Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 368 (2008) 606–613
BBRC
www.elsevier.com/locate/ybbrc

Initiation of the microgene polymerization reaction with non-repetitive homo-duplexes

Mark Itsko [a,*], Arieh Zaritsky [a], Avinoam Rabinovitch [b], Eitan Ben-Dov [c,d]

[a] Department of Life Sciences, Faculty of Natural Sciences, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 84105, Israel
[b] Department of Physics, Faculty of Natural Sciences, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 84105, Israel
[c] Department of Biotechnology Engineering, Faculty of Engineering Sciences, Ben-Gurion University of the Negev, PO Box 653, Be'er-Sheva 84105, Israel
[d] Achva Academic College, MP Shikmim, 79800, Israel

Received 19 January 2008
Available online 1 February 2008

Abstract

Microgene Polymerization Reaction (MPR) is used as an experimental system to artificially simulate evolution of short, non-repetitive homo-duplex DNA into multiply-repetitive products that can code for functional proteins. Blunt-end ligation by DNA polymerase is crucial in expansion of homo-duplexes (HDs) into head-to-tail multiple repeats in MPR. The propagation mechanism is known, but formation of the initial doublet (ID) by juxtaposing two HDs and polymerization through the gap has been ambiguous. Initiation events with pairs of HDs using Real-Time PCR were more frequent at higher HD concentrations and slightly below the melting temperature. A process molecularity of about 3.1, calculated from the amplification efficiency and the difference in PCR cycles at which propagation was detected at varying HD concentrations, led to a simple mechanism for ID formation: the gap between two HDs is bridged by a third. Considering thermodynamic aspects of the presumed intermediate "nucleation complex" can predict relative propensity for the process with other HDs.

Keywords: Blunt-end ligation by DNA polymerase; Head-to-tail primer polymerization; Repeated motifs; PCR artifacts; Molecularity

Generation of head-to-tail tandem repeats in PCR from a pair of oligonucleotides with a complementary region (homo-duplexes containing overhangs) was originally described a decade ago [1] and termed Microgene Polymerization Reaction (MPR). Expansion of DNA repeats is of great scientific interest because it is associated with a variety of neurological diseases [2–4] and presumably with the mechanism for gene evolution [5–7]. MPR is composed of two stages: initiation (Fig. 1A), forming an initial doublet (ID) of the homo-duplex (HD) (Fig. 1B), and propagation (Fig. 1C), leading to expansion of the ID into multiple head-to-tail repeats. The proposed mechanism [8] for the propagation stage involves staggered re-annealing of the repeat-containing DNA duplex, leading to overhangs that are replicated by the DNA polymerase (Fig. 1C). Propagation is more effective at temperatures closer to the melting point of the starting repeat-containing duplex [8]. The maximal end fraying and bulge forming at this temperature allow transient intermolecular base pairing (and sliding) between different sized duplexes in solution and their effective propagation. This process has thoroughly been studied under isothermal conditions [8–11] but it is analogous to PCR-related behaviour. The initiation stage implies existence of a most intriguing "illegitimate" activity of DNA polymerase allowing it to skip the gap between two discontinuous templates. To explain the initiation process, Shiba et al. [1] postulated that a molecule of DNA polymerase juxtaposes such templates and proceeds over the gap, either by itself or assisting another molecule to skip. DNA polymerases without proofreading such as Klenow fragment and *Taq* polymerase indeed serve as juxtaposing proteins themselves [12,13] and others replicate even across non-DNA segments [14], but the

* Corresponding author. Fax: +972 8 6278 951.
E-mail address: ariehz@bgu.ac.il (M. Itsko).

0006-291X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.01.108

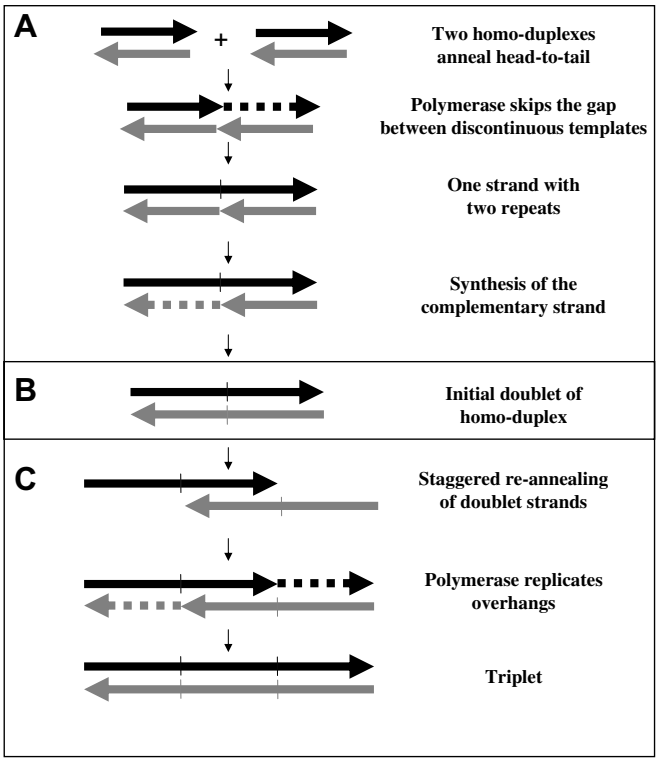

Fig. 1. A model for initial doublet generation from homo-duplex and propagation of repeats in MPR. (A) Initiation; (B) initial doublet; (C) propagation to long repeats in heat/cool cycles.

mechanism of initiation through such a bridge is still enigmatic.

Various homo-duplexes (HDs) demonstrate MPR ability with different efficiencies under long PCR conditions (at least 65 cycles; data not shown). Head-to-tail tandem repeats have recently been formed using the Quick Change protocol [15] with pairs of complementary oligonucleotides that do not contain any repeated motif [16]. MPR in this system has never been reported before, to the best of our knowledge, but when occasionally observed (unpublished data), it usually is ignored, and the "undesired" outcome is discarded. Here, such HDs were used as a system for consistent study of the MPR. Using Real-Time PCR with several HDs to follow the repeat propagation led us to unravel a simple mechanism for an ID formation in which juxtaposition of two discontinuous templates is accomplished not by DNA polymerase but rather by a DNA-HD.

Materials and methods

End-point detection PCR. A 25-μl reaction mixture contained 4–280 pmol of the HD-forming complementary strands (primers; Table 1) and either (a) 100–500 μM of each dNTP, ThermoPol buffer with 8 mM $MgSO_4$ and 1 U of Vent DNA polymerase, or (b) 100–1000 μM of each dNTP, Super-Therm Polymerase buffer with 8 mM $MgCl_2$ and 2.5 U of Taq DNA polymerase. The following conditions for T-Gradient Thermoblock cycler (Biotron, Germany) were employed: 10 min at 94 °C and 10 min at 72 °C, then 65 cycles of 94 °C for 10 s and 72 °C for 4 min. Amplicons obtained with 140 and 280 pmol primer were purified by Amersham kit, A-tailed by Taq DNA polymerase in the presence of 0.2 mM dATP at 70 °C during 30 min, and cloned for sequencing to pGEM-T easy vector (Promega).

Real-Time (RT) PCR. DNA obtained during RT-PCR with each of the primers (Table 1) was quantified in ABI Prism 7000 Sequence Detection System (Applied Biosystems), using Absolute QPCR SYBR Green ROX Mix (ABgene, Surrey, UK) in a 96-wells optical plate. Reaction mixture, in duplicates, consisted of 10 μl of the "Absolute Mix" with 100 μM of each dNTP and 10 μl of the appropriate primer (0.8–6.4 pmol). The following thermal cycling conditions were employed: 2 min at 50 °C, 15 min at 95 °C, followed by 65 rounds of 15 s at 95 °C and 3 min at 72 °C. DNA was detected by FRET with a SYBR Green fluorophore. The PCR products were verified with ethidium bromide-stained 0.8% and 2.5% agarose gels (SeaKemR LE Agarose; FMC BioProducts, Rockland, ME). HDs' T_m were determined using a dissociation protocol of the RT-PCR cycler.

Results

Concentration-dependence of MPR

When PCR was performed with EVNA primers (HD-forming complementary strands) (Table 1) using Vent DNA polymerase, a smear was obtained, reflecting heterogeneous amplicon sizes. Increased initial primer concentration led to decrease of the mean product size: much longer DNA than 10 kb at 0.16 μM, and shorter than 500 bp at 11.2 μM (Fig. 2A). The smear was resolved on 2.5% agarose gel to DNA bands with discrete lengths of ~42 bp-multiples (Fig. 2B). Similar results, albeit with bands that were less sharp, were obtained with Taq DNA polymerase (not shown). Higher primer concentration seems to promote more frequent generation of initial doublets, IDs (Fig. 1A and B), propagation of which is restricted by supply of dNTPs. The inverse dependence of the chain length on the initiation frequency, well known in polymer chemistry [17], is reflected here by the limited repeat-expansion due to a low [dNTP] to [HD] ratio. This was confirmed by performing MPR at different [dNTP]: as total

Table 1
The homo-duplexes used

Name	Sequence[a]	GC (%)	T_m (°C)[b]
EVNA	5′-GGTGATAGAAG**TG**CTTAAATCTTTATTAGGAATT**GCT**CTGGC-3′ 3′-CCACTATCTTC**AC**GAATTTAGAAATAATCCTTAA**CGA**GACCG-5′	38.1	73.7
NOMU	5′-GGTGATAGAAG**AA**CTTAAATCTTTATTAGGAATT**AAC**CTGGC-3′ 3′-CCACTATCTTC**TT**GAATTTAGAAATAATCCTTAA**TTG**GACCG-5′	33.3	70.8

[a] Bold type letters indicate differences in primer composition between EVNA and NOMU dimers.
[b] Derived from Fig. 3C.

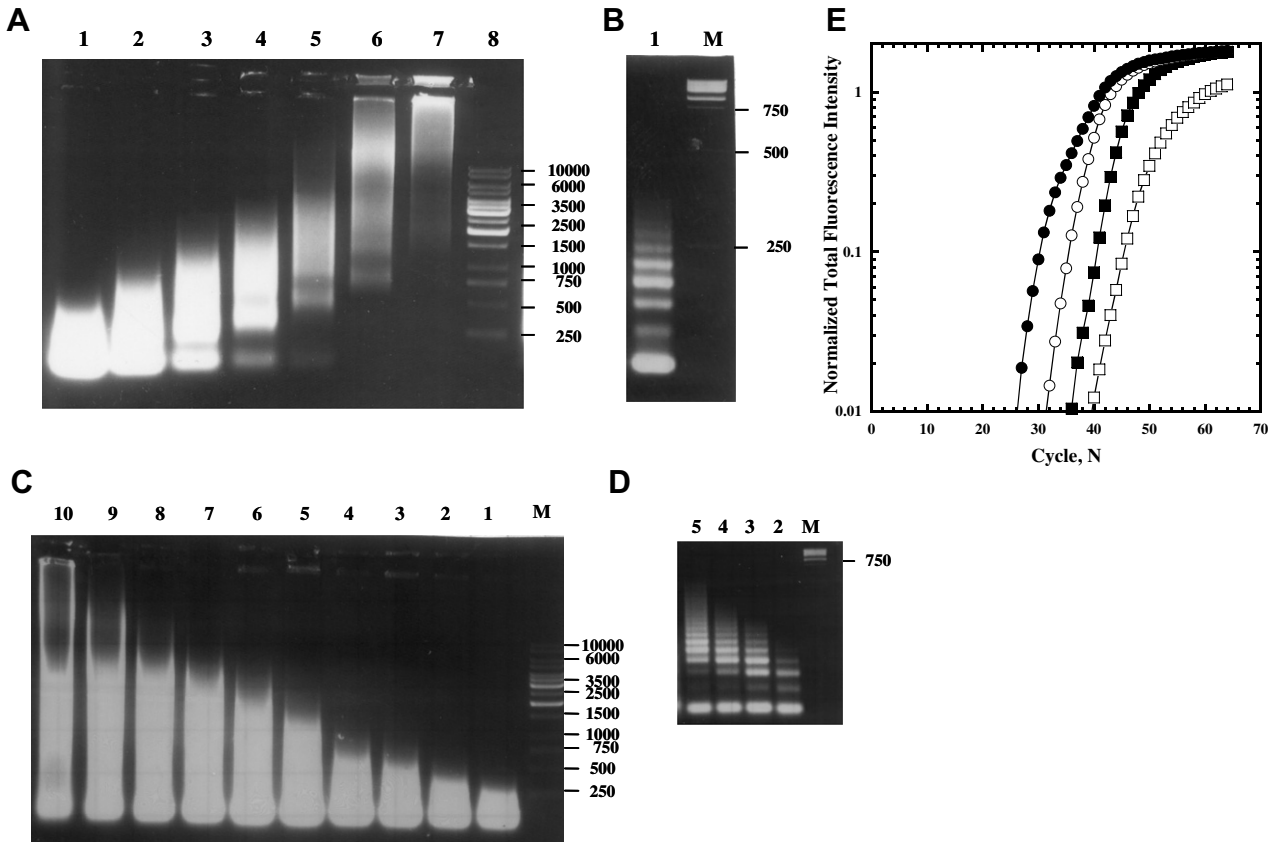

Fig. 2. Concentration-dependence of MPR with EVNA homo-duplex. Change of product length with concentrations of EVNA with Vent (A and B) and of dNTP (C with Taq, D with Vent) on 0.8% (A and C) and 2.5% (B and D) agarose gels. (A and B) Lanes 1–7 correspond to 10.4, 5.2, 2.6, 1.3, 0.64, 0.32, 0.16 μM, respectively, of EVNA. Lane 8, DNA ladder. (C and D) The following concentrations of each dNTP were used (in μM): lane 1 – 100; lane 2 – 200; lane 3 – 300; lane 4 – 400; lane 5 – 500; lane 6 – 600; lane 7 – 700; lane 8 – 800; lane 9 – 900; lane 10 – 1000. Concentration of the EVNA homo-duplex, 5.6 μM. (E) Relative total MPR products generated by RT-PCR with the following initial EVNA homo-duplex concentrations (in μM): closed circles, 0.32; open circles, 0.16; closed squares, 0.08; open squares, 0.04.

concentration increased from 100 to 1000 μM with Taq polymerase (Fig. 2C) and from 200 to 500 μM with Vent polymerase (Fig. 2D), the repetitive DNA lengthened. It is noteworthy that Vent polymerase is inactive at dNTP concentrations higher than 500 μM (data not shown).

To obtain MPR products with a small number of HD repeats, Vent polymerase was used with a high concentration (about 10 μM) of HDs. Of the 19 amplicons (with head-to-tail polymerized HDs) isolated under these conditions (two examples are shown in Fig. 4F and G), 5 were composed of 2 HD repeats, 3 of 3 HDs, 7 of 4 HDs, 1 of 5 HDs, 2 of 6 HDs and 1 with 7 HDs. These 19 multiply-repeated sequences thus contained total of 71 repetitions and 52 junctions. Nine junctions (17%) were exact, 18 (35%) contained deletions of one nucleotide, 10 (19%) contained deletions of two nucleotides and 15 (29%) contained deletions of three nucleotides, identical in the latter two cases. No additions were observed at the junctions.

Table 2
Experimentally obtained main MPR characteristics

HD nM[a] (10^{11} × molecules)	*N*th at products detection	Δ*N*th[b]	*E*[c]	Number of ID molecules at $N = 1$[d]
320 (40)	26 ± 2		0.8 ± 0.05	88×10^3–88×10^4
		5		
160 (20)	31 ± 2		0.8 ± 0.04	6.0×10^3–6.0×10^4
		3		
80 (10)	34 ± 2		0.68 ± 0.16	1.2×10^3–1.2×10^4
		4		
40 (5)	38 ± 2		0.57 ± 0.087	0.14×10^3–0.14×10^4

[a] Initial concentration (number of molecules) of homo-duplexes.
[b] The difference between two *N*th's obtained with 2-fold-different initial HD.
[c] Derived from Fig. 2E.
[d] Estimated from $X_N = \mathrm{ID} \times (1 + E)^N$ using average *E* of 0.71, *N*th value and the total number of repeats at *N*th (10^{11}–10^{12}).

The initial concentration of primers C_0 determines the cycle number (the operative threshold cycle) Nth (in cycle units N) at which the signal (the microgene expansion process) was first detected. A 2-fold decrease in C_0 delayed Nth by approximately 4 cycles (i.e., $\Delta N\text{th} = 4$; Fig. 2E, Table 2). Following the initiation process, total MPR products (X_N) are propagated exponentially, with an average amplification efficiency of about 0.7 (between 0.8 and 0.57; Table 2). The concentration of amplified molecules at cycle N during the exponential MPR propagation X_N was approximated by $X_N = \text{ID} \times (1 + E)^N$, where ID is the initial concentration of doublets (ID) and E is the amplification efficiency per cycle.

Temperature-dependence of MPR

A slight change in primer sequence affected the MPR efficiency: the five base changes in EVNA converting it to NOMU HD (Table 1) drastically decreased the efficiency of MPR (Fig. 3A and B). Microgene production with NOMU was detected 12 cycles later than with EVNA in RT-PCR under identical conditions (Fig. 3A). Consistently, repeats in end-point detection PCR could only be generated with NOMU concentrations 8-fold higher than with EVNA (Fig. 3B). This difference results most likely from the small difference in melting temperatures (T_m) between the homo-duplexes (as a consequence of their different GC contents; Table 1): 73.7 °C for EVNA, 70.8 °C for NOMU (Fig. 3C). At 72 °C, the extent of pairing between the complementary strands of NOMU was approximately 30%, whereas that of EVNA was about 70% (Fig. 3C). Decreasing the annealing temperature of NOMU to 69 °C allowed repeat generation by concentrations 4-fold lower, comparable with EVNA at 72 °C (compare Fig. 3B with Fig. 3D).

Discussion

The initiation process

The major hitherto unresolved question regarding the MPR phenomenon is the mechanism of its initiation. The

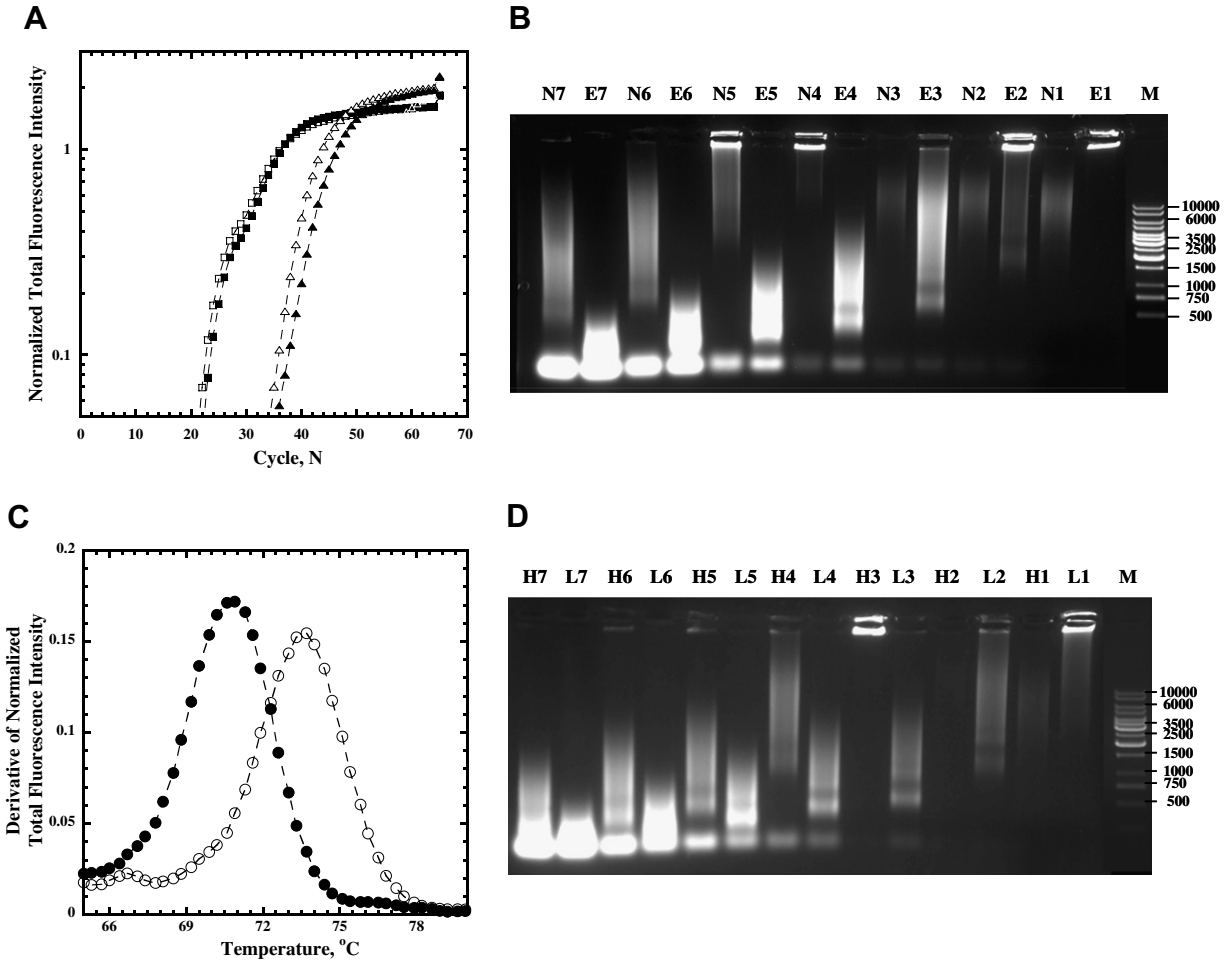

Fig. 3. Temperature-dependence of MPR. Relative total MPR products generated at 72 °C by (A) RT-PCR with EVNA (squares) and NOMU (triangles) homo-duplexes, both at 0.32 μM (closed and open symbols indicate corresponding duplicates), and by (B) end-point detection PCR with variable concentrations of EVNA (E) and NOMU (N). Lanes E(N)1–E(N)7, homo-duplex concentrations at 0.16, 0.32, 0.64, 1.3, 2.6, 5.2, and 10.4 μM, respectively. (D) The same respective concentrations of NOMU at 72 °C (lanes H1–H7), and 69 °C (L1–L7). (C) Derivatives of fluorescence intensities of EVNA (open circles) and NOMU (closed circles) (5.2 μM each) at different temperatures (dissociation protocol of RT-PCR).

high MPR ability of EVNA [16] resulted from the fact that its T_m (73.7 °C) is just about 2 °C above the elongation temperature (72 °C) (Fig. 3C). On the other hand, the lower melting temperature of NOMU (70.8 °C) does not allow effective initiation of MPR at 72 °C (Fig. 3A and B). Optimal initiation (as well as of propagation) [8] of a multiply-repeated product is achieved at a temperature slightly below its T_m (Fig. 3C and D).

The likelihood of initiation thus depends on the concentration of the homo-duplex C_0 rather than of its component primers: it is higher when the extent of pairing is larger, at temperatures below the T_m of the homo-duplex (Table 1 and Figs. 2E and 3), which precludes the alternative possibility. A third pair of primers with a higher T_m (78 °C) confirmed this conclusion (data not shown). The generation-rate of the initial doublet ID is assumed to be proportional to the power m (molecularity) of the concentration of the original homo-duplex HD ($dID/dN \propto C_0^m$) (Fig. 1A and B). The amplification efficiency during the exponential propagation (0.71 per cycle) is much higher than that of the initiation, and the consumption of HD molecules in the initiation process is negligible compared to that during the propagation stage. The initiation process would therefore not affect C_0, hence ID would be $\propto N \times C_0^m$. If the number of ID molecules is large, the MPR initiation is reasonably assumed (see below) to occur in all HD at the first cycle $N = 1$ (though the low sensitivity of RT-PCR detects the products at a later stage, i.e., at the Nth).

The experiments were performed with a series of 2-fold-C_0 values (Fig. 2E). The predicted ratio between ID of two successive dilutions is thus

$$\frac{ID_2}{ID_1} = \frac{(2 \times C_0)^m}{C_0^m} = 2^m \quad (1)$$

The exponential propagation of MPR total product (X_N) following generation of ID depends (as stated before) on the number of cycles and amplification efficiency: $X_N = ID \times (1 + E)^N$. If E (derived from the slope in each line of Fig. 2E) remains constant, the ratio between successive ID values in the RT-PCR experiments with 2-fold different initial concentrations of primers may be derived at points with equal amount of X_N, as for example the threshold points (Fig. 2E) according to

$$\frac{ID_2}{ID_1} = (1 + E)^{(N\mathrm{th}_1 - N\mathrm{th}_2)} \quad (2)$$

and the molecularity m of the reaction can be derived by equating Eq. (1) with Eq. (2):

$$m = (N\mathrm{th}_1 - N\mathrm{th}_2) \times \frac{\ln(1 + E)}{\ln 2} \quad (3)$$

The observed values of E (Table 2) however, do vary slightly (0.8–0.6) between successive lines (Fig. 2E). A molecularity of 3.1 (ranging between 2.6 and 3.4) for the initiation process was estimated, using the average E value (0.71). It can thus be concluded that three HDs must somehow interact to initiate doublet formation in the MPR.

The assumption used to derive Eq. (2), that initiations occur at $N = 1$ in all HD, is justified as follows. The absolute number of HD molecules in the 20 μl mixtures (with 40–320 nM) is $5–40 \times 10^{11}$. This number determines background fluorescence intensity in the mixtures. The value of Nth is determined by the cycle at which the fluorescence of the total number of repeats in the MPR products overgrows that of the background, which corresponds to 10^{11}–10^{12} repeats. Backwards extrapolations, using average E of 0.71 and Nth values yield the number of ID molecules at the first cycle ($N = 1$) in each case (Table 2). The calculated numbers are at about 10^2–10^6 molecules.

The model, consequences and implications

We propose a simple mechanism for the generation of ID: a rare and reversible association between three homo-duplexes (HD) generates a nucleation complex (NC), (Fig. 4A and C) which converts to ID according to $3HD \overset{I}{\leftrightarrow} NC \overset{II}{\rightarrow} ID + HD$, where I and II denote first (rate limiting) and second (enzymatic) stage in the MPR initiation process. One of these three HDs (labeled in gray) aligns and bridges the other two, fixing them in the required proximity for the DNA polymerase to skip the inter-template gap (Fig. 1A). This bridging occurs through occasional Watson–Crick bonds between aligning and aligned homo-duplexes.

Additional Watson–Crick bonds between 3′ end bases of the aligned homo-duplexes may assist in the generation of NC and in turn explain the deletion of 2 or 3 nucleotides at EVNA primer junctions in MPR products (Fig. 4D–G).

The extent of pairing between the complementary strands of EVNA (70%) was only about 2.5-fold higher than that of NOMU (30%) (Fig. 3C). Such a small difference cannot *per se* explain the difference of 12 cycles in their MPR threshold cycles Nths (Fig. 3A), yielding a ratio of about 625 ($\approx 1.7^{12}$) between EVNA and NOMU ID's (after substituting the average E in Eq. (2)). However, different thermodynamic stabilities of the putative NCs (expressed in ΔG^0 [$= -RT\ln K_{eq}$] of NC formation) of EVNA and NOMU HDs may underlie their different propensities to initiate MPR. Since the dissociation rate of NC is much faster than its rate of conversion to ID (due to low likelihood of its formation), the limiting, first stage of the initiation process is situated near its equilibrium. One can derive the difference between the free energies of NC formation from the ratio between equilibrium constants of this stage in the EVNA and NOMU cases: $K_{eq} = [NC]/[HD]^3$. Since [NC] is negligible compared to [HD], when C_0 of EVNA and NOMU is identical:

$$\frac{K_{eq(EVNA)}}{K_{eq(NOMU)}} = \frac{[NC]_{EVNA}}{[NC]_{NOMU}} = \frac{[ID]_{EVNA}}{[ID]_{NOMU}} = 625, \text{and hence}$$

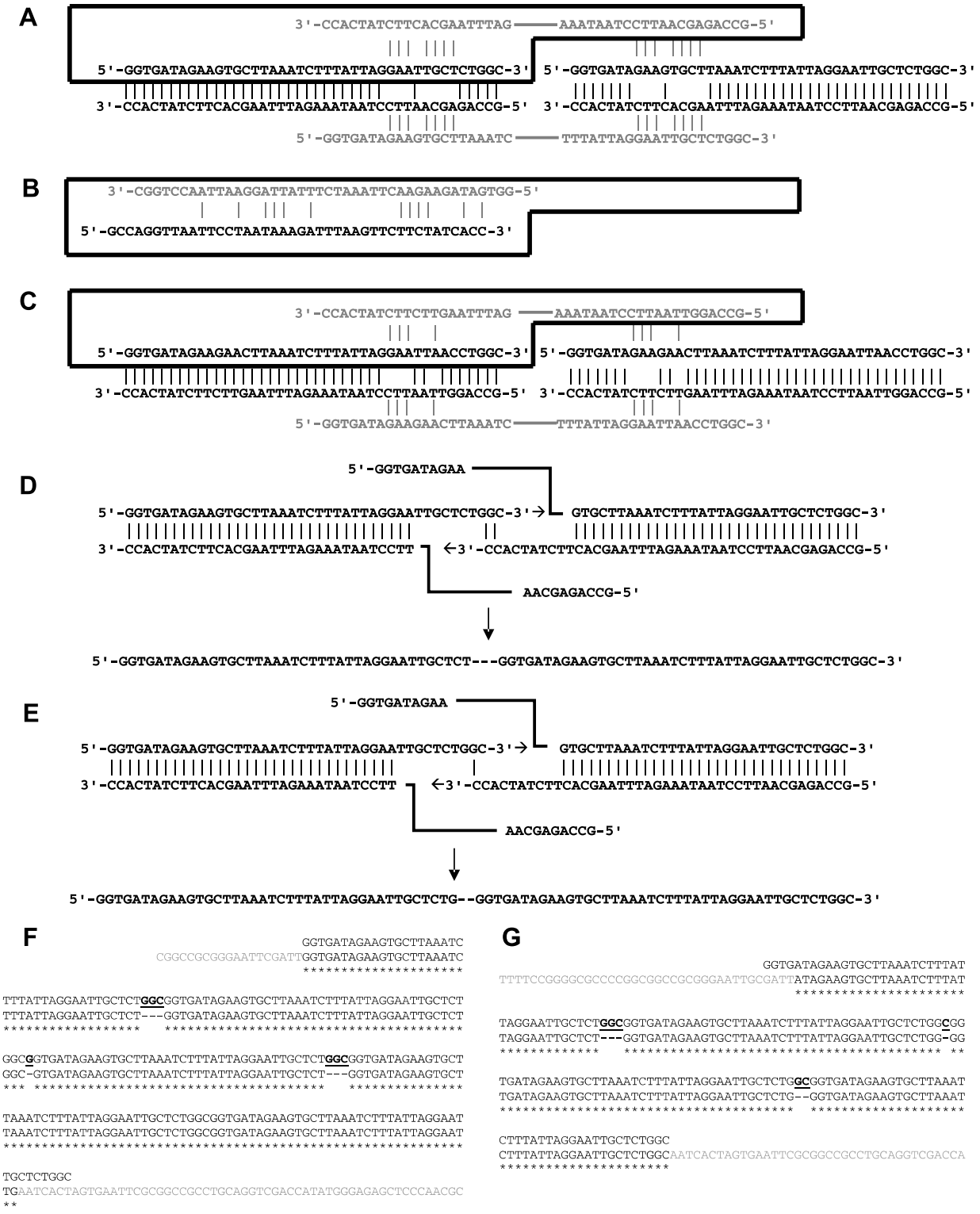

Fig. 4. Model for ID generation. Nucleation complexes of the primers EVNA (A) and NOMU (C), and an alternative (ineffective) structure of NOMU (B). Symbols in gray font indicate the primers (that align the juxtaposed HDs in A and C) and the hydrogen bonds involved. Putative arrangements of the juxtaposed EVNA within NC resulting in ID with 3 (D) and 2 (E) nucleotide deletions at the junctions are shown. Two examples of MPR products, with 5 (F) and 4 (G) repeats, obtained with EVNA (bottom sequences) and compared with head-to-tail EVNA primer repeats (top sequences). Gray font, sequence of vector pGEM; underlined nucleotides correspond to deletions.

$$\Delta\Delta G^0 = \Delta G^0_{\mathrm{EVNA}} - \Delta G^0_{\mathrm{NOMU}} = -RT \ln \frac{[\mathrm{ID}]_{\mathrm{EVNA}}}{[\mathrm{ID}]_{\mathrm{NOMU}}} = -4.4\ \mathrm{kcal/mol}$$

The two strands of a DNA molecule can generate the alternative partially complementary secondary structures with varying free energies of formation [18]. The most stable, fully complementary structures of EVNA and NOMU are of −78 and −74 kcal mol^{-1}, respectively (calculated with Oligo Analyzer 1.1.2) [19]. The free energies of the second-ranked structures are of, respectively, −4.68 and −2.35 kcal mol^{-1}. This EVNA structure (framed panel in Fig. 4A) but not that of NOMU (framed panel in Fig. 4B) can bridge a third HD to form an NC. The structure of NOMU that does allow such analogous NC (framed panel in Fig. 4C) is of −0.62 kcal mol^{-1} only. The difference between these free energies $\Delta\Delta G_0$ (−4.68 + 0.62) is −4.06 kCal mol^{-1}, which resembles the value obtained from the experimental results and the considerations of the model presented here.

The most abundant among three microgenic units in the combinatorial polymers recently generated from an equimolar mixture [20] is the one with the lowest free energy of formation of the NC-promoting second-ranked structure (analysis not shown), thus supporting the model proposed and described here.

This model requires neither involvement of two molecules of DNA polymerase to bridge the gap between two HDs to form an ID (model (i)) [1] nor a "putative repair mode of the enzyme" (model (ii)) [1]. It does propose a mechanism by which "a 5′-OH end of one template and a 3′-OH end of another" [1] are juxtaposed without direct involvement of the DNA polymerase (Fig. 4A and C).

A totally different observation, repair of chromosomal double-strand breaks in yeast cells, has recently led Resnick et al. to suggest that single-stranded oligo-deoxy-ribonucleotides [21,22] as well as oligo-ribonucleotides [23] complementary to the break-ends bridge the gap between the ends and hence provide a template to restore their contiguity. Kinetic analysis (not shown) of an analogous model, in which holding together two HDs (break-ends in their case) by one of its components (single-stranded oligo-deoxyribonucleotide), leads to molecularity of 2.5 rather than 3 obtained here. It is therefore very unlikely that the *in vitro* blunt-end ligation activity of DNA polymerase discovered here is performed through a single-stranded bridge.

The MPR-based technique allows relatively simple and fast creation of libraries of artificial genes coding for proteins containing multiple tandem motifs and has been successfully used to derive putative useful properties [24]. Mixing a number of microgene units in one reaction tube significantly enlarges the diversity of these libraries owing to differently shuffled peptide motifs encrypted in the products [20,25].

The numerous periodicities in genomic DNA, also encrypted in the encoded proteins, suggest that they derive from primordial oligomeric repeats [5–7]. MPR is used as an experimental system to artificially simulate evolution of short, non-repetitive homo-duplex DNA into multiply-repetitive products [26] that can code for functional proteins, which occasionally are the cause for certain diseases [2–4].

These empirical MPR phenomena can now be explained by the mechanism presented here for the essential, so far ambiguous event of this reaction, its initiation. Our results present a systematic approach to understand this event, culminating in a model that needs no accessory proteins and allows prediction of the relative propensities of different HDs to form an ID, essential for repeat propagation in MPR. It allows, moreover, to design and synthesize polypeptides with certain motifs multiplied within the frame of a given protein, as has already been implemented [16].

Acknowledgments

Thanks are due to David M. Chipman for critically reading the manuscript. This investigation was partially supported by Eshkol Fellowship, Ministry of Science, Jerusalem, Israel (to M.I.), and by two Grants, one from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel (No. 2001-042), and a seed grant from the Vice President for Research at Ben-Gurion University of the Negev and (both to A.Z.).

References

[1] K. Shiba, Y. Takahashi, T. Noda, Creation of libraries with long ORFs by polymerization of a microgene, Proc. Natl. Acad. Sci. USA 94 (1997) 3805–3810.
[2] G. Bates, H. Lehrach, Trinucleotide repeat expansions and human genetic disease, Bioessays 16 (1994) 277–284.
[3] S. Panzer, D.P. Kuhl, C.T. Caskey, Unstable triplet repeat sequences: a source of cancer mutations? Stem Cells 13 (1995) 146–157.
[4] G.R. Sutherland, R.I. Richards, Simple tandem DNA repeats and human genetic disease, Proc. Natl. Acad. Sci. USA 92 (1995) 3636–3641.
[5] S. Ohno, Birth of unique enzyme from an alternative reading frame of the preexisted, internally repetitious coding sequence, Proc. Natl. Acad. Sci. USA 81 (1984) 2421–2425.
[6] S. Ohno, Evolution from primordial oligomeric repeats to modern coding sequences, J. Mol. Evol. 25 (1987) 325–329.
[7] A.K. Bjorklund, D. Ekman, A. Elofsson, Expansion of protein domain repeats, PLoS Comput. Biol. (2006) 2 (8) e114.
[8] W. Tuntiwechapikul, M. Salazar, Mechanism of in vitro expansion of long DNA repeats: effect of temperature, repeat length, repeat sequence, and DNA polymerases, Biochemistry 41 (2002) 854–860.
[9] N. Ogata, T. Miura, Elongation of tandem repetitive DNA by the DNA polymerase of the hyperthermophilic archaeon *Thermococcus litoralis* at a hairpin-coil transitional state: a model of amplification of a primordial simple DNA sequence, Biochemistry 39 (2000) 13993–14001.
[10] N. Ogata, H. Morino, Elongation of repetitive DNA by DNA polymerase from a hyperthermophilic bacterium *Thermus thermophilus*, Nucl. Acids Res. 28 (2000) 3999–4004.
[11] X. Liang, K. Jensen, M.D. Frank-Kamenetskii, Very efficient template/primer-independent DNA synthesis by thermophilic DNA polymerase in the presence of a thermophilic restriction endonuclease, Biochemistry 43 (2004) 13459–13466.

[12] J.S. King, C.F. Fairley, W.F. Morgan, DNA end joining by the Klenow fragment of DNA polymerase I, J. Biol. Chem. 271 (1996) 20450–20457.
[13] J.S. King, C.F. Fairley, W.F. Morgan, Bridging the gap, J. Biol. Chem. 269 (1994) 13061–13064.
[14] A. Maor-Shoshani, V. Ben-Ari, Z. Livneh, Lesion bypass DNA polymerases replicate across non-DNA segments, Proc. Natl. Acad. Sci. USA 100 (2003) 14760–14765.
[15] QuickChange Site-Directed Mutagenesis Kit. Available from: <http://www.stratagene.com/manuals/200518.pdf>.
[16] M. Itsko, A. Zaritsky, Exposing cryptic antibacterial activity in Cyt1Ca from *Bacillus thuringiensis israelensis* by genetic manipulations, FEBS Lett. 581 (2007) 1775–1782.
[17] P.J. Flory, Principles of Polymer Chemistry, Cornell University Press, Ithaca, 1953, pp. 132–148.
[18] J. SantaLucia Jr., D. Hicks, The thermodynamics of DNA structural motifs, Annu. Rev. Biophys. Biomol. Struct. 33 (2004) 415–440.
[19] Free software. www.genelink.com.
[20] H. Saito, T. Minamisawa, K. Shiba, Motif programming: a microgene-based method for creating synthetic proteins containing multiple functional motifs, Nucl. Acids Res. 35 (6) (2007) e38.
[21] F. Storici, C.L. Durham, D.A. Gordenin, M.A. Resnick, Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast, Proc. Natl. Acad. Sci. USA 100 (2003) 14994–14999.
[22] F. Storici, J.R. Snipe, G.K. Chan, D.A. Gordenin, M.A. Resnick, Conservative repair of a chromosomal double-strand break by single-strand DNA through two steps of annealing, Mol. Cell. Biol. 26 (2006) 7645–7657.
[23] F. Storici, K. Bebenek, T.A. Kunkel, D.A. Gordenin, M.A. Resnick, RNA-templated DNA repair, Nature 447 (2007) 338–341.
[24] K. Shiba, MolCraft: a hierarchical approach to the synthesis of artificial proteins, J. Mol. Catal. B 28 (2004) 145–153.
[25] H. Saito, T. Honma, T. Minamisawa, K. Yamazaki, T. Noda, T. Yamori, K. Shiba, Synthesis of functional proteins by mixing peptide motifs, Chem. Biol. 11 (2004) 765–773.
[26] K. Shiba, Y. Takahashi, T. Noda, On the role of periodism in the origin of proteins, J. Mol. Biol. 320 (2002) 833–840.