J. Mol. Biol. (1995) **254**, 856–868

JMB

Modelling Viral Evolution *in Vitro* Using exo⁻ Klenow Polymerase: Continuous Selection of Strand Displacement Amplified DNA that Binds an Oligodeoxynucleotide to Form a Triple-helix

Nils G. Walter

Max-Planck-Institute for Biophysical Chemistry Department of Biochemical Kinetics, Am Fassberg D-37077, Göttingen Germany

Evolution comprises cycles of amplification, mutagenesis and selection. To study evolutionary phenomena, isothermal strand displacement amplification (SDA) of double-stranded DNA as an *in vitro* model for rolling-circle replication of viruses has been coupled to a positive selection procedure. First, two subsequent amplification reactions utilizing exo⁻ Klenow polymerase were performed under direct observation using the fluorescent dye thiazole orange. Under the chosen conditions, the mutation rate was 1.5×10^{-3} and 0.4×10^{-3} for base substitutions and deletions, respectively. Then, a 16mer oligodeoxynucleotide with an acridine moiety coupled to its 5' end was used to select for double strands that retained their ability to form a triple-helix with the oligodeoxynucleotide. Conditions for triple-helix formation were chosen such that only 10 to 40% of the SDA products were allowed to bind the third strand. Non-denaturing polyacrylamide gel electrophoresis was used to separate triple-helices from unmodified double strands, and only triplex strands were used to initiate a new round of error-prone amplification and selection. Nine such rounds with about 270 molecular generations were performed. The final mutant spectrum was characterized and compared with those of amplification reactions without additional selection pressure. While without selection pressure base substitutions and deletions throughout the initial wild-type rapidly produce a diverse mutant distribution, the consensus after nine selection rounds clearly shows two mutational hotspot positions. Using gel shift assays and a newly developed non-radioactive DNase I footprinting technique, it could be shown that both the initial wild-type and the final consensus do not differ significantly in their triplex formation ability. As opposed to this, they do show different amplification efficiencies. The final consensus sequence is amplified with the highest rate in the exponential reaction phase, while the most abundant clone, which is characterized by two additional point deletions, is the sequence with the highest amplification rate in the linear growth phase.

Keywords: DNA virus; quasispecies; rolling-circle replication; selection value; triple-helix

Introduction

Strand displacement amplification (SDA) is a novel isothermal amplification technique for ds-DNA that resembles rolling-circle replication (Walker *et al*., 1992; Walter & Strunk, 1994). It comprises cycles of primer annealing, elongation, nicking and strand displacement (Figure 1). As in rolling-circle replication of, for example, single-stranded DNA viruses and small (<10 kb) bacterial plasmids (Gilbert & Dressler, 1968), the crucial

Present address: c/o Professor John M. Burke, Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, Vermont 05405, USA.

Abbreviations used: c, complementary; ds, double-stranded; SDA, strand displacement amplification; PCR, polymerase chain reaction; dATPαS, 2'-deoxyadenosine-5'-*O*-(1-thiotriphosphate); ACR, 2-methoxy-6-chlor-9-aminoacridine; SSCP, single-strand conformational polymorphism; TMR, tetramethylrhodamine.

0022–2836/95/500856–13 $12.00/0

reaction steps of SDA are cleavage of only one strand of a DNA double strand and displacement of the downstream copy during elongation of the obtained 3' end. By using two primers complementary to both 3' ends of the double-stranded target molecule, the *in vitro* amplification reaction defines the sequence in between to be enriched exponentially under consumption of the primers. After the concentration of target has reached levels in the range of the primer concentrations, the reaction turns into a linear enrichment phase, where elongation and displacement of strands dominates over primer annealing (Walker *et al*., 1992; Walter & Strunk, 1994).

In contrast to the polymerase chain reaction, which requires temperature cycling for annealing and elongation of primers and separation of double strands (Saiki *et al*., 1988), all reaction steps in the isothermal SDA protocol occur simultaneously. As in other isothermal amplification techniques like the Qβ-reaction (Spiegelman, 1971) or the self-sustained sequence replication (Breaker & Joyce, 1994), this gives rise to competition among different target molecules for resources like primers and enzymes. The faster a specific sequence is copied, the higher its proportion in the final mutant distribution will be. This phenomenon has been termed *in vitro* evolution (Mills *et al*., 1967; Eigen, 1971).

In the case of SDA with exo⁻ Klenow polymerase, prolonged exponential amplification of DNA by continuous serial transfer of products into fresh reaction mixture leads to base substitutions and deletion formation and thereby loss of sequence information in the developing quasispecies of replicated molecules (Walter & Strunk, 1994). Similar deletion events have long been known from the *in vivo* replication of single-stranded DNA phages and bacterial plasmids that occurs *via* a rolling-circle mechanism (Horiuchi, 1983; Vilette *et al*., 1992). Thus, SDA resembles rolling-circle replication in its reaction mechanism, and in its evolution products. This led to the question of whether SDA could also be an *in vitro* model for further aspects of evolutionary processes *in vivo*. One question was, whether SDA could be coupled to a positive selection scheme that requires specific sequence information in the double-stranded DNA products.

A basic feature of double-stranded DNA is that specific nucleic acid sequences can be recognized by either proteins or other nucleic acids. For example, ligand binding to genomic sequences plays a major role in the regulation of gene expression. Since oligonucleotides can bind to the major groove of double-helical DNA *via* specific Hoogsteen or reverse Hoogsteen hydrogen bonds and are easily

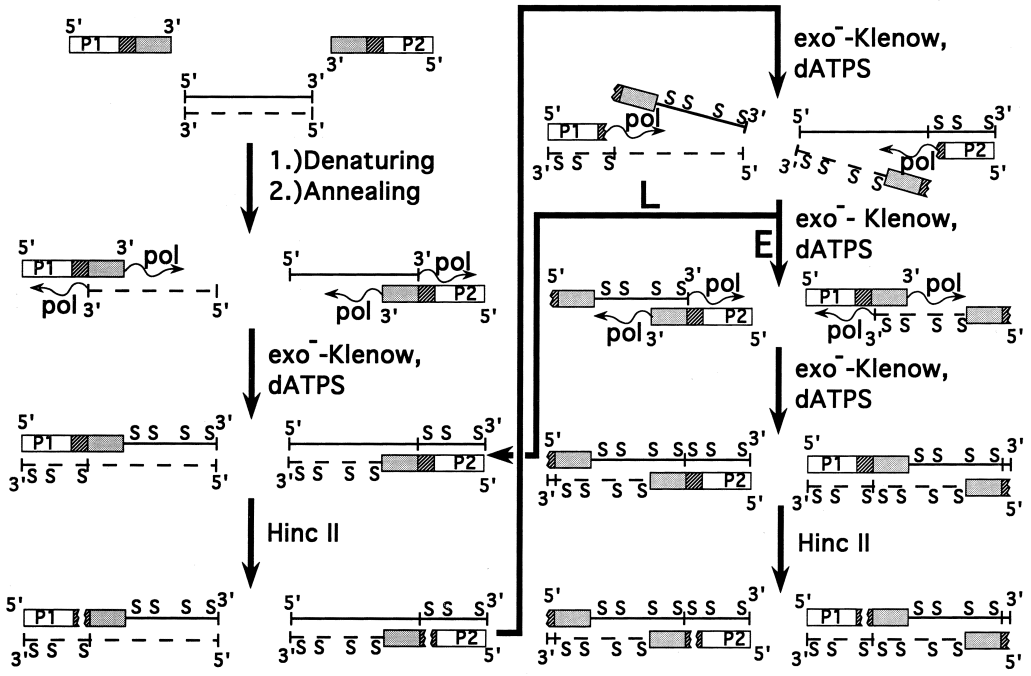

Figure 1. Reaction scheme of strand displacement amplification. Both primers, P1 and P2, contain 18 nt 5' overhangs (open bars), 6 nt of *Hin*cII restriction enzyme recognition sites (hatched bars) and 13 nt of 3' annealing sites (shaded bars). After annealing to sense (continuous line) and antisense (broken line) strands of the target, the priming 3' ends are elongated by exo⁻ Klenow polymerase (pol). The enzyme incorporates dATPαS (S in the strands), creating a hemiphosphorothioate *Hin*cII site, which is subsequently nicked by the endonuclease. The exo⁻ Klenow polymerase then extends the nick, displacing the downstream non-template strand. The displaced strands can hybridize the second primer and be copied likewise, resulting in an exponential enrichment of target sequences (pathway E). Linear amplification (pathway L) occurs with the formerly nicked, now restored double strand, that again contains a hemiphosphorothioate *Hin*cII site and can repeatedly be nicked and copied independent of primer binding.

available by synthesis, they are ideal model ligands to study DNA recognition. The obtained triple-helices are even discussed as a basis for an ''antigene'' approach of artificial gene control (Cooney *et al*., 1988; for a recent review, see Thuong & Hélène, 1993).

In the present work, strand displacement amplification utilizing exo$^-$ Klenow polymerase has been coupled to selection of double strands with ability to bind a 16mer 5'-acridine-modified oligodeoxynucleotide in a triple-helical form. Starting from a defined wild-type sequence, amplification was initiated, followed by incubation with the oligodeoxynucleotide and an electrophoretic mobility shift assay as the selection step for the formed triple-helices. Nine rounds of coupled amplification and selection with altogether 270 molecular generations were performed. Comparison of replicated molecule populations under selection for triple-helix formation with amplified populations without selection show different mutant distributions. The chosen system resembles the *in vivo* situation, where selection for a fast replication of the genome and for distinct features like specific recognition of genomic sequences by ligands are always coupled.

Results

Coupled amplification and selection system

Amplification and mutation under the use of exo$^-$ Klenow polymerase and selection for triple-helices were coupled by alternating two subsequent SDA reactions including a serial transfer and a separate selection step for oligodeoxynucleotide-binding dsDNA products as shown in Figure 2(a). A target molecule named TRIIS1 was used to inoculate the very first amplification reaction (Table 1). The amplification reactions were followed in a fluorometer by addition of the dye thiazole orange, that binds to the dsDNA products under increase of fluorescence intensity (Lee *et al*., 1986; Zhu *et al*., 1994; Walter & Strunk, 1994). Under the conditions chosen, both the exponential (identified in the fluorescence curve by an increasing slope after a lag-time) as well as the linear reaction phase (steep linear slope) of SDA were utilized to generate new molecule generations. Afterwards, the dsDNA products were selected for double strands binding the 16mer ACR-coupled oligodeoxynucleotide TRIPEL1 (Table 1, sequence with the three bases thymine, cytosine and guanine in analogy to Giovannangéli *et al*., 1992) by incubation in binding buffer and a subsequent mobility shift assay in a Mg^{2+}-containing electrophoresis buffer. The initial target molecule TRIIS1 (Table 1) contains a homopurine-tract that can form such a triple-helix with TRIPEL1 *via* Hogsteen base-pairs in its major groove, provided the cytosine residues of TRIPEL1 are protonated (Thuong & Hélène, 1993; Figure 2(b) and (c)). To increase the stringency of selection in the course of the performed nine selection rounds, the pH of the incubation buffer was increased and concentration of the oligodeoxynucleotide was lowered as depicted in the legend to Figure 3. Under all of these conditions, the corresponding oligodeoxynucleotide without the ACR-moiety did not bind to the TRIIS1 sequence (results not shown), for a 5'-acridine strongly stabilizes the triplex (Sun *et al*., 1989; Stonehouse & Fox, 1994).

Figure 3(a) shows a native gel with samples as used to select for triple-helices in each selection round. As expected (Giovannangéli *et al*., 1992; Thuong & Hélène, 1993), the yield of triple-helix decreases with decreasing concentration of oligodeoxynucleotide TRIPEL1 (from selection round 5) and increasing pH (from selection round 6; at a pH of 7.5 or above, triplex-formation could no longer be detected). In every case, the triplex band with lower electrophoretic mobility than the dsDNA was precisely cut out, diffusion eluted and the derived DNA used to initiate a subsequent amplification and selection round. Thus, only about 10 to 40% of the SDA products passed this selection criterion, as judged by band intensities in the ethidium bromide-stained gel.

In each round, the complexity of the product mixture was monitored using single-strand conformational polymorphism (SSCP) analysis (Orita *et al*., 1989; Hayashi & Yandell, 1993), showing that a diverse mutant distribution or quasispecies (Eigen *et al*., 1989) with broadly distributed electrophoretic mobilities of single strands was already obtained in the first four rounds (Figure 3(b)). The same can be seen in Figure 3(a), where the dsDNA bands up to selection round 4 become increasingly diffuse. After changing the stopping procedure for SDA from heating to addition of EDTA in selection round 5, the electrophoretic smears due to heteroduplex formation between different sequences of the quasispecies disappear.

Generation numbers of strand displacement amplified molecules could be estimated under the assumption of roughly exponential growth by determination of initial and final template concentrations. In the first amplification reaction, about 13 molecule generations were passed through (amplification factor 10^4), while all further reactions covered about 16 generations each (amplification factor 6×10^4). This led to about 270 molecular generations after altogether 17 subsequent amplification reactions (in nine selection rounds, one of them without serial transfer).

Sequence analysis of selected SDA products

To obtain appropriate amounts of clonable DNA, the SDA products and the purified triplex DNA, respectively, were amplified by PCR with low cycle number. The PCR product was cloned into pUC18 vector and sequenced with a reverse sequencing primer. Figure 4 shows the sequences of SDA products after one (clones A1S0), two (A2S0) or five (A5S0) subsequent SDA reactions without selection for triplex formation and after nine selection rounds

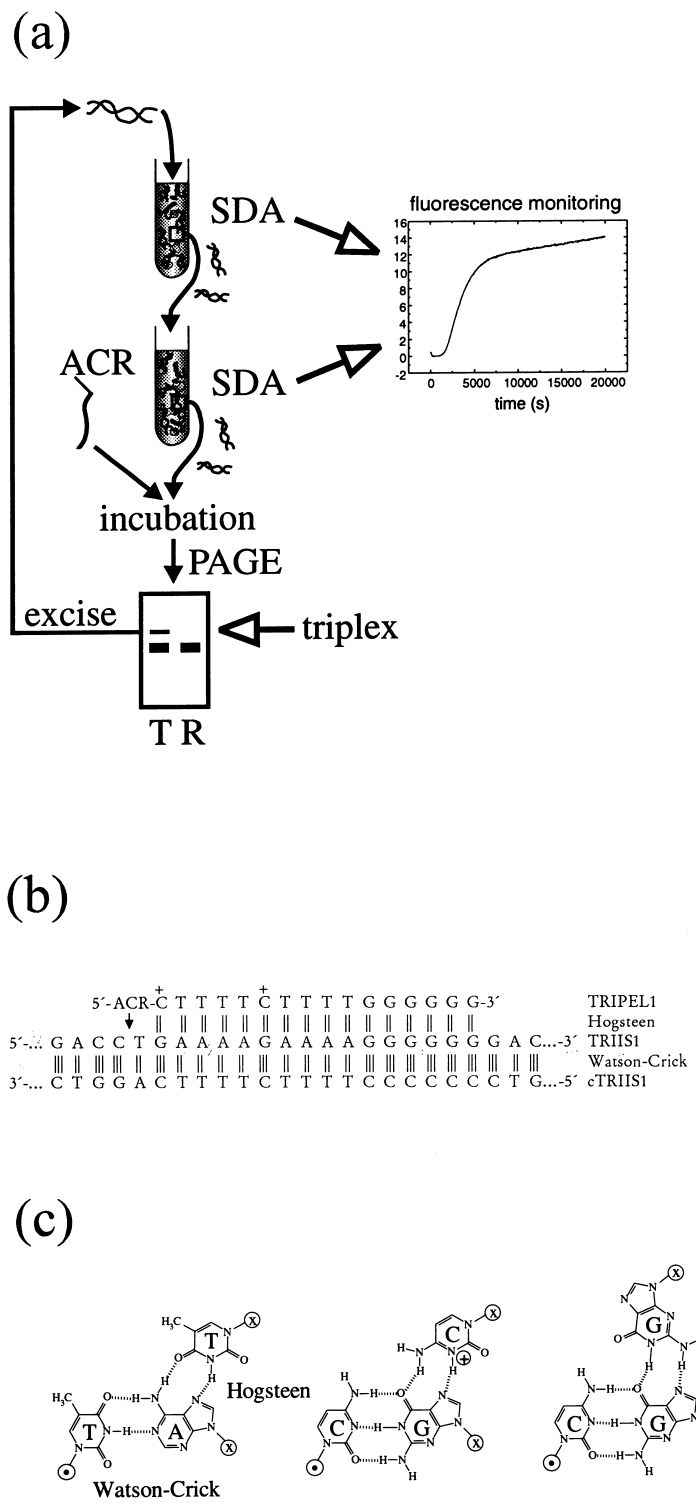

Figure 2. Experimental design for coupled amplification and selection. (a) A selection round consists of two subsequent SDA reactions under fluorescent control with the dsDNA reaction products of the first amplification being serially transfered into the second. Then, a 2-methoxy-6-chlor-9-aminoacridine(ACR)-coupled oligodeoxynucleotide is added in an appropriate buffer, incubated and the products are separated on a non-denaturating polyacrylamide gel. The obtained triplex in lane T is identified in comparison with a reference in lane R, excised, eluted and used to initiate the next selection round. (b) The ACR-coupled oligodeoxynucleotide TRIPEL1 can bind to the major groove of a homopurine·homopyrimidine tract, formed by Watson-Crick base-pairs between starting molecule TRIIS1 and its complement cTRIIS1. Provided the cytosine residues of TRIPEL1 are protonated (+), each base of the oligodeoxynucleotide can form two Hogsteen hydrogen bonds with the GC and AT base-pairs. The ACR moiety further stabilizes the triplex by intercalation into the double strand two base-pairs upstream (arrow). (c) Utilized base triplets $T \times AT$, $C^+ \times GC$ and $G \times GC$ of the triplex (after Thuong & Hélène, 1993).

(with 17 amplification reactions and nine triplex selections, A17S9). When the same analysis procedure was initiated with 0.38 pmol of target molecule TRIIS1 (Table 1), no mutation was found in eight sequenced clones, indicating a low error frequency of PCR amplification and cloning. Thus, the observed sequence heterogeneity of analyzed clones should be mostly due to strand displacement amplification by exo$^-$ Klenow polymerase.

After the first amplification reaction with about 13 molecule generations, among the 22 clones of sequence length 33 (excluding constant primer binding sites), eight base substitutions and two deletions (one point deletion and an eight base deletion) are found (clones A1S0. . . in Figure 4): 14 clones (64%) still have the original sequence of TRIIS1, the initial target. The error rate *r* (defined as errors per base per replication cycle) of exo$^-$ Klenow

polymerase under the described SDA conditions can then be estimated as

$$r \approx 2(p/(nd)) - r'$$

(Saiki *et al*., 1988), where p is the number of mutated positions among n sequenced bases after d cycles of replication. r' is the error rate for the background amplification process (PCR and cloning), which is dominated by the error-prone *in vitro* amplification by *Taq* DNA polymerase in PCR. r' can thus be identified as the error rate of *Taq* DNA polymerase, determined to be 2×10^{-4} (Saiki *et al*., 1988; Tindall & Kunkel, 1988). These values lead to an error rate of exo$^-$ Klenow polymerase during SDA of 1.5×10^{-3} for base substitutions and 0.4×10^{-3} for deletions, which is about one order of magnitude higher than the error rate determined under standard conditions (Bebenek *et al*., 1990).

After two subsequent amplification reactions or about 29 molecular generations beyond initiation with TRIIS1, among nine analyzed clones (A2S0. . . in Figure 4) only three (33%) still exhibit the correct TRIIS1 sequence. The other six mutants carry altogether seven base substitutions and three large deletions, indicating a broadened quasispecies. In accordance with the above calculations, the error rate of exo$^-$ Klenow polymerase estimated from these values would be 1.7×10^{-3} for base substitutions and 0.4×10^{-3} for deletions, respectively. After five subsequent SDA reactions or 77 molecule generations without selection for triplex formation (clones A5S0. . . in Figure 4), only one of ten analyzed clones (10%) is still identical with TRIIS1. The other clones carry mainly large deletions, indicating that the latter become dominating in the quasispecies during prolonged amplification, since shorter templates with less information content are copied more rapidly in SDA (Walter & Strunk, 1994).

Analysis of the products after nine selection rounds (clones A17S9. . . in Figure 4) shows a dramatically different quasispecies of replicated molecules. Though amplification was performed here for 17 SDA reactions or about 270 molecule generations, the sequence information of replicated templates is strongly retained. Selection for triplex-formation tolerated only single base exchanges at the end of the triplex-forming region in TRIIS1 (underlined in the complementary strand cTRIIS1 in Figure 4). Ten of altogether 33 positions (1, 2, 24, and 26 to 32 in Figure 4) are completely conserved, while five positions (4, 10 and 12 to 14) are changed in ten or more of the analyzed 30 clones. The final majority consensus after nine selection rounds carries a G-to-T base transversion in position 13, a point deletion of an adenosine in position 14 (Figure 4) and is unchanged in its triplex-forming region. Thus, molecular *in vitro* evolution from an initial wild-type sequence to a slightly changed final consensus did occur.

Characterization of selected sequence variants

The final majority consensus sequence after nine selection rounds (FC, Table 1) and its complementary strand (cFC) were synthesized to compare them with the initiating template TRIIS1 (Table 1) and its complement cTRIIS1. First, the ability of the formed double strands dsTRIIS1 and dsFC to bind the oligodeoxynucleotide TRIPEL1 used for the selection step were studied. Figure 5(a) shows that both double strands form a triplex in the presence of TRIPEL1 with about the same yield.

Second, the regions of dsTRIIS1 and dsFC binding TRIPEL1 were characterized. A novel fluorescent footprinting assay was established, with the dsDNA being labeled with tetramethylrhodamine (TMR) in a PCR using the 5' modified primer SDARHOP3 (Table 1), then incubated with TRIPEL1 and digested with DNase I. The nuclease is known to bind to the minor groove of dsDNA and to cleave the phosphodiester backbone in a random fashion in the presence of Mg^{2+}. Cleavage is inhibited by the alteration of duplex structure induced by a triplex-forming oligodeoxynucleotide binding to the major groove of the DNA (Stonehouse & Fox, 1994). The fluorescent label on the homopyrimidine-carry-

Table 1. Oligodeoxynucleotides used in this study

Oligodeoxynucleotide	Sequence (5' → 3')
SDARSAP1	TTGAATAGTCGGTTACTTgttgacACTCGACCTGAAA
SDARSAP3	GCATTATGAATCCTGTCTgttgacACTGAGATCCCCT
SDARHOP3	TMR-GCATTATGAATCCTGTCTgttgacACTGAGATCCCCT
SDAPCRP1	TTGAAT AGTCGGTTACTTaagcttACTCGACCTGAAA
SDAPCRP3	GCATTATGAAT CCTGTCTggattcACTGAGATCCCCT
TRIPEL1	ACR-CTTTTCTTTTGGGGGG
TSDAB1	CGATCGAGCAAGCCA
TRIIS1	ACTCGACCTGAAAAGAAAAGGGGGGGACGTTA***TC***CACCATACGGATAGGGGATCTCAGT
cTRIIS1	ACTGAGATCCCCTATCCGTATGGTG***GA***TAACGTCCCCCCCTTTTCTTTTCAGGTCGAGT
FC	ACTCGACCTGAAAAGAAAAGGGGGGGACGTTAACA***C***CATAC***G***GATAGGGGATCTCAGT
cFC	ACTGAGATCCCCTATC***C***GTATG***G***TGTTAACGTCCCCCCCTTTTCTTTTCAGGTCGAGT
MC	ACTCGACCTGAAAAGAAAAGGGGGGGACGTTAACACATACGATAGGGGATCTCAGT

The sequence of TRIIS1 is derived from *IS6110*, an insertion element of *Mycobacteria* already used by Walker *et al*. (1992) as target for SDA, with addition of a potential triplex-forming region; underlined, potential triplex-forming region; lower-case letters, *Hin*cII, *Hin*dIII and *Bam*HI recognition sites; italics and bold, in TRIIS1 and its complement cTRIIS1 the two bases are marked that are exchanged or deleted in FC and cFC, respectively, while in FC and cFC the two bases are marked that are further deleted in MC; TMR and ACR are tetramethylrhodamine and acridine labels, respectively.

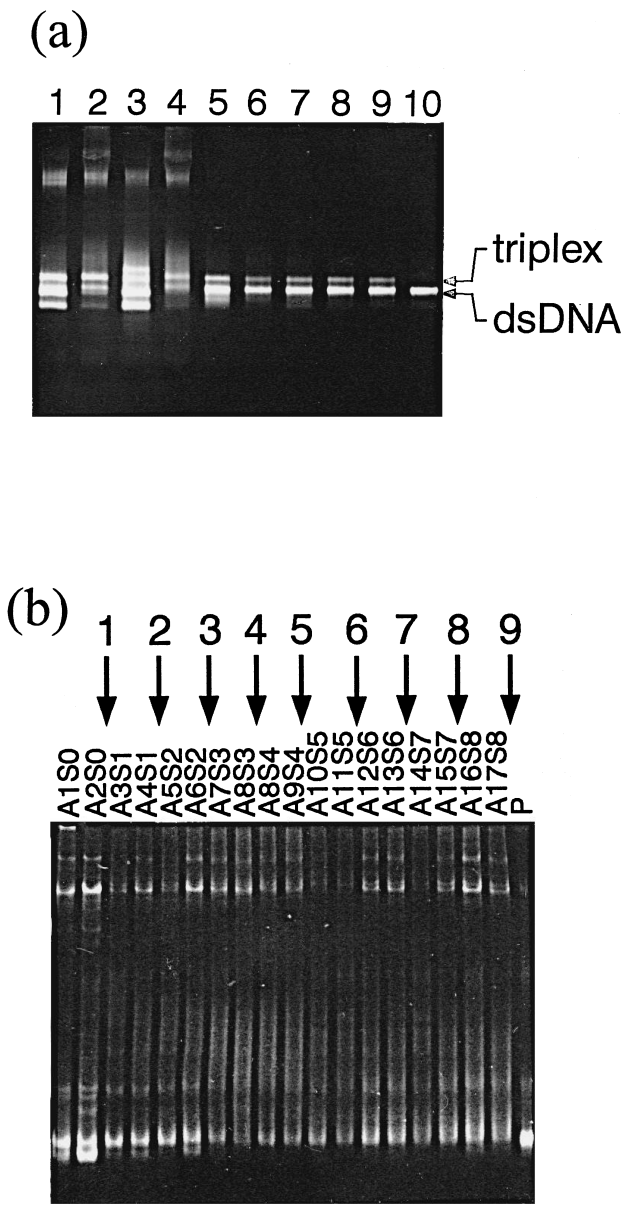

Figure 3. Gel analysis of the evolution experiment. (a) Stopped SDA reaction mixtures were incubated with the triplex-forming oligodeoxynucleotide TRIPEL1 and loaded onto a nondenaturing polyacrylamide gel in 50 mM Tris-borate, 1 mM $MgCl_2$ as described in Materials and Methods. Lanes 1 to 9 correspond to the incubation mixtures that were used for triplex selection in rounds 1 to 9. Conditions were as in Materials and Methods with the following details: selection rounds 1 to 3, serial transfer dilution 1:60,000; selection round 4, no serial transfer; selection rounds 5 to 9, serial transfer dilution 1:15,000; selection rounds 1 to 4, SDA stopped by two minutes at 95°C; selection rounds 5 to 9, SDA stopped by addition of 10 mM EDTA; selection rounds 1 to 4, 5 μM TRIPEL1, selection rounds 5 to 9, 2 μM TRIPEL1; selection rounds 1 to 3, sodium cacodylate buffer; selection rounds 4 to 9, potassium phosphate buffer; selection rounds 1 to 4, pH 6.5; selection rounds 5 to 9, pH 7.0. Lane 10 shows the same SDA reaction mixture as in lane 9 without addition of TRIPEL1. The obtained triplex exhibits a lower electrophoretic mobility than the unmodified dsDNA. (b) A volume of 2 μl each of terminated SDA reactions A1S0 to A17S8 (named after the number of subsequent amplification reaction and triplex selection cycles they had undergone) were subjected to SSCP analysis as described by Orita *et al*. (1989). In principle, double strands were melted in a formamide and EDTA-containing buffer and the obtained single strands subjected to non-denaturing gel electrophoresis in a high-resolution polyacrylamide gel. Arrows indicate triplex selection. In the fourth selection round, A7S3 was used (instead of A8S3) for triplex selection. Lane P was loaded with 2 μl of A17S8 without preceding denaturation.

ing strand (corresponding to cTRIIS1) of the duplex made possible the identification of the cleavage products of this strand on a non-radioactive automated sequencer. Figure 5(b) shows the obtained fragment patterns. When the triplex-forming oligodeoxynucleotide TRIPEL1 was added at a pH below 6, the produced cleavage pattern showed a gap in the expected triplex-forming region. Very similar gaps were found for both dsTRIIS1 and dsFC, indicating that both duplexes bind TRIPEL1 in that same conserved region.

Third, the putative secondary structures of TRIIS1 and FC were computationally analyzed using the algorithms of Hofacker *et al*. (1994). Figure 5(c) shows that the G-to-T base transversion and the point deletion in FC (Table 1) strongly destabilize a stem-loop-structure of TRIIS1. The resulting less compactly folded structure is accompanied by a slightly lower electrophoretic mobility of FC and its complement cFC compared with TRIIS1 and cTRIIS1 on a native gel (Figure 5(a)). Since secondary structure of single strands plays an important role in strand displacement amplification (Walter & Strunk, 1994), it seemed likely that the amplification efficiencies of the different targets might differ. This difference could then account for the selection of the consensus sequence FC in the course of the evolution experiment.

To test this hypothesis, the standard SDA reaction mixture was initiated with either 1 nM template TRIIS1 or FC and the amplification kinetics were followed by fluorescence detection. The difference between both samples was quite significant, with TRIIS1 being characterized by a faster increase in fluorescence intensity (Figure 6(a)). When 1 nM TRIIS1 and 1 nM FC were used to initiate the same SDA sample, always TRIIS1 grew out of the mixture (results not shown). To exclude differences due to unequal concentrations of template stock solutions, concentrations were controlled independently by absorbance measurement at 260 nm and gel electrophoresis followed by ethidium bromide staining (Sambrook *et al*., 1989). Moreover, the number of temperature cycles necessary to obtain equal amplification factors for TRIIS1 and FC in a PCR reaction were determined and found to be identical, proving that both templates were comparably free of lesions. The differences in strand displacement amplification could therefore be explained only by different replication efficiencies

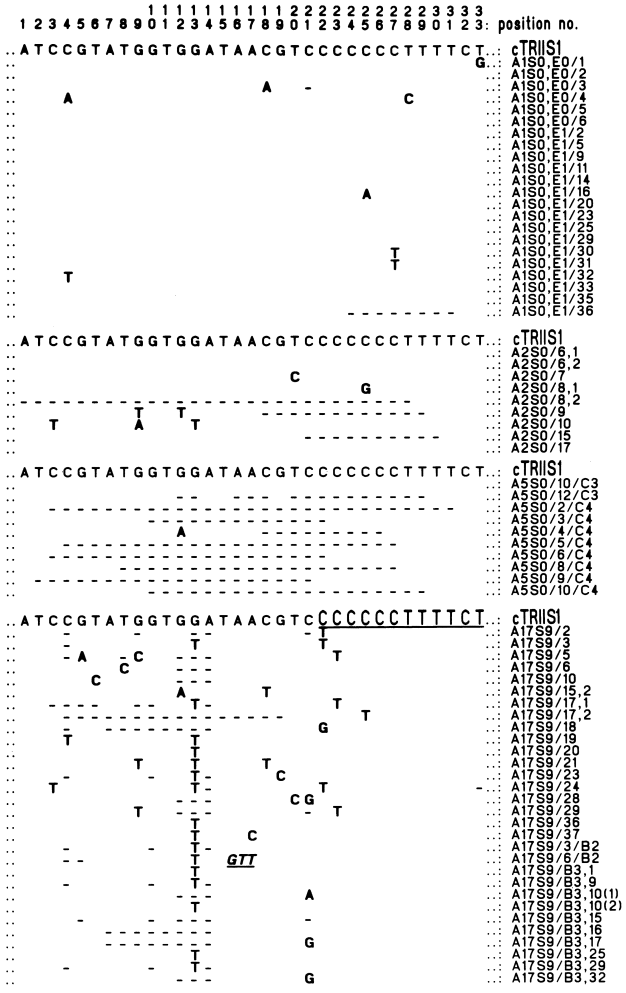

Figure 4. Sequence analysis of amplification reactions A1S0, A2S0, A5S0 (without selection) and A17S9 (after nine selection rounds). The sequences were obtained using a reverse sequencing primer and are given in the 5' to 3' direction excluding in every clone the unchanged primer binding sites (upstream of position 1 and downstream of position 33). Only differences from the complementary sequence of starting molecule TRIIS1 are shown, with dashes indicating deletions. One insertion is highlighted, underlined and in italics with its first base under its insertion site. Underlined and enlarged is the homopyrimidine sequence in cTRIIS1, to which the oligodeoxynucleotide binds during triplex selection. Without selection all positions are equally exchanged or deleted with ongoing amplification, while under selection pressure the triplex-forming region in the cTRIIS1 sequence is strongly conserved and only two hotspot positions are exchanged in more than 50% of the clones. Large deletions might be slightly overepresented by a bias of the cloning procedure for shorter sequences.

of the template molecules during SDA, with TRIIS1 being surprisingly faster amplified under the chosen conditions than the final consensus FC.

To explain these findings, one has to keep in mind that the majority consensus sequence FC itself is not represented among the characterized clones after selection (Figure 4). This is quite typical for a quasispecies distribution of mutants, since the consensus or wild-type is only a representation of the most probable base for each position (Eigen *et al*., 1989; Rohde *et al*., 1995). We therefore also analyzed the most abundant clone MC (Table 1) that was found in 10% of the clones after selection (namely A17S9/3/B2, -/B3,9 and -/B3,29 in Figure 4). Initiating SDA with either 1 nM or 1 pM MC led to a faster fluorescence increase as in the case of TRIIS1 or FC (Figure 6(a)). To extract selection values for the amplification efficiency under exponential and linear growth from the fluorescence curves, the method of Biebricher *et al*. (1985, 1991) was applied. As shown in Figure 6(b) for TRIIS1, κ as a measure for exponential growth was obtained from the displacement τ_{dif} of the curves for 1 nM and 1 pM template on the time axis according to $\kappa = \ln 1000 / \tau_{dif}$, while ρ was calculated from the slope of the linear part of the fluorescence profiles. Table 2 represents the values of κ and ρ for the three templates TRIIS1, FC and MC in arbitrary units, showing that FC is replicated best in the exponential, but slower than TRIIS1 in the linear growth phase. MC is replicated well in the exponential and most rapidly in the linear phase. Since most of the amplification in the selection experiment occurred in the linear phase with only a short initial exponential growth period, it becomes plausible that TRIIS1 evolved into a quasispecies that is dominated by the replicationally advantageous mutant MC.

Discussion

Evolution consists of repeated cycles of amplification, mutation (often coupled with amplification) and selection of advantageous mutants. Evolution has classically been observed *in vitro* with RNA using the replicase of the *Escherichia coli* RNA phage Qβ (Mills *et al*., 1967; Spiegelman, 1971; Kramer *et al*., 1974; Biebricher *et al*., 1985). This RNA polymerase is able to copy both the plus and minus strand of its template in a cell-free environment with a high error frequency of about 3×10^{-4} per base per replication cycle (Eigen, 1993). By applying the method of serial transfer, replication can be prolonged indefinitely (Mills *et al*., 1967). Error accumulation results in the formation of a broad quasispecies of nearly neutral mutants that is able to adapt to a variety of different selection pressures (Spiegelman, 1971; Eigen, 1993; Rohde *et al*., 1995). This experience of evolutionary mechanisms at the molecular level has highlightened our understanding of RNA virus evolution; e.g. quasispecies distributions have been shown to make possible the fast adaption of RNA virus populations to the immune response of their host (Domingo *et al*., 1985; Eigen, 1993).

Here, an *in vitro* evolution experiment on double-stranded DNA with selection for triplex formation ability has been performed. Amplification and mutagenesis were done by isothermal strand displacement amplification (SDA; Walker *et al*., 1992) starting from a defined sequence.

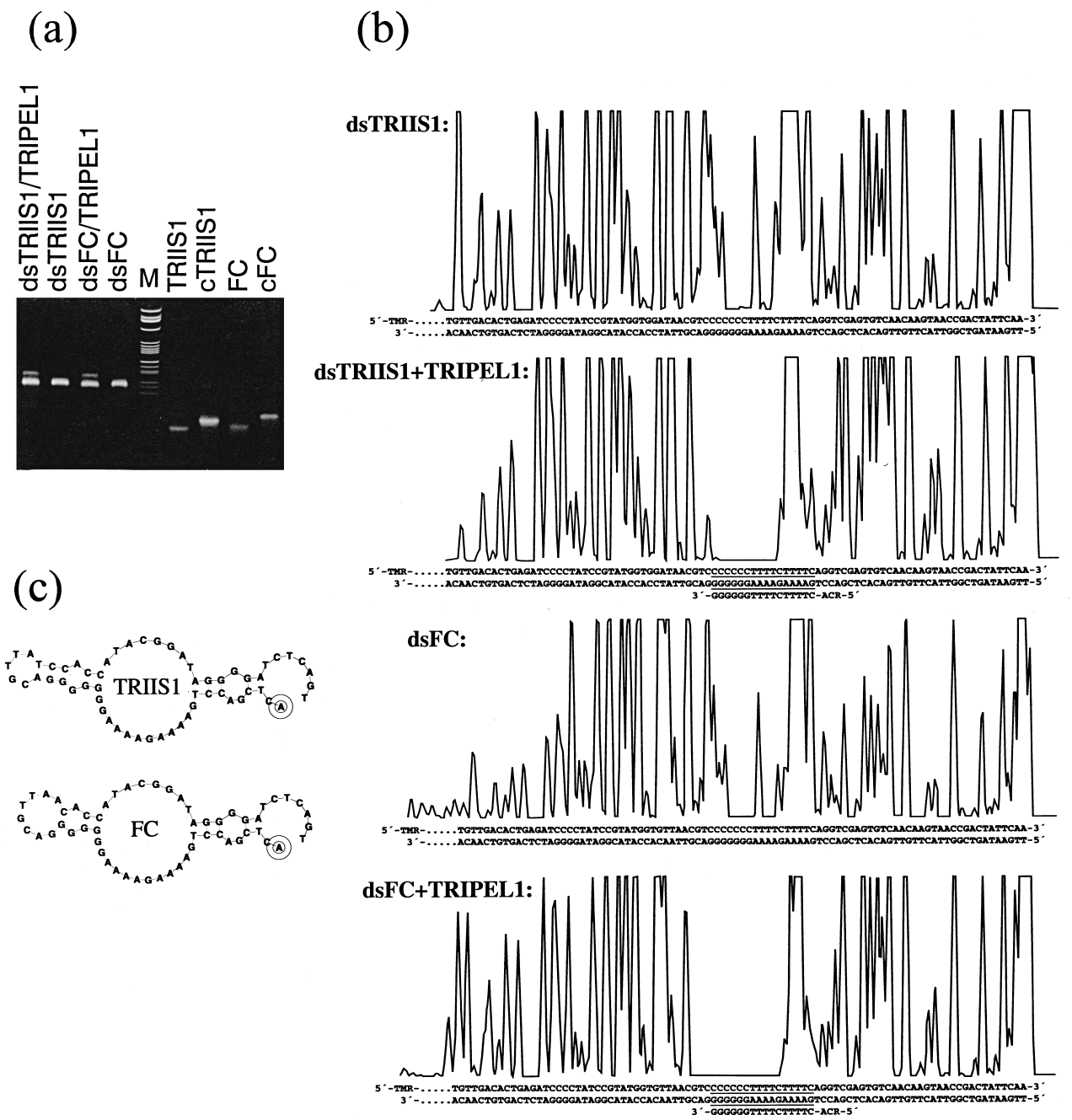

Figure 5. Characterization of selected sequence variants. (a) TRIIS1 and cTRIIS1 as well as FC and cFC (Table 1) were mixed in equal concentrations and annealed to give dsTRIIS1 and dsFC, respectively. The double strands were incubated with TRIPEL1 and analyzed on a non-denaturing polyacrylamide gel in 50 mM Tris-borate, 1 mM $MgCl_2$ as described in Materials and Methods, together with the single strands as controls. Lane M was loaded with DNA marker V from Boehringer (Mannheim, Germany). (b) The fragment pattern of a non-radioactive DNase I footprinting assay utilizing a TMR-label in the homopyrimidine strand (see Materials and Methods) shows that both dsTRIIS1 and dsFC bind the ACR-coupled oligodeoxynucleotide TRIPEL1 in the same region (underlined). (c) The most likely secondary structures of TRIIS1 and FC at 20°C were calculated by the algorithms of Hofacker *et al*. (1994). The 5' end is marked by circles. Each 13 nucleotides at the 5' and 3' ends are predefined by the primers. At 37°C, the TRIIS1 structure is still found, while cFC is completely unstructured.

Previously it has been shown that SDA resembles rolling-circle replication, its cycles comprising nicking of a double-stranded DNA, elongation of the derived 3' end and displacement of the downstream copy strand, followed by conversion of the displaced single-stranded intermediate to a double strand again (Figure 1; Walker *et al*., 1992; Walter & Strunk, 1994). Rolling-circle replication and closely related mechanisms play an important role in the replication of DNA genomes (Gilbert & Dressler, 1968; Ilyina & Koonin, 1992). There is evidence that they account for the *in vivo* generation

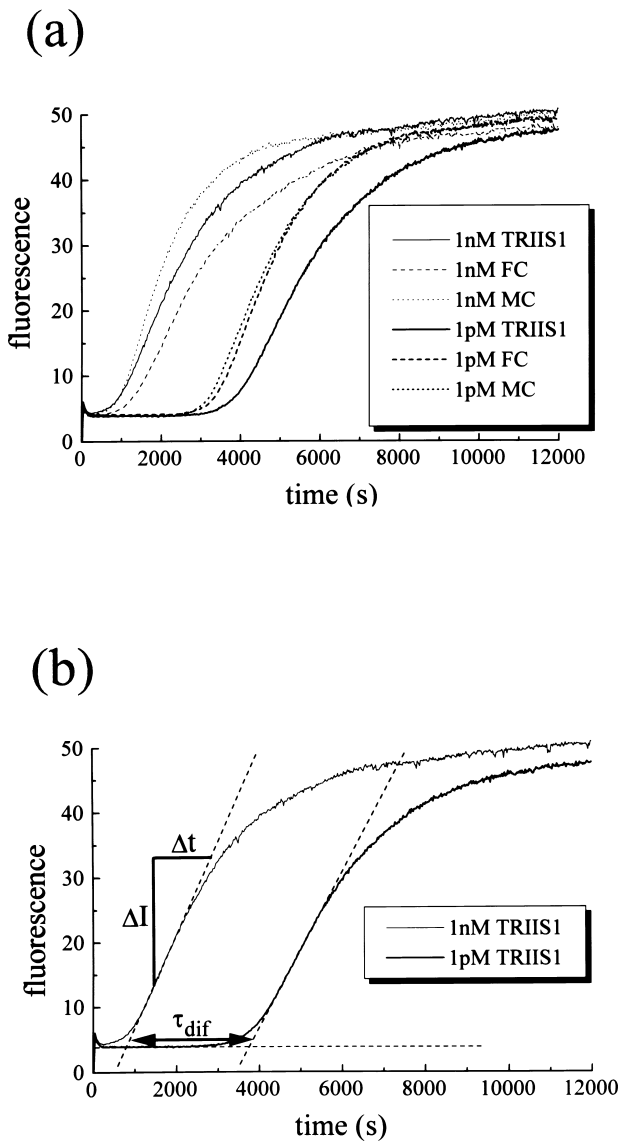

Figure 6. Fluorescence curves of strand displacement amplified sequence variants. (a) A standard SDA reaction mixture was initiated with the indicated concentration of mutants TRIIS1, FC and MC, and the increase of TO fluorescence upon replication monitored in a fluorometer (see Materials and Methods). Since the three mutants varied in sequence length, the identity of replication products could be confirmed by high-resolution denaturing polyacrylamide gel electrophoresis. (b) From the fluorescence curves, rate constants in arbitrary units could be extracted by the method of Biebricher *et al*. (1985, 1991). Briefly, the kinetics in the linear phase were characterized by the slope $\Delta I/\Delta t$, while kinetics in the exponential phase, that mainly occurs below the fluorescence detection limit, were analyzed using the displacement τ_{dif} between two curves initiated with different mutant concentrations. When the concentrations differ by a factor of 1000, then the exponential kinetics can be measured by $\kappa = \ln 1000/\tau_{dif}$.

Table 2. Selection values of initial and selected templates in the exponential (κ) and linear growth phase (ρ)

Sequence	$\kappa/10^{-3}$ (s^{-1})	$\rho/10^{-2}$ (s^{-1})
TRIIS1	2.4 ± 0.3	1.5 ± 0.2
FC	3.2 ± 0.3	1.2 ± 0.2
MC	3.1 ± 0.3	2.0 ± 0.2

κ and ρ were extracted from the fluorescence curves as described for Figure 6(b).

of plasmids in prokaryotes (Gruss & Ehrlich, 1989) and eukaryotes (Dasgupta *et al*., 1992), of single-stranded DNA viruses like filamentous bacteriophages (Baas & Jansz, 1988), plant geminiviruses (Laufs *et al*., 1995) or animal parvoviruses (Berns, 1990), of double-stranded DNA viruses like bacteriophage λ (Hendrix *et al*., 1983) and herpes simplex virus (Skaliter & Lehman, 1994) and of certain transposons (Mendiola *et al*., 1994).

In the case of *Escherichia coli* plasmid and bacteriophage replication, DNA polymerase I takes part in lagging strand synthesis by filling the gaps between the Okazaki fragments generated by DNA polymerase III, a process that leads to a higher susceptibility to mutations than in leading strand synthesis (Veaute & Fuchs, 1993). In the present study, the enzyme applied for replication and mutagenesis is the exonuclease-deficient large (Klenow) fragment of *Escherichia coli* DNA polymerase I (exo⁻ Klenow polymerase). Its proofreading activity is suppressed by five orders of magnitude due to a double mutation in the 3'-5'-exonuclease domain of the large fragment (Derbyshire *et al*., 1988). Since proofreading by excision of misincorporated nucleotides plays a major role in replication fidelity (Goodman *et al*., 1993), the error rate of exo⁻ Klenow polymerase under standard conditions is with 10^{-4} errors per base per replication about four to seven times higher than that of Klenow polymerase (dependent on the error of interest; Bebenek *et al*., 1990). Under the *in vitro* conditions applied in the present study, the determined error frequency is even an order of magnitude higher than this value. This may be due to the artificial conditions of SDA reaction under fluorescent detection, utilizing dATPαS and thiazole orange. dATPαS is known to suppress any residual exonuclease activity (Kunkel *et al*., 1981), while fluorescent dyes are mutagens (Kramer *et al*., 1974). Both effects could increase the error frequency. Though they may also limit the direct applicability of acquired results on to the *in vivo* situation, the obtained high error rates are the only chance to produce a broad quasispecies of mutants and to observe evolutionary processes with short nucleic acids replicated *in vitro* (Eigen *et al*., 1989).

After SDA in the exponential and linear growth phase without selection pressure other than for fast replication, large deletions dominate the obtained quasispecies, as was already found for continuous serial transfer under exponential growth conditions (Walter & Strunk, 1994). Deletions can be formed by primer misalignments (often after nucleotide misincorporation; Bebenek & Kunkel, 1990), *via* a ''slipped mispairing'' mechanism, as postulated for single-stranded phage and bacterial plasmids

(Horiuchi, 1983; Vilette *et al*., 1992) or in a more complex form by ''strand switching'' in quasi-palindromic sequence contexts (Ripley, 1990). Deletions are favored for low processive polymerases like the Klenow fragment (Polesky *et al*., 1990; Wang & Ripley, 1994) and may be induced by fluorescent dyes like thiazole orange or acridine (Streisinger *et al*., 1966). They inevitably lead to loss of information content and are natural antagonists for complex functionality of genomes.

To counteract loss of sequence information, nine selection rounds with SDA replication coupled to selection for triplex formation were performed. The obtained quasispecies demonstrates that alternating replication and triplex selection yields mutants adapted to both selection pressures. The system essentially conserves the homopurine·homopyrimidine sequence tract capable of oligonucleotide binding that was already contained in the starting molecule TRIIS1 (Figure 2(b)), indicating that it did not find a superior alternative. This is not unexpected, because it is known that the utilized natural nucleobases exhibit only a limited number of possible triplets (Thuong & Hélène, 1993). Moreover, triplex selection *via* gel electrophoresis implies selection of strands that show a defined mobility shift and thus a roughly fixed sequence length. Heavily deleted sequences are selected against. Nevertheless, the selected sequences indicate (Figure 4) that single base exchanges or deletions at the end of the triplex-forming region seem to be tolerated under the experimental conditions (even a C-to-T change in position 25, four bases downstream from the duplex/triplex junction, is found in one clone). This is in accordance with earlier footprinting data (Stonehouse & Fox, 1994).

In order to essentially maintain their triplex-forming tract and at the same time enhance replication efficiency, mutants to be selected had to optimize their sequence for amplification by SDA. The obtained final consensus sequence FC (Table 1) carries one G-to-T base substitution and an adjacent point deletion in comparison with TRIIS1, resulting in a less stable secondary structure (Figure 5(c)). The G-to-T base substitution was also frequently found earlier under exclusively exponential growth conditions without triplex selection (Walter & Strunk, 1994). In the present study, the two mutations of FC were found to be accompanied by an increase of amplification efficiency in the exponential growth phase by 33% (Table 2). The easiest explanation for these findings would be facilitated binding of primers due to the less compact secondary structure as the dominant step in exponential growth. At the same time, amplification efficiency in the linear phase is reduced for FC against TRIIS1 by about 20%, which could be explained by a less efficient primer elongation and strand displacement of the downstream complement (Figure 1). As a consequence, TRIIS1 outgrowths FC under the SDA conditions used in the selection experiment. MC (Table 1), the most abundant and obviously most rapidly copied mutant, carries two more point deletions. This enhances its replication efficiency against the starting molecule in the exponential phase by 30% and in the linear phase by 33% (Table 2). MC seems to be the best adapted mutant, eventually selected after 270 molecular generations of amplification, mutation and selection.

Observing the described evolutionary phenomena in an *in vitro* amplification system for DNA like SDA has important implications. First, it proves that evolution in principle can take place in any isothermal reproductive system, as already shown for RNA replication by Qβ replicase (Spiegelman, 1971) or by self-sustained sequence replication (Breaker & Joyce, 1994). When the error rate of polymerization is high, a quasispecies of mutants is rapidly produced. When all steps of amplification occur simultaneously at a single temperature, different mutants start competing for resources and optimized amplicons will outgrow the mixture. This process can be coupled to a separate selection step by simple alternation with amplification. Selected mutants are then adapted to both selection pressures and it remains to be analyzed which selection pressure accounts for which mutation. Moreover, the most abundant mutant may be the best adapted sequence rather than the averaged consensus sequence. These findings have to be taken into account when interpreting evolution experiments on nucleic acids (Joyce, 1994).

Second, it gives a hint how selection in the analoguous rolling-circle replication and related strategies of DNA virus and plasmid replication *in vivo* may contribute to evolution of rapidly reproducing genomes. As demonstrated *in vitro*, single base exchanges and point deletions can result in higher amplification efficiencies. As long as they are compatible with functionality of the genome in gene expression (as with triplex formation ability in the *in vitro* experiment), they will be incorporated into the sequence by evolutionary forces. Being able to model these properties of natural replicative systems *in vitro* could give the chance for a detailed understanding and finally the development of novel antiviral strategies as was already the case with the Qβ replicase system.

Materials and Methods

Materials

Exonuclease-deficient large fragment of *Escherichia coli* DNA polymerase I (exo$^-$ Klenow polymerase; Derbyshire *et al*., 1988) was purchased from U.S. Biochemicals Corporation (Cleveland, OH) at a concentration of 5 to 10 units/μl, *Hin*cII (at 50 to 75 units/μl) from New England Biolabs (Beverly, MA), DNase I from Boehringer (Mannheim, Germany), and *Taq* DNA polymerase from Perkin-Elmer (Weiterstadt, Germany). 2'-Deoxyadenosine-5'-*O*-(1-thio-triphosphate; dATPαS; diastereomer mixture) and all other nucleotides were obtained from Pharmacia (Freiburg, Germany). Thiazole orange (TO) was from Molecular Probes (Eugene, OR),

and its concentration was determined by absorption measurement at 476 nm (ϵ_{476}(TO) = 30,000 M^{-1} cm^{-1}; Lee *et al*., 1986). All oligodeoxynucleotides were either synthesized on a Milligene Expedite Synthesizer and purified by gel electrophoresis or synthesized and HPLC-purified by NAPS (Göttingen, Germany). The 2-methoxy-6-chlor-9-aminoacridine (ACR) moiety of oligodeoxynucleotide TRIPEL1 was attached in-line using the phosphoramidite, while tetramethylrhodamine (TMR) was attached post-synthetically *via* an amino-hexyllinker (NAPS, Göttingen, Germany). Aerosol Resistant Tips (ART) from Biozym (Hessisch Oldendorf, Germany) were routinely used to avoid contamination of SDA reaction mixtures with previously amplified products.

SDA reaction with fluorescent on line detection

The very first SDA reaction was initiated with the 59mer oligodeoxynucleotide TRIIS1 as template (Table 1). SDA buffer contained 50 mM potassium phosphate (pH 7.4), 3 mM $MgCl_2$, 4 mM dithiothreitol, 0.01% (v/v) Triton X-100, 3% (v/v) 1-methyl-2-pyrrolidinone (NMP), 0.25 mM each of dGTP, dCTP, dTTP and dATPαS, 0.5 μM each of primers SDARSAP1 and SDARSAP3 (Table 1), 2 μM TO, 56 μg/ml acetylated bovine serum albumin, 1.67 units/μl *Hin*cII and 0.056 units/μl exo$^-$ Klenow polymerase (6.7 mM NaCl, 9.5 μg/ml bovine serum albumin and 2.2% (v/v) glycerol were further added through the enzyme mixture). The addition of the fluorescent dye TO (excited at 490 nm, fluorescence detected at 530 nm) that binds to double-stranded nucleic acids under increase of fluorescence quantum yield allows the amplification reaction to be recorded on-line. The reaction sample was incubated with typically 2×10^{10} template strands in a final volume of 60 μl prior to addition of any proteins and incubated for two minutes at 95°C, followed by one minute at 39°C. A serial transfer was likewise performed by adding 1 μl of an appropriate dilution of the products of the preceding reaction into a fresh reaction mixture. Upon addition of proteins, the amplification mixture was incubated for either 12,000 or 20,000 seconds at 39°C in a thermostatted fluorometer cuvette. The increase in fluorescence was monitored using a Perkin-Elmer LS5B fluorometer. The reaction was stopped either by heating at 95°C for two minutes or by addition of EDTA to a final concentration of 10 mM and the reaction mixtures were stored at − 20°C.

Selection for double-stranded SDA products that bind an oligodeoxynucleotide

A volume (2 μl) of a terminated SDA reaction mixture or of a dsDNA solution was incubated with 2 to 5 μM of the 5' ACR-labeled oligodeoxynucleotide TRIPEL1 (Table 1) in 10 μl of 50 mM sodium cacodylate or potassium phosphate buffer (pH 6.5 or 7.0; see the legend to Figure 3), 10 mM $MgCl_2$, and 100 mM NaCl for 15 minutes at 37°C, followed by 15 minutes at 16°C. Then, the complete mixture was loaded at 4°C onto a nondenaturing polyacrylamide gel in 50 mM Tris-borate (pH 8.0), 1 mM $MgCl_2$. The gel was run at 10 V/cm for 3.5 hours and stained with ethidium bromide. The triplex band was excised, diffusion eluted, desalted, and 0.1 of the DNA was used to initiate the next SDA reaction.

Analysis of SDA products

SSCP analysis of SDA products was performed in an MDE-hydrolink gel (Serva, Heidelberg, Germany) as described by the manufacturer. For sequence analysis, either 1 μl of a terminated SDA reaction solution or the purified triplex DNA was further amplified in a final volume of 100 μl PCR reaction mixture (10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% (w/v) gelatine, 0.04 units/μl *Taq* DNA polymerase) containing 0.5 μM each of primers SDAPCRP1 and SDAPCRP3 with *Hin*dIII and *Bam*HI restriction sites, respectively (Table 1). Thermocycling was performed in a TRIO-Thermoblock cycler (Biometra, Göttingen, Germany) for a total of 16 cycles with a temperature program of 94°C for 30 seconds, 30°C (in the first two cycles) or 68°C (in the following 14 cycles) for 30 seconds and 72°C for 30 seconds. The products were purified by non-denaturing polyacrylamide gel electrophoresis, restricted with *Hin*dIII and *Bam*HI, cloned into pUC18 vector and transformed into *Epicurian coli* SURE™ competent cells (Stratagene, Heidelberg, Germany). Finally, the plasmid DNA from single clones was prepared using the WIZARD™ miniprep protocol (Promega, Heidelberg, Germany) according to the manufacturer's instructions, further purified by precipitation in ethanol, and sequenced as described (Walter & Strunk, 1994).

Non-radioactive DNase I footprinting

For non-radioactive detection, an isotopic footprinting assay (Stonehouse & Fox, 1994) was adapted to fluorescence detection of the digested fragments. TMR-labeled dsDNA was generated by PCR-amplification of the templates TRIIS1 and FC as described above using 1 μM primers SDARSAP1 and SDARHOP3 (Table 1; the latter is a 5' TMR-labeled form of primer SDARSAP3). Seven thermal cycles with 94°C for 30 seconds, 30°C (in the first two cycles) or 68°C (in the following five cycles) for 30 seconds and 72°C for 30 seconds were performed, the main product band identified by non-denaturing polyacrylamide gel electrophoresis, excised, diffusion eluted, and desalted. 25 ng (0.36 pmol) of this PCR-product were incubated with 0 to 5 μM TRIPEL1 and 20 to 15 μM of carrier oligodeoxynucleotide TSDAB1 (Table 1) in 10 μl of 50 mM sodium acetate (pH 5.5), 10 mM $MgCl_2$ and 100 mM NaCl for 15 minutes at 37°C, followed by 15 minutes at 16°C and five minutes at 22°C. A volume (2 μl) of a solution of 0.025 units/μl DNase I dissolved in 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$ was added, the mixture incubated for two minutes at 22°C and digestion stopped by addition of 138 μl of 20 mM EDTA (pH 8.0). Two-thirds of this mixture was twice extracted with phenol/chloroform, precipitated with ethanol, the precipitate washed, and loaded onto a 6% sequencing gel to be analyzed by electrophoresis on a model 373A DNA sequencer (Applied Biosystems, Weiterstadt, Germany). TMR-labeled fragments showed up in the ''T''-signal. The ACR-labeled oligodeoxynucleotide TRIPEL1 (visible in the ''G''-signal) does not interfere with the TMR-signal.

Acknowledgements

This work has been supported in part by a grant from the Bundesministerium für Forschung und Technologie (0310248A4) and by a Kekulé fellowship by the Stiftung Stipendien-Fonds des Verbandes der Chemischen Indus-

trie for N.G.W. The author very warmly thanks S. Völker for excellent technical assistance, M. Meyer for careful preparation of photographic prints, O. Matzura for programming structure calculation algorithms, C. Biebricher for critical reading of the manuscript, G. T. Walker for helpful discussions and M. Eigen for a very stimulating environment.

References

Baas, P. D. & Jansz, H. S. (1988). Single-stranded DNA phage origins. *Curr. Top. Microbiol. Immunol.* **136**, 31–70.

Bebenek, K. & Kunkel, T. A. (1990). Frameshift errors initiated by nucleotide misincorporation. *Proc. Natl Acad. Sci. USA*, **87**, 4946–4950.

Bebenek, K., Joyce, C. M., Fitzgerald, M. P. & Kunkel, T. A. (1990). The fidelity of DNA synthesis catalazed by derivatives of *Escherichia coli* DNA polymerase I. *J. Biol. Chem.* **265**, 13878–13887.

Berns, K. I. (1990). Parvovirus replication. *Microbiol. Rev.* **54**, 316–329.

Biebricher, C. K., Eigen, M. & Gardiner, W. C. (1985). Kinetics of RNA replication: Competition and selection among self-replicating RNA species. *Biochemistry*, **24**, 6550-6560.

Biebricher, C. K., Eigen, M. & Gardiner, W. C. (1991). Quantitative analysis of selection and mutation in self-replicating RNA. In *Biologically Inspired Physics* (Peliti, L., ed.), pp. 317–337, Plenum Press, New York.

Breaker, R. R. & Joyce, G. F. (1994). Emergence of a replicating species from an *in vitro* RNA evolution reaction. *Proc. Natl Acad. Sci. USA*, **91**, 6093-6097.

Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988). Site-specific oligonucleotide binding represses transcription of the human c-*myc* gene in vitro. *Science*, **241**, 456–459.

Dasgupta, S., Zabielski, J., Simonsson, M. & Burnett, S. (1992). Rolling-circle replication of a high-copy BPV-1 plasmid. *J. Mol. Biol.* **228**, 1–6.

Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M. & Steitz, T. A. (1988). Genetic and crystallographic studies of the 3',5'-exonucleolytic site of DNA polymerase I. *Science*, **240**, 199–201.

Domingo, E., Martínez-Salas, E., Sobrino, F., de la Torre, J. C., Portela, A., Ortín, J., López-Galindez, C., Pérez-Brena, P., Villanueva, N., Nájera, R., VandePol, S., Steinhauer, D., DePolo, N. & Holland, J. (1985). The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. *Gene*, **40**, 1–8.

Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, **58**, 465–523.

Eigen, M. (1993). The origin of genetic information: viruses as models. *Gene*, **135**, 37–47.

Eigen, M., McCaskill, J. & Schuster, P. (1989). The molecular quasispecies. *Advan. Phys. Chem.* **75**, 149–263.

Gilbert, W. & Dressler, D. (1968). DNA replication: The rolling circle model. *Cold Spring Harbor Symp. Quant. Biol.* **33**, 473–484.

Giovannangéli, C., Rougée, M., Garestier, T., Thuong, N. T. & Hélène, C. (1992). Triple-helix formation by oligonucleotides containing the three bases thymine, cytosine, and guanine. *Proc. Natl Acad. Sci. USA*, **89**, 8631–8635.

Goodman, M. F., Creighton, S., Bloom, L. B. & Petruska, J. (1993). Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol.* **28**, 83–126.

Gruss, A. & Ehrlich, S. D. (1989). The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53**, 231–241.

Hayashi, K. & Yandell, D. W. (1993). How sensitive is PCR-SSCP? *Human Mutat.* **2**, 338–346.

Hendrix, R. W., Roberts, J. W., Stahl, F. W. and Wesiberg, R. A. (1983). Editors of *Lambda II*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M. & Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Chem. Monthly*, **125**, 167–188.

Horiuchi, K. (1983). Co-evolution of a filamentous bacteriophage and its defective interfering particles. *J. Mol. Biol.* **169**, 389–407.

Ilyina, T. V. & Koonin, E. V. (1992). Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucl. Acids Res.* **20**, 3279–3285.

Joyce, G. F. (1994). *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.* **4**, 331–336.

Kramer, F. R., Mills, D. R., Cole, P. E., Nishihara, T. & Spiegelman, S. (1974). Evolution *in vitro*: sequence and phenotype of a mutant RNA resistant to ethidium bromide. *J. Mol. Biol.* **89**, 719–736.

Kunkel, T. A., Eckstein, F., Mildvan, A. S., Koplitz, R. M. & Loeb, L. A. (1981). Deoxynucleoside[1-thio]triphosphates prevent proofreading during *in vitro* DNA synthesis. *Proc. Natl Acad. Sci. USA*, **78**, 6734–6738.

Laufs, J., Traut, W., Heyraud, F., Matzeit, V., Rogers S. G., Schell, J. & Gronenborn, B. (1995). *In vitro* cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl Acad. Sci. USA*, **92**, 3879–3883.

Lee, L. G., Chen, C. H. & Chiu, L. A. (1986). Thiazole orange: a new dye for reticulocyte analysis. *Cytometry*, **7**, 508–517.

Mendiola, M. V., Bernales, I. & de la Cruz, F. (1994). Differential roles of the transposon termini in IS*91* transposition. *Proc. Natl Acad. Sci. USA*, **91**, 1922–1926.

Mills, D. R., Petersen, R. L. & Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl Acad. Sci. USA*, **58**, 217–224.

Nüesch, J., Cotmore, S. F. & Tattersall, P. (1995). Sequence motifs in the replicator protein of parvovirus MVM essential for nicking and covalent attachment to the viral origin: identification of the linking tyrosine. *Virology*, **209**, 122–135.

Orita, M., Suzuki, Y., Sekiya, T. & Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics*, **5**, 874–879.

Polesky, A. H., Steitz, T. A., Grindley, N. D. & Joyce, C. M. (1990). Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. *J. Biol. Chem.* **265**, 14579–14591.

Rohde, N., Daum, H. & Biebricher, C. K. (1995). The mutant distribution of an RNA species replicated by Qβ replicase. *J. Mol. Biol.* **249**, 754–762.

Ripley, L. S. (1990). Frameshift mutation: determinants of specificity. *Annu. Rev. Genet.* **24**, 189–213.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Skaliter, R. & Lehman, I. R. (1994). Rolling circle DNA replication *in vitro* by a complex of herpes simplex virus type 1-encoded enzymes. *Proc. Natl Acad. Sci. USA*, **91**, 10665–10669.

Spiegelman, S. (1971). An approach to the experimental analysis of precellular evolution. *Quart. Rev. Biophys.* **4**, 213–253.

Stonehouse, T. J. & Fox, K. R. (1994). DNase I footprinting of triple helix formation at polypurine tracts by acridine-linked oligopyrimidines: stringency, structural changes and interaction with minor groove binding ligands. *Biochim. Biophys. Acta*, **1218**, 322–330.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966). Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* **31**, 77–84.

Sun, J. S., Francois, J. C., Montenay-Garestier, T., Saison-Behmoaras, T., Roig, V., Thuong, N. T. & Hélène, C. (1989). Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA via major groove recognition by oligonucleotide-intercalator conjugates. *Proc. Natl Acad. Sci. USA*, **86**, 9198–9202.

Thuong, N. T. & Hélène, C. (1993). Sequence-specific recognition and modification of double-helical DNA by oligonucleotides. *Angew. Chem. Int. Ed. Engl.* **32**, 666–690.

Tindall, K. R. & Kunkel, T. A. (1988). Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry*, **27**, 6008–6013.

Veaute, X. & Fuchs, R. P. P. (1993). Greater susceptibility to mutations in lagging strand of DNA replication in *Escherichia coli* than in leading strand. *Science*, **261**, 598–600.

Vilette, D., Uzest, M., Ehrlich, S. D. & Michel, B. (1992). DNA transcription and repressor binding affect deletion formation in *Escherichia coli* plasmids. *EMBO J.* **11**, 3629–3634.

Walker, G. T., Little, M. C., Nadeau, J. G. & Shank, D. D. (1992). Isothermal *in vitro* amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc. Natl Acad. Sci. USA*, **89**, 392–396.

Walter, N. G. & Strunk, G. (1994). Strand displacement amplification as an *in vitro* model for rolling-circle replication: deletion formation and evolution during serial transfer. *Proc. Natl Acad. Sci. USA*, **91**, 7937–7941.

Wang, F. J. & Ripley, L. S. (1994). DNA sequence effects on single base deletions arising during DNA polymerization *in vitro* by *Escherichia coli* Klenow fragment polymerase. *Genetics*, **136**, 709–719.

Zhu, H., Clark, S. M., Benson, S. C., Rye, H. S., Glazer, A. N. & Mathies, R. A. (1994). High-sensitivity capillary electrophoresis of double-stranded DNA fragments using monomeric and dimeric fluorescent intercalating dyes. *Anal. Chem.* **66**, 1941–1948.

Edited by J. Karn

(Received 7 August 1995; accepted 26 September 1995)