could then either recreate the old situation, or result in an elongation of the complementary strand.

The in vitro experiments suggest that slippage occurs frequently and at a high rate within simple sequence stretches. It is however difficult to estimate how often such a slippage event would also result in a length change of the respective region in the DNA, since this depends on the characteristics of the enzyme machinery. In vivo rate estimates for slippage driven mutations in E.coli suggest a rate of about one mutation in 100 replication events (11). Similar experiments have not been performed for eukaryotes so far. However, the fact that simple sequence stretches are on the one hand usually stable during several generations, but are on the other hand hypervariable in length when analysed in unrelated individuals, suggests that the mutation rate in simple sequence regions in eukaryotes may be similar to that measured in prokaryotes.

**REFERENCES**