Directed evolution of polymerase function by compartmentalized self-replication

Farid J. Ghadessy, Jennifer L. Ong, and Philipp Holliger†

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

Communicated by Max F. Perutz, Medical Research Council, Cambridge, United Kingdom, February 1, 2001 (received for review December 15, 2000)

We describe compartmentalized self-replication (CSR), a strategy for the directed evolution of enzymes, especially polymerases. CSR is based on a simple feedback loop consisting of a polymerase that replicates only its own encoding gene. Compartmentalization serves to isolate individual self-replication reactions from each other. In such a system, adaptive gains directly (and proportionally) translate into genetic amplification of the encoding gene. CSR has applications in the evolution of polymerases with novel and useful properties. By using three cycles of CSR, we obtained variants of *Taq* DNA polymerase with 11-fold higher thermostability than the wild-type enzyme or with a >130-fold increased resistance to the potent inhibitor heparin. Insertion of an extra stage into the CSR cycle before the polymerase reaction allows its application to enzymes other than polymerases. We show that nucleoside diphosphate kinase and *Taq* polymerase can form such a cooperative CSR cycle based on reciprocal catalysis, whereby nucleoside diphosphate kinase produces the substrates required for the replication of its own gene. We also find that in CSR the polymerase genes themselves evolve toward more efficient replication. Thus, polymerase genes and their encoded polypeptides cooperate to maximize postselection copy number. CSR should prove useful for the directed evolution of enzymes, particularly DNA or RNA polymerases, as well as for the design and study of *in vitro* self-replicating systems mimicking prebiotic evolution and viral replication.

Polynucleotide polymerases occupy a central role in the maintenance, transmission, and expression of genetic information (1). They also have enabled core technologies of molecular biology like sequencing, PCR, site-directed mutagenesis, and cDNA cloning. However, polymerases available from nature are often not optimally suited for these applications and attempts have been made to tailor polymerase function by using either design or selection strategies. Structural studies have greatly advanced our understanding of polymerase function (2–4) and together with insights gained from site-directed mutagenesis have allowed the rational design of some polymerase variants with improved properties. Among these are polymerases with improved dideoxynucleotide incorporation for cycle-sequencing (5) or increased processivity (6). Other "designer" polymerases include truncation variants (7, 8), some of which show improved thermostability and fidelity, although at the cost of reduced processivity. Despite these advances, our ability to engineer designer polymerases for specific applications remains limited.

Repertoire selection methods have proven to be an effective means to obtain biopolymers with desired properties. Polymerases have been selected successfully for activity by phage display (9) and by complementation of a DNA polymerase I-deficient *Escherichia coli* strain (10). Careful screening of the complementing polymerase mutants yielded polymerases with a range of interesting properties such as altered fidelity (11). However, *in vivo* complementation strategies are limited in that selection conditions, and hence selectable phenotypes, cannot be chosen freely and are constrained further by limits of host viability.

We reasoned that a strategy allowing a direct interrogation of polymerase activity under a range of conditions (e.g., high temperatures, presence of inhibitors) would greatly benefit our ability to study polymerase function. Furthermore, it could be expected to be of considerable practical importance, as it should allow the directed evolution of designer polymerases tailor-made for both existing applications like PCR and sequencing, as well as novel applications in emerging technologies like DNA computing and nanotechnology.

We have developed a strategy, called compartmentalized self-replication (CSR), for the evolution of polymerases. In CSR, individual polymerase variants are isolated in separate compartments. Provided with nucleoside triphosphates and appropriate flanking primers, each polymerase replicates only its own encoding gene to the exclusion of those in other compartments (Fig. 1). Consequently, only genes encoding active polymerases are replicated, whereas inactive variants fail to amplify their own genes and disappear from the gene pool. Among differentially active variants, the more active can be expected to produce proportionally more "offspring," correlating postreplication copy number with enzymatic turnover. Reactions take place in individual aqueous compartments of a water-in-oil emulsion that are stable for prolonged periods at temperatures exceeding 90°C. This approach allows selection for enzymatic activity under a wide range of conditions. We foresee many applications of this technology.

As an initial test, we have applied CSR to the isolation of designer polymerases directly from a diverse repertoire of mutant polymerases. As a model system, we have used the DNA polymerase I from *Thermus aquaticus* (*Taq* polymerase), the enzyme most commonly used for sequencing and PCR.

Materials and Methods

DNA Manipulation and Protein Expression. *Taq* polymerase was amplified from *T. aquaticus* by using primers 1 (5′-GGC GAC TCT AGA TAA CGA GGG CAA AAA ATG CGT GGT ATG CTT CCT CTT TTT GAG CCC AAG GGC CGC GTC CTC CTG-3′) and 2 (5′-GCG GTG CGG AGT CGA CTC ACT CCT TGG CGG AGA GCC AGTC-3′), cut with *Xba*I and *Sal*I and cloned into pASK75 (12). Constructs were transformed into *E. coli* TG1 and expressed as described (12). Cells were harvested and washed once in 1× *Taq* buffer [50 mM KCl/10 mM Tris·HCl (pH 9.0)/0.1% Triton X-100/1.5 mM $MgCl_2$]. For purification, *Taq* clones were subcloned with an N-terminal hexahistidine tag by using primers 2 and 3 (5′-GGC GGC TCT AGA TAA CGA GGG CAA AAA ATG CAT CAT CAT CAC CAT CAC GGT GCC ATG GCT CTT CCT CTT TTT GAG CCC AAG GGC-3′) expressed as above, lysed as described (13), and purified by using Ni-NTA Agarose (Qiagen, Chatsworth, CA). Error-prone PCR mutagenesis was used to generate libraries L1 (14) and L2 (15).

Abbreviations: wt*Taq*, wild-type *Taq* polymerase; sf*Taq*, *Taq* polymerase Stoffel fragment; CSR, compartmentalized self-replication; NDK, nucleoside diphosphate kinase.

†To whom reprint requests should be addressed. E-mail: ph1@mrc-lmb.cam.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

www.pnas.org/cgi/doi/10.1073/pnas.071052198

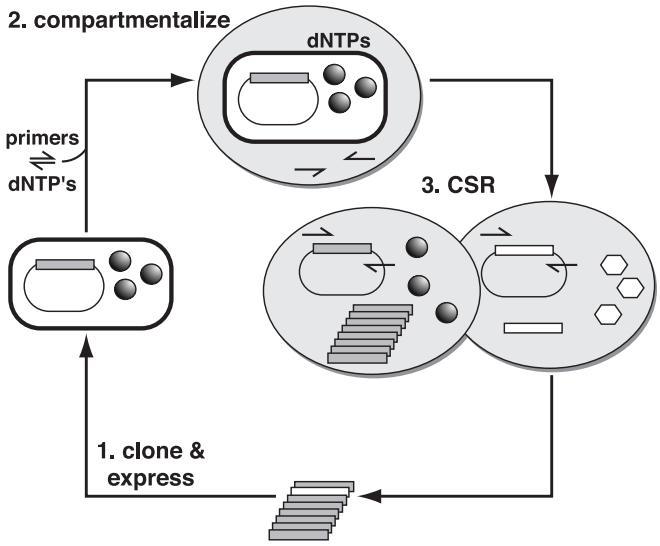

Fig. 1. General scheme of CSR. (*1*) A repertoire of polymerase genes is cloned and expressed in *E. coli*. Spheres represent active polymerase molecules. (*2*) Bacterial cells containing the polymerase and encoding gene are suspended in reaction buffer containing flanking primers and dNTPs and segregated into aqueous compartments. (*3*) The polymerase enzyme and encoding gene are released from the cell, allowing self-replication to proceed. Poorly active polymerases (white hexagons) fail to replicate their encoding gene. The offspring polymerase genes are released and recloned for another cycle of CSR.

Emulsification of CSR. Reactions were emulsified and extracted as described (16) except for alterations to the oil phase and water-to-oil ratio (see below). Briefly, 0.2 ml of CSR mix [primers (1 μM), dNTPs (0.25 mM), 50 μM tetramethylammonium chloride, and 0.05% (vol/vol) DNAse-free pancreatic RNAse (Roche) in 1× *Taq* buffer as well as either induced *Taq* expresser cells (2×10^8) or *Taq* polymerase (10 units) and template DNA were added to 0.4 ml of the oil phase [4.5% (vol/vol) Span 80 (Fluka), 0.4% (vol/vol) Tween 80, and 0.05% (vol/vol) Triton X-100 in light mineral oil (all Sigma)] under constant stirring (1,000 rpm). After addition of the aqueous phase, stirring continued for 5 min more before thermocycling. Compartment dimensions were determined by light microscopy and by laser diffraction spectroscopy as described (16).

Selection. For selecting thermostable mutants, emulsions were incubated at 99°C for increasing amounts of time (up to 15 min) before CSR. For heparin selections, heparin (Sigma) was included at increasing concentrations. CSRs (94°C for 5 min, 20 times at 94°C for 1 min, 55°C for 1 min, and 72°C for 5 min) were carried out as above by using primers 4 (5′-CAG GAA ACA GCT ATG ACA AAA ATC TAG ATA ACG AGG GCAA-3′) and 5 (5′-GTA AAA CGA CGG CCA GTA CCA CCG AAC TGC GGG TGA CGC CAA GCG-3′). The aqueous phase was extracted as described (16) and purified selection products were reamplified by using primers 6 (5′-CAG GAA ACA GCT ATG AC-3′) and 7 (no. 1211; New England Biolabs) and recloned as above. Between rounds, selected clones were rediversified by staggered extension process (StEP) PCR shuffling (17) with activity preselected libraries L1* and L2*.

Screening and Characterization. Clones were screened and ranked by a PCR assay. Briefly, 2 μl of induced cells were added to 30 μl of PCR mix and amplification of a 0.4-kb fragment was compared under standard and selection conditions (e.g., increasing amounts of heparin). Thermostability of purified His_6-tagged wild-type (wt) and mutant *Taq* clones was determined as described in ref. 8 by using activated salmon sperm DNA and normalized enzyme concentrations. K_d for DNA was determined by using BIAcore. Briefly, a 68-mer (18), biotinylated at the 5′ end, was bound to a streptavidin (SA) sensorchip and binding of polymerases was measured in 1× *Taq* buffer (see above) at 20°C. K_d was also estimated by the PCR assay with decreasing amounts of template. Steady-state kinetics were measured as described (19) by using the homopolymeric substrate $poly(dA)_{200}$ (Amersham Pharmacia) and an $oligo(dT)_{40}$ primer at 50°C. Mutation rates were determined by using the mutS ELISA assay (Genecheck, Ft. Collins, CO) as described (20).

Cooperative CSR with Nucleoside Diphosphate Kinase (NDK). The NDK gene from *Myxococcus xanthus* (21) was expressed in QL1387 (22). The inactive mutant *ndk*Δ has an H117A mutation and lacks flanking sequences. Cooperative CSRs (65°C for 10 min, 37°C for 10 min, and 15 times at 94°C for 0.25 min, 55°C for 0.5 min, and 72°C for 1.5 min) were performed as above except at an oil-to-water ratio of 3:1 by using primers 6 and 7 in 1× *Taq* buffer including 100 μM of each dNDP, 400 μM ATP, and 0.1 unit/μl *Taq*.

Results and Discussion

Principles Underlying CSR. Segregation of self-replication into separate, noncommunicating compartments ensures the linkage of phenotype and genotype during CSR. For self-replication of *Taq* polymerase, compartments must remain stable at the high temperatures of PCR thermocycling. Encapsulation of PCRs has been described earlier for lipid vesicles (23) and fixed cells and tissues (24, 25) but with low efficiencies. We used the recently described water-in-oil emulsions (16) but modified the composition of the surfactants as well as the water-to-oil ratio. This modification greatly increased the heat stability and allowed PCR yields in the emulsion to approach those of PCR in solution. Compartments had average diameters of 15 μm and proved heat-stable, with no coalescence or changes in compartment size after 20 cycles of PCR as judged by laser diffraction and light microscopy (Fig. 2*A*). Once formed, compartments did not permit exchange of macromolecules like DNA or protein (Fig. 2*B*). Presumably, their large molecular weight and electric charge prevent their diffusion across the hydrophobic surfactant shell, even at high temperatures.

To colocalize the gene and encoded protein within the same compartment, we used bacteria (*E. coli*) overexpressing *Taq* polymerase as "delivery vehicles." *E. coli* cells (diameter 1–5 μm) fit readily into the emulsion compartments while leaving room for sufficient amounts of PCR reagents like nucleoside triphosphates and primers (Fig. 2*A*). The denaturation step of the first PCR cycle ruptures the bacterial cell and releases the expressed polymerase and its encoding gene into the compartment, allowing self-replication to proceed while simultaneously destroying background bacterial enzymatic activities. In analogy to hot-start strategies (26), this cellular "subcompartmentalization" also prevents the release of polymerase activity at ambient temperatures and the resulting nonspecific amplification products.

Polymerase Selection by CSR. Genotype–phenotype linkage during CSR was probed by mixing cells expressing either wt *Taq* polymerase (wt*Taq*) or the (under the buffer conditions) less active Stoffel fragment (sf*Taq*) (8) at a 1:1 ratio and comparing CSR products in solution or in emulsion (Fig. 3*B*). Although the smaller sf*Taq* was amplified preferentially in solution, in emulsion there was almost exclusive self-replication of the full-length wt*Taq* gene (Fig. 2*C*). Cells were emulsified in such a way that most of the emulsion compartments contained only a single bacterial cell but, because cells are distributed randomly, some would have contained two or more cells. As compartments do

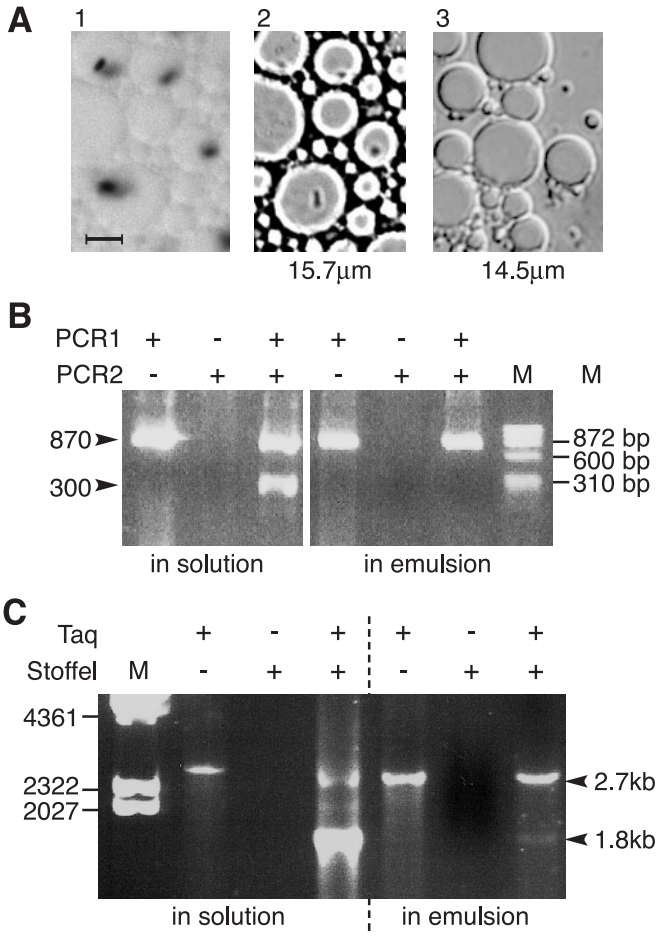

Fig. 2. Compartment dimensions and stability. (*A*) Aqueous compartments of the heat-stable emulsion containing *E. coli* cells expressing green fluorescent protein (GFP) before (*1* and *2*) and after (*3*) thermocycling, as imaged by light microscopy. *1* and *2* represent the same frame, imaged at 535 nm for GFP fluorescence (*1*) and in visible light to visualize bacterial cells within compartments (*2*). Smudging of the fluorescent bacteria (*1*) is caused by Brownian motion during exposure. Thermocycling lyses bacterial cells within compartments. Lysed cells are not visible (*3*), presumably because cell debris disperses throughout the compartment. Average compartment dimensions as determined by laser diffraction before and after thermocycling are given below. The black bar (*1*) measures 10 μm. (*B*) Crossover between emulsion compartments. Two standard PCRs, differing in template size [R1 (0.9 kb) and R2 (0.3 kb)] and presence of *Taq* enzyme (R1, + *Taq* and R2, no enzyme), are amplified (25 cycles) either individually or combined. When combined in solution, mixing occurs and both templates are amplified. When emulsified separately, before mixing, only the template from R1 is amplified. Thus in emulsion, amplification is associated with enzyme activity, indicating that neither template DNA nor protein (i.e., enzyme) is exchanged between compartments during thermocycling. M, *Hae*III-digested phage ϕX174 DNA marker. (*C*) Crossover between emulsion compartments. Bacterial cells expressing wt*Taq* polymerase (2.7 kb) or the less active sf*Taq* (1.8 kb) are mixed 1:1 before emulsification. In solution, mixing occurs and the shorter *sfTaq* gene is amplified preferentially. In emulsion, there is predominant amplification of the *wtTaq* gene (arrow). M, *Hin*dIII-digested λ phage DNA marker.

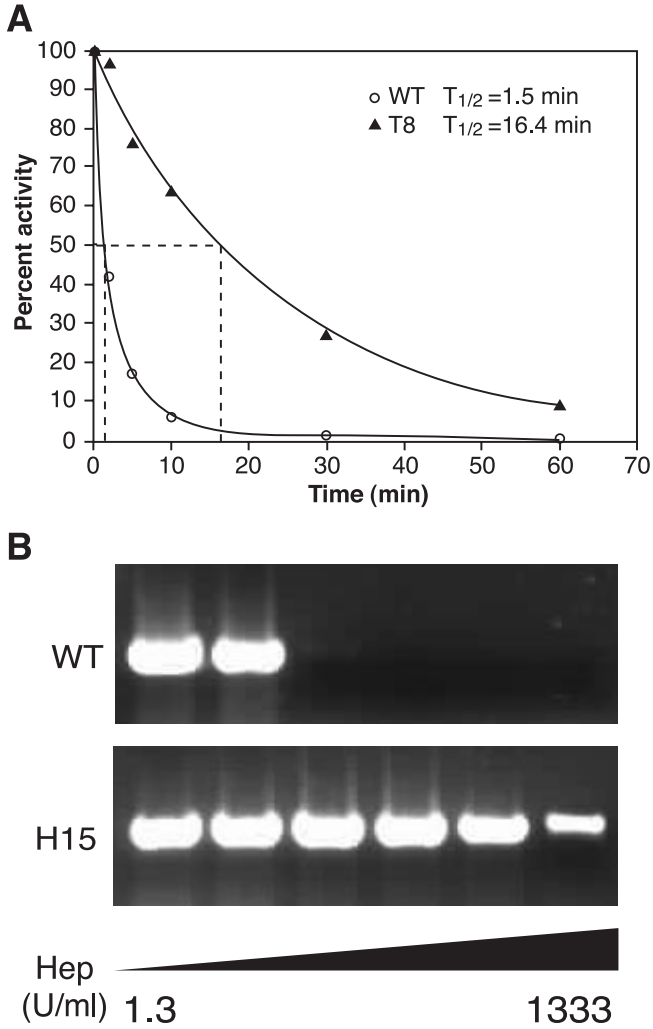

Fig. 3. Properties of selected polymerases. (*A*) Thermal stability of polymerases; percentage of remaining activity is plotted against time. The half-life of enzymatic activity at 97.5°C of wt*Taq* is 1.5 min and of T8 is 16.4 min. (*B*) Polymerase activity in PCR in the presence of increasing amounts of heparin. Heparin concentration increases stepwise by a factor of 4 starting at 1.3 units/ml. The activity of wt*Taq* is inhibited at concentrations >5.2 units/ml, whereas H15 remains active at concentrations of up to 1,350 units/ml.

not exchange template DNA (see above), the small amount of sf*Taq* amplification in emulsion is likely to originate from these compartments. Their abundance is low and does not appear to affect selections. In a trial selection, a single round of CSR was sufficient to isolate wt*Taq* from a 10^6-fold excess of an inactive *Taq* mutant (not shown).

To explore the potential of CSR for the evolution of polymerases with novel properties, we used two different methods of error-prone PCR (see *Materials and Methods*) to generate two repertoires (L1 and L2) of random *Taq* mutants with a diversity of 5×10^7 clones each. *Taq* clones in L1 and L2 had an average of 4–5 mutations and only 1–5% were active in PCR. A single round of CSR selection for polymerase activity under standard conditions increased the proportion of active clones to 81% (L1*) and 77% (L2*) (not shown).

Polymerases Selected by CSR. Increased thermostability is a useful property in a polymerase. It reduces activity loss during thermocycling and allows higher denaturation temperatures for the amplification or sequencing of difficult templates. We first used CSR for the directed evolution of *Taq* variants with increased thermostability, starting from preselected libraries (L1* and L2*) and progressively increasing the duration of the initial thermal denaturation at 99°C. After three rounds of CSR selection, we isolated T8 (F73S, R205K, K219E, M236T, E434D, A608V), a *Taq* clone with an 11-fold longer half-life at 97.5°C compared with the already thermostable wt*Taq* enzyme (Fig. 3*A*). When preincubated at different temperatures for 20 min before PCR, T8 remained active after incubation at 99°C,

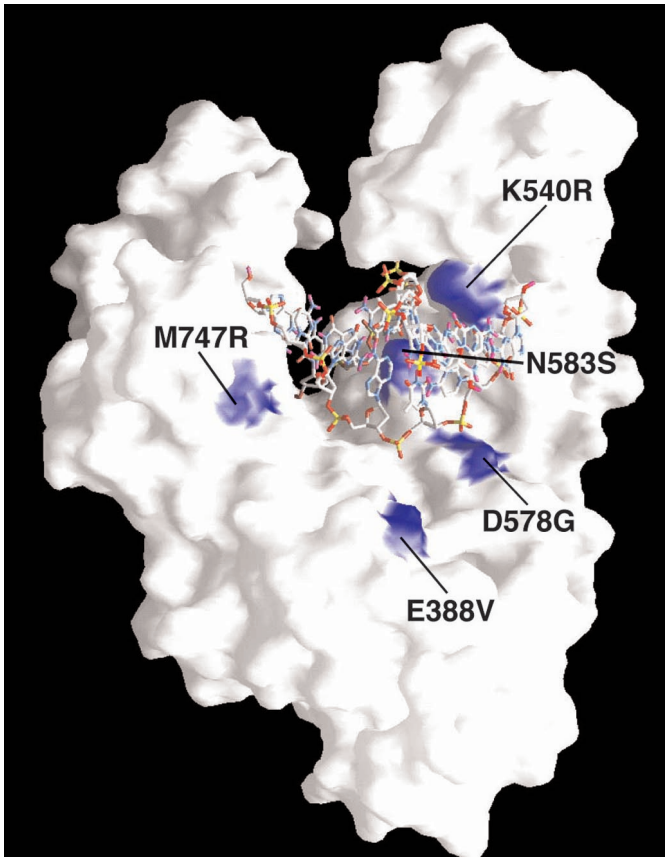

Fig. 4. Structure of heparin resistance. Surface representation of the main polymerase domain in *Taq* polymerase (30). Residues conferring heparin resistance when mutated are shown in blue and labeled. The template–primer hybrid is shown in stick representation, with phosphate atoms shown in yellow, oxygen atoms in red, nitrogen atoms in blue, and carbon atoms in white. Hydrogen atoms are not displayed.

whereas wt*Taq* lost detectable activity at 96°C (not shown). Although less thermostable than some polymerases isolated from hyperthermophilic archaea (27), T8 appears to be the most thermostable member of the DNA polymerase I (pol I) family. It will be interesting to see whether the pol I architecture allows stabilities approaching those of the archaeal pol α-polymerases and whether similar strategies of stabilization [like disulfide bonds (28)] will be used to achieve it.

Mutations conferring thermostability are often difficult to rationalize, but in T8 (and many less thermostable mutants) they cluster in the 5′-3′ exonuclease domain. Truncation variants of *Taq* polymerase (7, 8) lacking the exonuclease domain show improved thermostability, indicating that the exonuclease may be less thermostable than the main polymerase domain. Truncation of the 5′-3′ exonuclease domain also markedly reduces processivity. We found that several clones (although not T8) with mutations in the 5′-3′ exonuclease domain displayed reduced processivity (not shown). Increased processivity might also be evolved through mutations targeting the 5′-3′ exonuclease domain.

Heparin is a widely used anticoagulant but is also a potent inhibitor of all known polymerases, which creates difficulties for PCR from clinical samples (29). Polymerases with resistance to heparin should therefore prove useful in clinical and forensic applications of PCR. By using CSR selection in the presence of increasing amounts of heparin, we isolated H15 (K225E, E388V, K540R, D578G, N583S, M747R), a variant of *Taq* polymerase that remains fully functional in PCR at up to 130 times the inhibitory concentration of heparin (Fig. 3*B*). Heparin resistance-conferring mutations also cluster—in this case in the base of the finger and thumb polymerase subdomains, regions involved in binding the primer–template duplex (30, 31). Four of six residues (K540, D578, N583, M747) mutated in H15 (Fig. 4) directly contact either the template or primer strand in both the open and the closed form of the ternary polymerase–template-primer complex (31). H15 mutations appear to be neutral (or mutually compensating) as far as affinity for duplex DNA is concerned (while presumably reducing affinity for heparin), as judged by BIAcore (Table 1) and a template-dilution assay. The molecular basis of polymerase inhibition by heparin is not known but heparin is widely believed to mimic and compete with duplex DNA for binding to the polymerase. Our results, which suggest overlapping (and mutually exclusive) binding sites for DNA and heparin in the polymerase active site, support this concept.

Genotype–Phenotype Coevolution in CSR. The adaptive phenotype in CSR extends to the polymerase genes themselves, as indicated by a striking bias in the direction of silent mutations [G·C to A·T (64) vs. A·T to G·C (1)] toward a decreased GC content, which promotes replication by facilitating strand separation and destabilizing secondary structures. Coding mutations show no such bias [G·C to A·T (35) vs. A· T to G·C (42)], indicating that this effect is not caused by the fortuitous selection of amino acids encoded by A- and T-rich codons. Polymerase and encoding gene seem to form a unit on which selection acts, resulting in the evolution of genes that code for polymerases better adapted to the selection conditions and are also more efficient templates for replication. CSR thus displays features predicted for a hypercycle, which proposes the coevolution of the components of a self-replicating system to maximize overall fitness (32).

CSR should select not only for optimal adaptation but also for optimal adaptability, i.e., optimal rates of self-mutation. Increased mutation rates can be favorable traits under strong selective pressure. Indeed, mutators can arise spontaneously in asexual bacterial populations under adaptive stress (33) and high mutation rates are common among parasites and viruses. In analogy, CSR might favor polymerase variants that are more error prone and hence capable of faster adaptive evolution. However, none of the selected polymerases displayed significantly increased error rates (Table 1). *Taq* polymerase lacks a 3′-5′ exonuclease proofreading activity and consequently displays an already elevated error rate (34) but its error rate is still at least 10-fold lower than those of the error-prone polymerases of RNA viruses like HIV (35). CSR seems to exert a strong selective pressure on polymerase catalytic activity, as neither improved thermostability nor heparin resistance evolved at the cost of reduced catalytic activity (k_{cat}/K_m) compared with wt*Taq*

Table 1. Properties of selected *Taq* clones

Taq clone	$t_{1/2}$ (97.5°C), min*	Heparin resistance, units/ml	K_d, nM^{-1}	k_{cat}, s^{-1}	K_m for dTTP, μM	Mutation rate†
wt*Taq*	1.5 (7)‡	5.2‡	0.8§	9.0¶	45.0¶	1
T8	16.5 (75)‡	2.6‡	1.2	8.8	48.6	1.2
H15	0.3	1,350‡	0.79	6.8	47.2	0.9

*With N-terminal His_6 tag, as measured by [^{32}P]dCTP incorporation into salmon sperm DNA. Presence of the N-terminal His_6 tag reduces overall thermal stability about 5-fold.
†In relation to wt*Taq* [*Pfu* polymerase (Stratagene), 0.2].
‡No tag, as measured by PCR assay.
§Published value from ref. 27, 1 nM^{-1}.
¶Commercial *Taq* preparation (HT Biotechnology, Cambridge, U.K.): k_{cat}, 4.0 s^{-1}; K_m for dTTP, 43.5 μM.

EVOLUTION

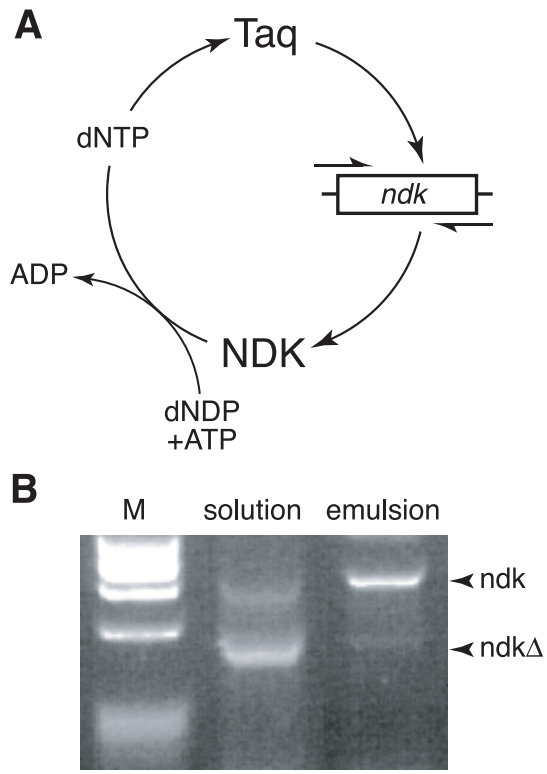

Fig. 5. General scheme of cooperative CSR. (*A*) NDK converts dNDPs to dNTPs, which in turn allow the polymerase to replicate the *ndk* gene. (*B*) In solution, the inactive *ndk*Δ (0.4 kb) is amplified preferentially; in emulsion, the replication is linked to NDK activity and the gene encoding active *ndk* (0.8 kb) is amplified preferentially. M, *Hae*III-digested phage ϕX174 DNA marker.

(Table 1). Increased error rates in *Taq* polymerase may come at the price of reduced catalytic efficiency and therefore may be selected against. Furthermore, selected clones were recombined by PCR shuffling (17) between CSR cycles. Recombination is thought to prevent the selection of mutators (36), because it disrupts the association of the mutator alleles with the favorable traits they generate.

Cooperative CSR. Insertion of an extra stage into the CSR cycle allows the application of CSR to enzymes other than polymerases. For example, two enzymes can cooperate through reciprocal catalysis as shown for NDK (21) and *Taq* polymerase. NDK converts dNDPs to dNTPs, which are the substrates for the polymerase. Hence, only *ndk* genes encoding active NDK are replicated (Fig. 5).

In principle, replication would be expected to cease when the dNTPs within a compartment are exhausted. However, we found that our emulsion mix allowed some diffusion of small molecules, including the negatively charged dNTPs, between compartments during thermocycling (not shown). Permeability of lipid bilayers to charged molecules inversely relates to the length of the aliphatic chain (37), and the aliphatic chain length of the emulsion surfactants would permit diffusion of charged molecules. We therefore reasoned that exchange took place through contacts between compartments rather than through the oil phase, and that a quasibilayer membrane formed at these contact sites. Increasing the separation of the emulsion compartments from each other, breaking intercompartmental contacts, should therefore reduce the diffusion of small molecules. Indeed, by increasing the oil-to-water ratio of the emulsion, dNTP diffusion could be controlled (not shown).

Controlled ''communication'' between compartments offers prospects for the extension of CSR technology. For example, replication reactions may be ''fed'' from the outside or from ''feeder'' compartments and proceed much longer, essentially until there is no room left within the compartments for further replication. It may also be possible to alter buffer conditions after emulsification, e.g., adjusting the pH or starting/stopping reactions with the addition of substrates or inhibitors. Finally, in analogy to a scenario proposed for early living cells (38), it may be possible to produce a ''symbiotic'' CSR mix in which compartments comprising different enzymes cooperate to mutual benefit by a reciprocal exchange of substrates and products to promote the replication of their genes.

Conclusion. We have developed a technology for the study and directed evolution of enzymes, especially polymerases. Our results show that CSR selection can be used to rapidly evolve polymerases with useful properties from a modestly sized repertoire of random polymerase mutants. Future repertoires, in which the diversity is targeted to discrete regions or derives from the molecular breeding (39) of polymerases, may further accelerate the directed evolution of desired polymerase phenotypes. In this study we have focused on an enzyme from a thermophilic organism, but CSR could be applied to mesophilic enzymes, e.g., by *in situ* expression or lysis of the expression host at ambient temperatures.

In its present form, CSR is readily applied to the directed evolution of enzymes (or auxiliary factors) involved in DNA replication or gene expression such as polymerases, ligases, helicases, etc., but we envisage more diverse uses of the technology. For example, CSR may be used for the functional cloning of enzymes from genomic libraries or directly from diverse cellular populations. The PCR-stable emulsion developed for CSR contains an estimated 10^8 to 10^9 PCR-competent compartments per milliliter. Such ''emulsion-PCR'' may allow large-scale single-cell PCR applications with uses extending from gene linkage analysis to repertoire construction.

CSR need not be limited to enzymes involved in nucleic acid transactions. Our results show that two (and presumably more) enzymes can cooperate in CSR through reciprocal catalysis. Further stages may be added to a cooperative CSR cycle, allowing the evolution of both single enzymes and reaction pathways. We anticipate a generic selection system for catalysis built around the ''polymerase engine'' of CSR, whereby coupled catalytic reactions either produce replicase substrates or consume inhibitors, thus allowing replication of the genes encoding the enzymes to proceed.

We thank our colleagues G. Winter, Y. Choo, G. Varani, and A. Klug for comments and suggestions on the manuscript.

1. Steitz, T. A. (1999) *J. Biol. Chem.* **274,** 17395–17398.
2. Joyce, C. M. & Steitz, T. A. (1994) *Annu. Rev. Biochem.* **63,** 777–822.
3. Brautigam, C. A. & Steitz, T. A. (1998) *Curr. Opin. Struct. Biol.* **8,** 54–63.
4. Doublié, S., Sawaya, M. R. & Ellenberger, T. (1999) *Struct. Fold. Des.* **7,** 31–35.
5. Li, Y., Mitaxov, V. & Waksman, G. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9491–9496.
6. Bedford, E., Tabor, S. & Richardson, C. C. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 479–484.
7. Barnes, W. M. (1992) *Gene* **112,** 29–35.
8. Lawyer, F. C., Stoffel, S., Saiki, R. K., Chang, S. Y., Landre, P. A., Abramson, R. D. & Gelfand, D. H. (1993) *PCR Methods Appl.* **2,** 275–287.
9. Jestin, J. L., Kristensen, P. & Winter, G. (1999) *Angew. Chem. Int. Ed. Engl.* **38,** 1124–1127.
10. Suzuki, M., Baskin, D., Hood, L. & Loeb, L. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 9670–9675.
11. Suzuki, M., Yoshida, S., Adman, E. T., Blank, A. & Loeb, L. A. (2000) *J. Biol. Chem.* **275,** 32728–32735.
12. Skerra, A. (1994) *Gene* **151,** 131–135.

13. Engelke, D. R., Krikos, A., Bruck, M. E. & Ginsburg, D. (1990) *Anal. Biochem.* **191,** 396–400.
14. Vartanian, J. P., Henry, M. & Wain-Hobson, S. (1996) *Nucleic Acids Res.* **24,** 2627–2631.
15. Zaccolo, M. & Gherardi, E. (1999) *J. Mol. Biol.* **285,** 775–783.
16. Tawfik, D. S. & Griffiths, A. D. (1998) *Nat. Biotechnol.* **16,** 652–656.
17. Zhao, H., Giver, L., Shao, Z., Affholter, J. A. & Arnold, F. H. (1998) *Nat. Biotechnol.* **16,** 258–261.
18. Astatke, M., Grindley, N. D. & Joyce, C. M. (1995) *J. Biol. Chem.* **270,** 1945–1954.
19. Polesky, A. H., Steitz, T. A., Grindley, N. D. & Joyce, C. M. (1990) *J. Biol. Chem.* **265,** 14579–14591.
20. Debbie, P., Young, K., Pooler, L., Lamp, C., Marietta, P. & Wagner, R. (1997) *Nucleic Acids Res.* **25,** 4825–4829.
21. Muñoz-Dorado, J., Inouye, M. & Inouye, S. (1990) *J. Biol. Chem.* **265,** 2702–2706.
22. Lu, Q., Zhang, X., Almaula, N., Mathews, C. K. & Inouye, M. (1995) *J. Mol. Biol.* **254,** 337–441.
23. Oberholzer, T., Albrizio, M. & Luisi, P. L. (1995) *Chem. Biol.* **2,** 677–682.
24. Embleton, M. J., Gorochov, G., Jones, P. T. & Winter, G. (1992) *Nucleic Acids Res.* **20,** 3831–3837.
25. Haase, A. T., Retzel, E. F. & Staskus, K. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4971–4975.
26. Chou, Q., Russell, M., Birch, D. E., Raymond, J. & Bloch, W. (1992) *Nucleic Acids Res.* **20,** 1717–1723.
27. Perler, F. B., Kumar, S. & Kong, H. (1996) *Adv. Protein Chem.* **48,** 377–430.
28. Hopfner, K. P., Eichinger, A., Engh, R. A., Laue, F., Ankenbauer, W., Huber, R. & Angerer, B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 3600–3605.
29. Satsangi, J., Jewell, D. P., Welsh, K., Bunce, M. & Bell, J. I. (1994) *Lancet* **343,** 1509–1510.
30. Eom, S. H., Wang, J. & Steitz, T. A. (1996) *Nature (London)* **382,** 278–281.
31. Li, Y., Korolev, S. & Waksman, G. (1998) *EMBO J.* **17,** 7514–7525.
32. Eigen, M. & Schuster, P. (1979) *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Berlin).
33. Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. (1997) *Nature (London)* **387,** 703–705.
34. Tindall, K. R. & Kunkel, T. A. (1988) *Biochemistry* **27,** 6008–6013.
35. Domingo, E. & Holland, J. J. (1997) *Annu. Rev. Microbiol.* **51,** 151–178.
36. Tenaillon, O., Le Nagard, H., Godelle, B. & Taddei, F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 10465–10470. (First Published September 5, 2000; 10.1073/pnas.180063397)
37. Chakrabarti, A. C., Breaker, R. R., Joyce, G. F. & Deamer, D. W. (1994) *J. Mol. Evol.* **39,** 555–559.
38. Woese, C. R. & Fox, G. E. (1977) *J. Mol. Evol.* **10,** 1–6.
39. Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. (1998) *Nature (London)* **391,** 288–291.