In vitro evolution of molecular cooperation in CATCH, a cooperatively coupled amplification system

Thomas Ellinger*, Ralf Ehricht* and John S McCaskill

Background: One of the key issues in the investigation of evolution is how complex systems evolved from simple chemical replicators. Theoretical work proposed several models in which complex replicating systems are kinetically stabilized. The development of powerful isothermal amplification technique allows complex nucleic acid based evolving *in vitro* systems to be set up, which may then serve to verify experimentally current theories of evolution. Recently such a system based on the 3SR (self-sustained sequence replication) reaction has been established to investigate the evolution of cooperation: the *trans*-cooperatively coupled CATCH (cooperative amplification by cross hybridization).

Results: Over four rounds of serial transfer, the cooperatively coupled two species CATCH system evolved into a more complex cooperative four species system, which then was overgrown by CATCH-derived RNA-Z-like hairpin species. In contrast to the classical RNA-Z species, these molecules have complementary loop sequences and self-amplify using a dual mechanism that includes concentration-dependent phases of noncooperative and cooperative amplification.

Conclusions: The evolution of a cooperative system, under conditions that were alternately unfavorable and favorable for cooperative amplification, led to a system showing facultative cooperation. This principle of facultative cooperation preserves the complexity of the system investigated and could have general implications for the evolution and stabilization of cooperation under oscillating reaction conditions.

Address: Department of Molecular Information Processing, Institute of Molecular Biotechnology, Beutenbergstrasse 11, 07745 Jena, Germany.

*Present address: Clondiag Chip Technologies, Löbstedter Strasse 105, D 07743 Jena, Germany.

Correspondence: Thomas Ellinger
E-mail: telli@clondiag.com

Key words: cooperation, isothermal amplification, *in vitro*, molecular evolution

Received: **17 August 1998**
Revisions requested: **16 September 1998**
Revisions received: **30 September 1998**
Accepted: **5 October 1998**

Published: **19 November 1998**

Chemistry & Biology December 1998, **5**:729–741
http://biomednet.com/elecref/1074552100500729

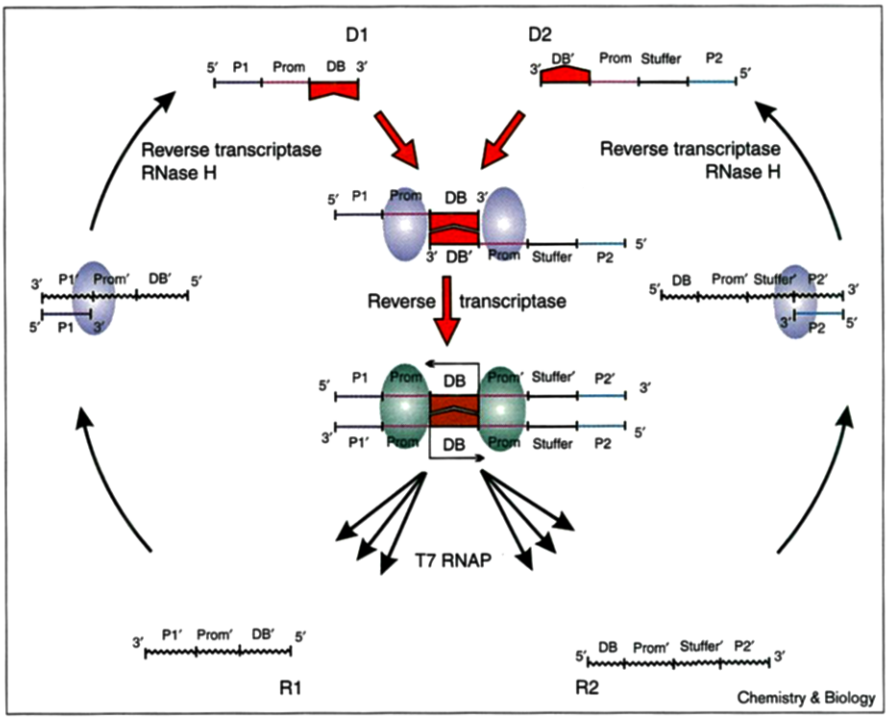
Introduction

The short generation times, large population numbers and limited copying fidelity of *in vitro* DNA amplification systems [1–4] have turned molecular evolution into an experimental science in recent years. DNA-amplification reactions have been used to explore the basic pathways of evolution of independent molecular species on a laboratory time scale [5]. Moreover, with the appropriate choice of primers, coupled amplification systems of interacting DNA templates that exhibit predator–prey and symbiotic behavior have been generated [6,7]. These systems can serve as models for the study of current and early co-evolution and are opening the way to the new field of *in vitro* evolutionary ecologies, to which this paper contributes.

The theoretical implications of the results presented here concern problems ranging from molecular cooperation at the origin of life to the evolutionary stability of modern molecular ecologies. How does molecular information accumulate beyond the simplest replicators to more complex systems and how is it sustained? The investigation of co-evolution of molecular systems is important today for an understanding of host–parasite relationships, the evolution of drug resistance and, ultimately, for the evolutionary successful strategies exploitative cellular proliferation (as in cancer). Furthermore, for biotechnological applications, the range of biomolecular functions that are optimizable through evolution might be generically extendable to *trans*-acting catalysts if one understands experimentally the conditions favoring molecules that produce exploitable resources.

With respect to the early stages of evolution, the maximum length of the first self-replicating molecules (i.e. the information content) would have been restricted by the accuracy of their error-prone archaic replication machinery. In addition to the bootstrapping mechanism (whereby each incremental improvement in replication fidelity allows for a slightly larger genome, which in turn can code for the next incremental improvement of fidelity), two competing mechanistic views have prevailed for dealing with limited fidelity in replicating systems that lack a complex self-sustaining reproductive apparatus (such as a cellular replication and translation machinery). Both mechanisms involve cooperative interaction between different molecular species from a combinatorial family (such as biopolymers) in the process of replication.

Figure 1



Amplification pathway of the cooperatively coupled *in vitro* system. In CATCH the amplification cycles of the DNA species D1 and D2 are coupled *trans*-cooperatively. Coupling occurs via the common double-stranded intermediate, which is formed after mutual priming of single-stranded D1 and D2 in the complementary DB region (red pentagons) and elongation of the hybrid by reverse transcriptase (blue ellipse). Transcription by T7 RNAP (green ellipse) produces the transcripts R1 and R2 specified by the two T7 promoters (purple lines) situated at either side of DB. After primers P1 (blue line) and P2 (cyan line) hybridize to R1 and R2 reverse transcriptase elongates the complex, the single-stranded DNA species D1 and D2 are regenerated. The sequence modules are primers (P1, P2), T7 promoters (Prom), template hybridization region (DB) and a sequence without particular function ('stuffer', black line). Antiparallel sequences of the modules are indicated by a prime after the module name. Straight lines refer to DNA strands, wavy lines refer to RNA.

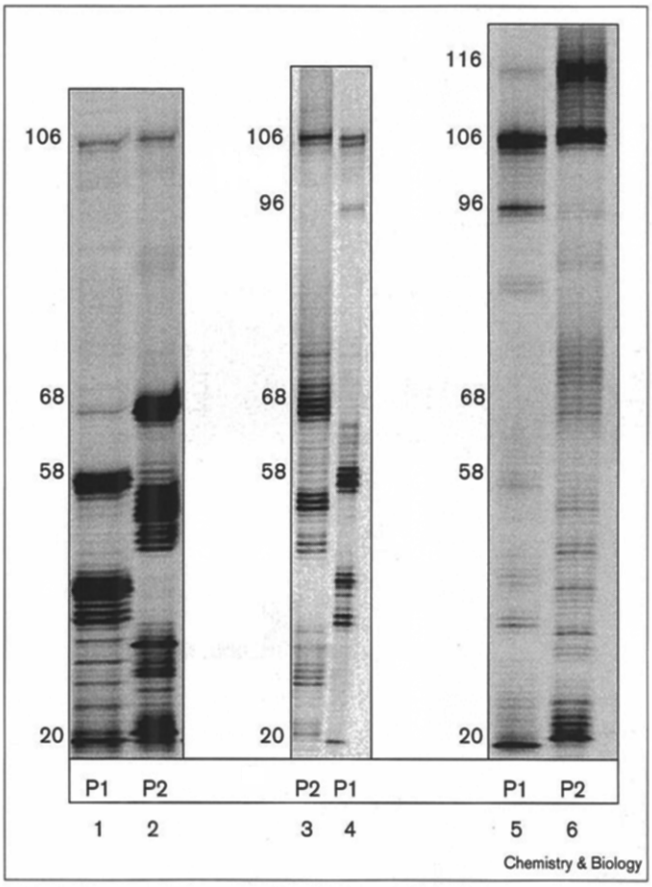
Eigen's hypercycle theory [8,9] argues that the cyclic cooperative kinetic coupling of independent genetic replicators not only provides kinetic stability, but also ensures the continued evolvability of the coupled system. Kauffman [10,11] argues that complex self-replicating cycles involving nonreplicating catalysts are found in large random combinatorial reaction networks of nonreplicating molecules such as proteins. Systems based on nonreplicating molecules are expected to show only limited evolvability [8] and indeed no convincing simulation of continued evolvability in autocatalytic sets has been shown. On the other hand, hypercyclic networks are expected to be unstable towards evolving parasitic species, which enhance their own rate of synthesis without contributing catalytically to the synthesis of other hypercycle components.

A possible solution for this central problem was elaborated in theoretical studies [8,12–14], which predict that cooperation amongst individual molecular species can be stabilized to a certain extent by putting coupled systems in compartments, so that selection can work at higher levels. More recently, different scenarios for the evolutionary stabilization of cooperation in diffusion-limited extended systems without compartments through spiral and self-replicating spot patterns have been suggested [15–17].

Experimental systems are needed to test whether such predictions are relevant to the behavior of real evolving cooperative systems; these systems might also show if evolution will lead to other more simple strategies to stabilize cooperation.

We previously presented the biochemical behavior and the molecular biology of CATCH (cooperative amplification of templates by cross hybridization), a *trans*-cooperatively coupled *in vitro* amplification system [7]. CATCH amplification is based on the self-sustained sequence replication (3SR) amplification reaction, also known as NASBA (nucleic-acid sequence-based amplification) [1,2]. In contrast to the polymerase chain reaction (PCR) [18], the 3SR process is isothermal and supports autonomous amplification without external synchronization, so that different species can amplify at different rates. In CATCH, the basic 3SR scheme is modified to obtain a two species *in vitro* system in which the amplification of the species, the single-stranded DNA templates D1 and D2, is coupled cooperatively through a common double-stranded DNA intermediate (Figure 1). This double-stranded intermediate contains two T7 promoters that are oriented towards each other and separated by a 20-nucleotide DNA sequence (DB). T7 RNAP (T7 RNA polymerase) produces two transcripts, the RNA species R1 and R2. After hybridization of primers P1 and P2 to R1 and R2, and elongation by reverse transcriptase, the single-stranded DNA species D1 and D2 are formed. Coupling occurs by mutual priming of D1 and D2 via the complementary DB and DB′ modules situated at the 3′ end of D1 and D2,

Figure 2



Effect of different reaction conditions on the efficiency of CATCH amplification. The cooperatively coupled system was amplified under previously optimized conditions using AMV RT (lanes 1 and 2), HIV-1 RT (lanes 3 and 4) and under the newly optimized reaction conditions (lanes 5 and 6). The reactions were labeled with 80 nM IRD-41 fluorescent labeled primers P1 (lanes 1,4 and 5) or P2 (lanes 2,3 and 6). One quarter (lanes 1–4) or 1/25 (lanes 5 and 6) of the reactions were loaded on a 10% denaturing polyacrylamide gel and analyzed on a LI-COR sequencing device. The AMV reaction preferentially produced the 58- and 68-nucleotide single-stranded templates D1 and D2, along with shorter abortive DNA products. The amount of the common double-stranded DNA intermediate (106-mer) is only marginal. Use of HIV-1 RT under the same conditions resulted in a higher relative amount of the 106-mer and led to the production of an additional 96 base-pair recombination product. Under optimized conditions, cooperatively coupled amplification was much more efficient and produced almost exclusively the expected 106-mer and two recombination products of 96 and 116 base pairs.

respectively. After priming the common double-stranded intermediate is regenerated by reverse transcriptase.

The CATCH system allows the experimental investigation of evolving molecular cooperation *in vitro*. Here we report our initial findings of the basic pathways of CATCH evolution under batch conditions (i.e. the reaction without continued supply of resources). To this end, serial transfer experiments [5] (i.e. the repeated transfer of small aliquots of an amplification reaction to a fresh solution which then supports further amplification) were performed. Sequences of the progeny of four serial transfer rounds were analyzed and fitted to models of amplification cycles and evolution pathways. The major result of this work was the identification of a new mechanism that stabilizes cooperation under serial transfer conditions.

Results

Optimization of reaction conditions is the key to serial transfer with the CATCH system

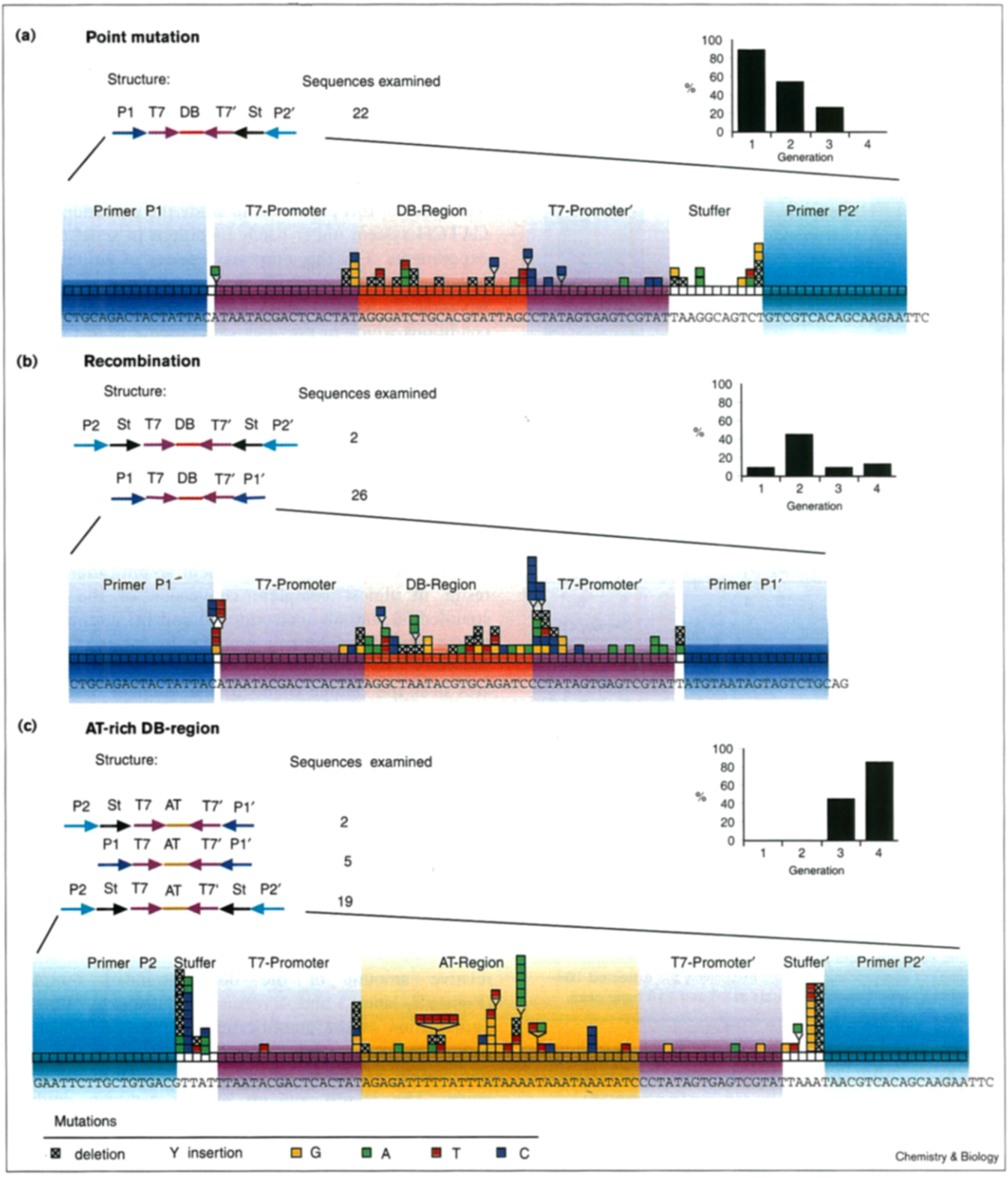
The focus of this paper was to study the evolution of the CATCH system under batch conditions in serial transfer experiments. For this purpose, a series of amplification reactions was set up with the set of enzymes (*E. coli* RNaseH, T7 RNAP and AMV RT) and under the reaction conditions which in previous experiments were found to be optimal [7]. Each single reaction of the series was 'seeded' with 1/50 or 1/10 volume of the previous one. The amplification rate of the CATCH system in these experiments proved to be insufficient (i.e. less than tenfold) to support amplification over several transfers, however.

Analysis of the amplification products on denaturing polyacrylamide gel (Figure 2, lanes 1 and 2) revealed a major bottleneck in CATCH amplification. According to the scheme shown in Figure 1, efficient amplification would result in almost complete conversion of the single-stranded 58- and 68-nucleotide D1 and D2 intermediates to the final 106 base-pair product. In contrast, the gel analysis shows a large excess of single-stranded intermediates over the double-stranded final product, illustrating that CATCH amplification was stalled at the step that converts the D1 and D2 intermediates into the double-stranded product. Inefficient amplification of CATCH was also a result of the extensive production of short abortive products of up to 38 or 48 nucleotides in length (detected in the P1- and P2-labeled lanes; Figure 2).

Because the rate-limiting steps for amplification were inefficient reverse-transcription reactions, alternative reverse transcriptase enzymes were tested. An HIV-1 RT with an amino-terminal His_6 tag [19] produced the largest relative amount of the double-stranded template (Figure 2, lanes 3 and 4 versus lanes 1 and 2). Further optimization of the amplification rate was achieved by varying the concentration of each component of the reaction system, as well as temperature and pH, over a wide range. The largest amount of the common double-stranded 106 base-pair DNA was produced at 42°C in a 40 mM Tris–HCl buffer pH 8.1 containing 30 mM $MgCl_2$, 5 mM KCl, 1 mM dNTPs and 2 mM rNTPs, 2 U/μl His-tagged HIV-1 RT, 0.8 U/μl T7 RNAP and 0.16 U/μl *E. coli* RNase H (Figure 2, lanes 5 and 6).

After optimization, the CATCH cycle ran more efficiently. D1 and D2 intermediates were almost completely

Figure 3



converted into the common 106 base-pair double-stranded product (lanes 5 and 6 versus lanes 1 and 2). In addition to the expected bands, unexpected products of 96 base pairs and 116 base pairs are visible in the P1 and the P2 lanes, respectively. The efficient amplification of CATCH allowed us to perform serial transfer experiments.

Figure 3 legend

Three families of sequences of CATCH templates evolved during serial transfer amplification. Aliquots of each of the four rounds of the serial transfers were cloned. Sequence analysis revealed three structural families of sequences that are related to the input template by **(a)** point mutations (family 1), **(b)** recombination events (family 2) and **(c)** switch of the DB-sequence to an AT-rich palindromic structure (family 3). The left upper corner of each part of the figure shows the overall structure and number of sequences identified belonging to the respective families. Sequence modules are shown as colored lines with primers P1 and P2 shown in blue and cyan, respectively, the stuffer (St) in black, T7 promoters (T7) in purple, the DB region (DB) in red and the AT region (AT) in yellow. The orientation of the modules is indicated by arrows and the antiparallel modules are indicated using primes. The most abundant structure of each family is shown as a consensus sequence in more detail. Sequence modules are labeled with the same color code as described above. Each nucleotide of the sequence is symbolized by a square. All nucleotides that were identified in any of the four serial transfer rounds to deviate from the consensus sequence are indicated as colored squares above (point mutations or deletions) or in between (insertions) the respective positions. Mutations to A are indicated in green, to T in red, to C in blue and to G in yellow. The analysis shows that mutations preferentially hit the nonessential stuffer and the DB region but seldom the central part of the promoters essential for their function. Whereas sequences of families 1 and 2 amplify according to the cooperative mechanism, single-stranded intermediates formed from sequences with AT-rich DB region shown in (c) replicate independently using an RNA-Z-like hairpin mechanism. The graphs in the upper right corner of each part of the figure indicate the percentage of the respective family among the total clones obtained by blunt-end cloning in each generation. They show that the population is overtaken by AT-rich sequences during serial transfer.

Apparent loss of cooperation in the evolution of CATCH to RNA Z

Serial dilution experiments were carried out under the optimized conditions achieved above. After a 2 h reaction, 1/50 of the respective solution was transferred into a fresh reaction mix. After four such transfer generations, the reaction products of each generation were ethanol precipitated and double-stranded CATCH templates were cloned as described in the Materials and methods section.

Sequence analysis of the clones revealed that the resulting molecules are clustered in three families. Figure 3 shows the overall structure of the double-stranded intermediates that belong to each family, as well as the location and type of mutations found in the double-stranded intermediates of the most frequent member of each family.

The first family (Figure 3a) consists of sequences with the original CATCH structure. Although some of these sequences contain point mutations, the general structural organization is the same. As intended in the reaction scheme, the double-stranded intermediates consist of one copy of both single-stranded templates, D1 and D2, each containing a T7 promoter adjacent to the DB region. Mutations were mainly detected within the DB region and in the D2 stuffer region or the weakly conserved distal parts of the promoters that are not essential for CATCH amplification. The structure and the low frequency of mutation in essential parts of the sequence clearly suggest that these templates amplify according to the CATCH mechanism shown in Figure 1.

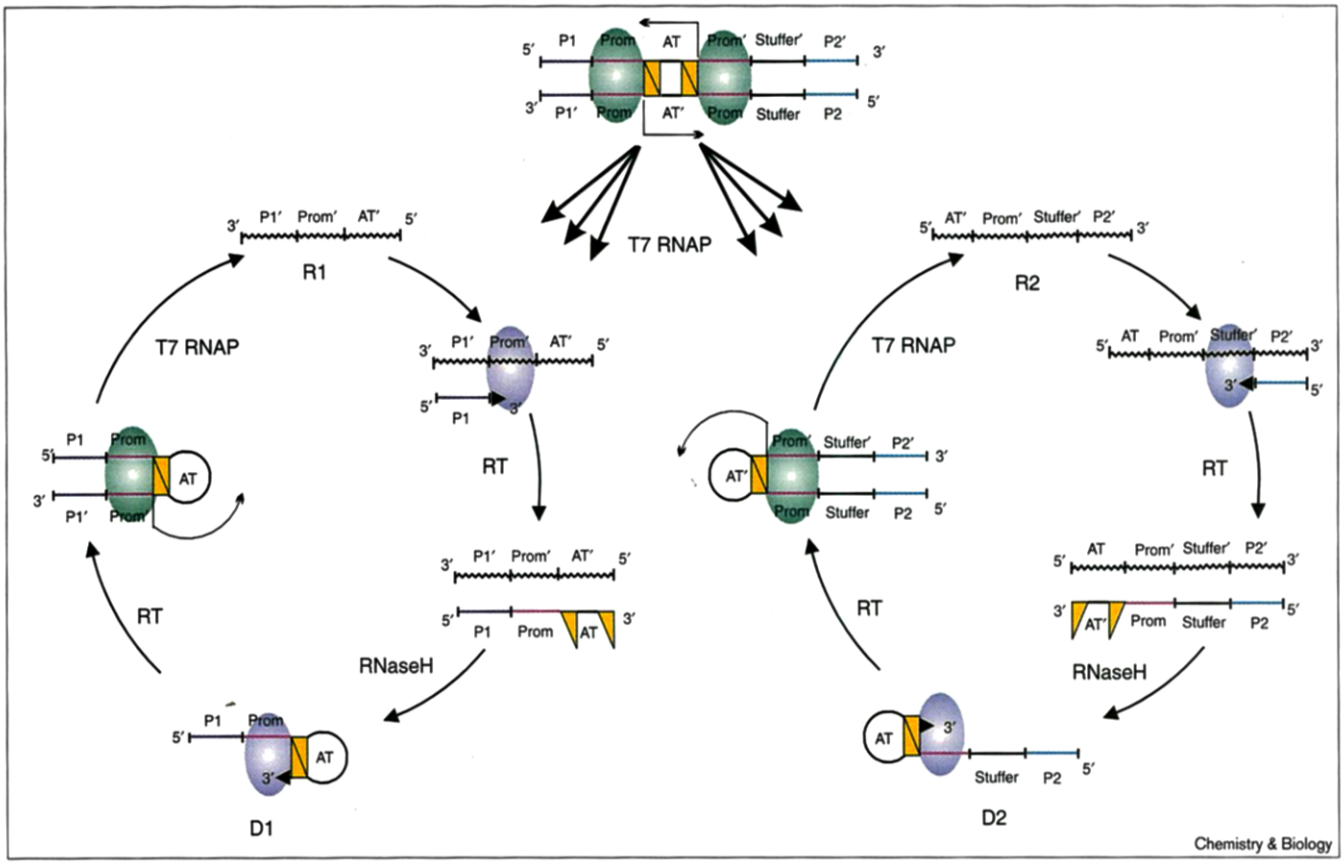
Evolution led to a major reorganization in the second family, which encompasses sequences with a similar structure as the input sequence, except that they are made up of two D1 or two D2 single-stranded template portions that are hybridized via complementary DB and DB′ regions (Figure 3b). The double-stranded intermediates of these sequences had an average length of 96 base pairs (D1 double strand) and 116 base pairs (D2 double strand), respectively. They incorporate primers P1 or P2 only during amplification, which is in good agreement with the data shown in Figure 2 where the 96 base-pair product occurred in the P1-labeled lane 5 and the 116 base-pair intermediate in the P2-labeled lane 6.

The formation of family 2 molecules is not consistent with the original CATCH amplification scheme. Formation of these molecules requires recombination events that combine the D1 sequence with the DB′ module and the D2 sequence with the DB module, respectively.

The overall structure and distribution of mutations in the double-stranded templates of this family suggest, however, that once they have formed, they can amplify according to the cooperative CATCH mechanism. Alternatively, the highly palindromic structure of all sequences found in family 2 suggests that they could be the result of, and amplified by, a noncooperative hairpin replication mechanism [20,21] that would involve hairpin formation after self-hybridization of the single-stranded intermediates D1 and D2 in the DB region, the ends of which are complementary. This possibility, however, was excluded because 3SR reactions seeded with only D1 or D2 did not result in any amplification product (data not shown).

A third family (Figure 3c) was also detected, initially by only a very few cloned sequences. Outside the DB region, they all closely resembled the original CATCH sequence. As in family 2, some templates were products of recombination events: molecules consisting of two D1 or two D2 modules were found, in addition to those with a combination of D1 and D2 modules. In contrast to the rest of the molecule, however, the family 3 DB-region sequence had completely changed and consisted of clusters of oligo A and T. Moreover, the AT clusters within the DB region

Figure 4



AT-rich sequences evolved in the CATCH system can amplify using an RNA Z-like hairpin mechanism. The central part of the sequences, evolved during serial transfer of CATCH, consist of an AT-rich module that contains two complementary regions, indicated by yellow triangles. After transcription by T7 RNAP, hybridization of primers P1 and P2, reverse transcription of the two transcripts by reverse transcriptase and decay of the RNA by RNase H, single-stranded DNA intermediates are formed that contain the self-complementary AT region at their 3′ends. Self-complementary bases can hybridize quickly to form a stem–loop structure, the 3′ end of which is filled in by reverse transcriptase to a complete hairpin. The hairpin contains a double-stranded T7 promoter (Prom, purple) that can be recognized by T7 RNAP, producing RNA transcripts identical to the one produced from the double-stranded template shown above. As a result, the cooperative CATCH amplification mechanism is replaced by two independent RNA-Z-like amplification cycles that are fast and include a concentration-independent self-hybridization step. Abbreviations as in Figure 1.

are highly self-complementary and resemble the stem–loop region of the RNA Z hairpin replicator [20,21]. In fact, the structure of the single-stranded DNA molecules that are produced from these templates suggests that they might form hairpins and could amplify independently according to the RNA Z mechanism (Figure 4). Such hairpins, however, are not expected to be accessible to conventional cloning techniques. The results of a serial transfer experiment support this idea. The reaction was carried out using a single synthetic RNA Z species exhibiting a stem–loop structure identical to the consensus of the DB region of the AT-rich sequences. It wasn't possible to clone the products of each generation despite the appearance of the expected gel bands.

Dual structure of the RNA Z like sequence family

The fact that it was possible to clone sequences of family 3, as well as the resulting structures, indicates that CATCH evolution did not only lead to the formation of RNA Z molecules that amplify according to the hairpin mechanism shown in Figure 4. It must also have led to pairs of hairpins that, in addition to the RNA Z structure, are complementary to one another within the loop sequence. This special characteristic would allow them to form alternative structures — that is either an RNA Z hairpin or, after hybridizing to each other, a double-stranded intermediate similar in overall organization to the input CATCH sequence.

If this assumption holds true, a major proportion of the AT-rich sequences in the reaction should have a hairpin structure and should not be detectable by conventional cloning procedures. Such molecules, however, should be accessible using a modified cloning technique: aliquots of the reaction products of each generation were first incubated with the Klenow enzyme and then cloned as

described above. If sequences with complementary AT-rich DB regions also exist as hairpin structures, they should be able to hybridize via their complementary loop regions to form a heterodimer that might contain mismatches in the rest of the sequence. The Klenow enzyme would repair these mismatches and convert the heterodimer to a double-stranded molecule, which can easily be cloned.

In fact, this cloning procedure led to an increase in the number of clones in the third and fourth (but not in the first) generations. Sequence analysis revealed that most of these clones in the last two generations represented molecules with AT-rich DB sequences. The significantly increased number of clones obtained by the modified cloning procedure suggested that many of the cloned sequences in the reaction mixture had a hairpin rather than a CATCH-like structure.

Further conclusions on the direction of the evolutionary pathway can be drawn from analysis of the distribution of clones among the three families in each generation. Because all but the blunt-end cloning approaches were directed to clone a specific type of template, only clones obtained by blunt-end cloning were taken into account for the calculation. This analysis, the results of which are shown in Figure 3, revealed a continuous shift of the population from the original CATCH templates via recombinant sequences to AT-rich sequences. Whereas sequences with the original structure represent 90% of the templates in the first generation, in the second generation recombinant and original sequences are represented equally. Finally, the population is overtaken by AT-rich molecules which, even if underestimated due to the low cloning efficiency of molecules with hairpin structure, represent 65% and 85% of generations 3 and 4 respectively. Taken together, evolution of the CATCH system resulted in sequences with AT-rich DB sequences that can form both RNA-Z-like hairpin structures and double-stranded templates that exhibit a CATCH-like organization. Further analysis was necessary to determine the functional consequences of the structural features for the evolutionary stability of cooperation (see below).

Result of CATCH evolution is independent of the input sequence

Serial transfer experiments starting from a fixed DB sequence resulted in a shift of the population from cooperatively amplifying molecules to species with an RNA-Z-like structure. This could be the result of a specific inability of the fixed DB sequence to support efficient cooperative amplification. To investigate whether a different result would be obtained starting from different DB sequences, a CATCH system was designed in which the 16 central bases in the DB region were randomized. This sequence was synthesized as single-stranded DNA of 106 nucleotides in length and converted into a double-stranded template using PCR.

Randomization of DB led to a dramatic decrease in the concentration of single-stranded templates with matching ends during amplification, which could prove detrimental to the cooperative amplification scheme. Nevertheless, as revealed by gel analysis, the system amplified efficiently using the CATCH mechanism, generating the typical single-stranded products 58, 68 and 106 nucleotides in length (data not shown).

When this randomized system was subjected to serial transfer over four generations, the system ended up preferentially with the same AT-rich stem–loop structures in the DB region [22]. This illustrates that the result of the serial transfer was determined by the evolution process and not by the input sequence.

The evolved CATCH system retains stable cooperation by using a new dual amplification mechanism

Evolution of the cooperatively coupled system in serial dilution experiments resulted in sequences that can form both an RNA-Z-like hairpin structure carrying one T7 promoter adjacent to the AT-rich stem–loop and a double-stranded template with two convergent T7 promoters situated on each side of the AT-rich sequence. The existence of these alternative structures suggests that the evolved CATCH system amplifies using a dual mechanism. Self-hybridization of the single-stranded DNA intermediates in the AT-rich region would result in a noncooperative RNA Z amplification as shown in Figure 4, whereas *trans*-hybridization of two single-stranded DNA intermediates results in cooperatively coupled amplification. To trace the path of the amplification of the evolved CATCH system, two single-stranded hairpin templates D1-AT and D2-AT, which had exactly the same sequence as the single-stranded intermediates of one of the AT-rich sequences found, were synthesized. The specific structure of these templates makes them suitable for both RNA-Z-like and cooperative CATCH-like amplification.

Because D1-AT and D2-AT differ in size, the products of the different amplification pathways can be distinguished using gel analysis. RNA-Z-like amplification of D1-AT and D2-AT would be expected to produce hairpins of 109 and 117 nucleotides, respectively, whereas the combination of the templates during CATCH-like amplification would be expected to produce an intermediate DNA product of 113 nucleotides in length.

Figure 5 shows the gel analysis of experiments in which different amounts of D1-AT and D2-AT were used, either alone or in combination, as templates in 3SR amplification reactions. Amplification of D1-AT (lanes 1 and 5) and D2-AT (lanes 2 and 6) alone resulted in DNA molecules that

corresponded, in terms of length, to the expected RNA Z-hairpins.

These products were also seen when both templates were present in one reaction, also illustrating that, under these conditions, a significant proportion of the templates amplified using the RNA Z mechanism. In addition, however, under these conditions a DNA product of intermediate size appeared irrespective of whether the reaction was labeled with primer P1 or P2. This product corresponds to the double-stranded template of the cooperative CATCH-like amplification pathway. The distribution of the molecules between the two amplification pathways depended significantly on the concentration of single-stranded templates in the reaction. When the amplification reaction was started with a low template concentration of 20 nM (lanes 3 and 4), templates almost exclusively amplified using the concentration-independent RNA Z mechanism, whereas at input concentrations of 1 μM (lanes 7 and 8) a significant part of the template amplified using the cooperative mechanism. Taken together, evolution of the CATCH system led to templates amplifying according to the dual mechanism shown in Figure 6.

The stabilization during evolution of structural elements for both amplification mechanisms suggests that they are both used and essential for efficient amplification under batch conditions.

Discussion

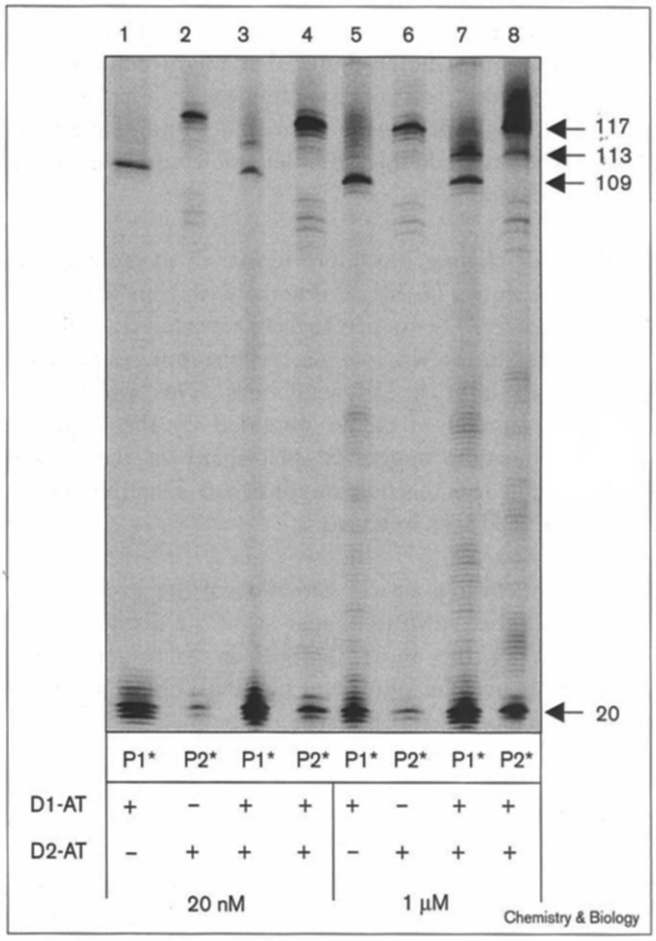
The three evolutionary pathways found in CATCH evolution

In this paper evolution of a cooperatively coupled system was followed over four serial transfer generations corresponding to an average progeny of 10^6–10^7 molecules produced from each single input template. This study has to be regarded as a short-term analysis. Nevertheless, the study is conclusive in that three evolutionary pathways, including the most important evolutionary transition, switching CATCH to a system that seems perfectly adapted to the serial transfer conditions, were observed during this short period.

As expected for any evolving system, CATCH evolution produced molecules with point mutations. These mutations frequently hit nonessential nucleotides in the stuffer or DB regions. In contrast, only a few mutations were found within the promoter sequences, none of which is expected to completely destroy promoter activity (see Figure 3a) [23]. Because CATCH is *trans*-cooperative with respect to the promoter function, the protection of the promoters from mutation was not expected. Inactive promoter mutants could act as parasites that would be efficiently amplified by the cooperating template.

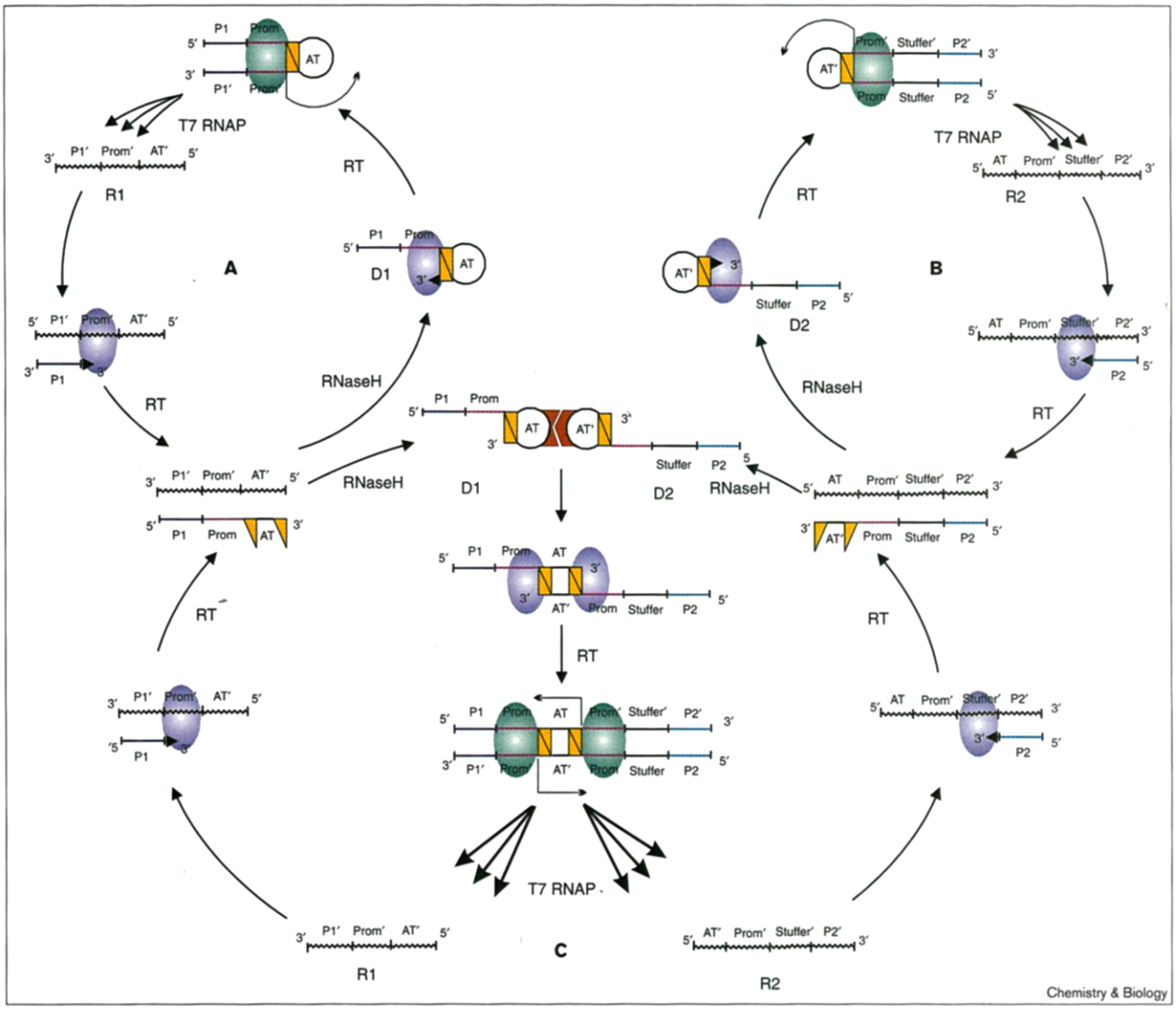
The specific capability of CATCH to exchange whole sequence modules during amplification opened the way for a second evolution pathway, which resulted in sequences representing recombination products consisting of two D1 or two D2 portions (see Figure 3b). In spite of its efficiency,

Figure 5



Concentration-dependent cooperative amplification of RNA-Z-like CATCH evolution products. 3SR amplification reactions were performed using synthetic oligonucleotides of single-stranded hairpin intermediates D1-AT and D2-AT as input templates that were synthesized after the sequence of an AT-rich double-stranded template that had evolved during CATCH serial transfer. The reactions were labeled with 5′ IRD-41 fluorescent primers and after 2 h of incubation subjected to 10% denaturing PAGE on a LI-COR sequencing device as described in the Materials and methods section. Reactions were started either with D1-AT or D2-AT alone (lanes 1 and 5 or 2 and 6, respectively) or with both templates (lanes 3, 4, 7 and 8) and labeled with P1-IRD-41 (lanes 1, 3, 5 and 7) or P2-IRD-41 (lanes 2, 4, 6 and 8). In reactions where only one template was present, the formation of a specific single product of 109 or 117 nucleotides in length for D1-AT or D2-AT, respectively, clearly indicates that amplification occurs according to the RNA-Z-like mechanism shown in Figure 4. As expected for RNA Z amplification, the product pattern was independent of the input template concentration, which was 20 nM in lanes 1 and 2 or 1 μM in lanes 5 and 6. In reactions where both templates were present simultaneously, an additional product of 113 nucleotides was formed, which perfectly matches the length of a product formed by a CATCH mechanism after mutual priming of D1-AT and D2-AT. This is also supported by the detection of the product irrespective of whether the reaction was labeled with P1-IRD or P2-IRD and by its efficient formation at high (1 μM, lanes 7 and 8) but not at low (20 nM, lanes 3 and 4) input template concentrations.

Figure 6

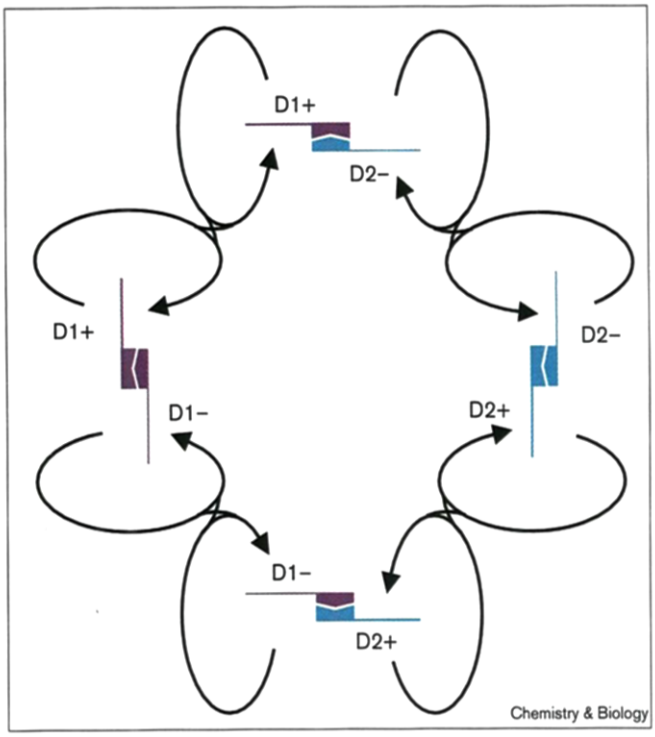


Dual replication mechanisms for species containing an AT-rich DB sequence. As revealed by the experiment shown in Figure 5, evolved CATCH templates can replicate via a dual mechanism. At low concentrations single-stranded templates amplify predominantly independently (Cycles A and B) according to the RNA-Z-like mechanism shown in Figure 3, taking advantage of self-complementary sequences in the AT-rich region at their 3′ end (yellow triangles). The conservation of complementary sequences (red) during evolution, and their presentation as stable single strands by loop formation opens the way to cooperative amplification (cycle C) similar to the CATCH mechanism shown in Figure 1. Abbreviations as in Figure 1.

recombination is not a part of the basic CATCH amplification scheme. It was most probably induced by the hybridization of RNA products R1 and R2 via the complementary DB and DB′ sequences at their 5′ ends. Under nonoptimized reaction conditions (Figure 1, lanes 1 and 2), reverse transcription of these hybrids leads to the formation of abortive D1 and D2 molecules that lack the DB sequence. These can freely hybridize, via their promoter sequence, to any of the RNA templates irrespective of whether this will be R1 or R2. After hybridization, the recessed 3′ DNA ends will be elongated to full length D1 and D2 products, which then carry a DB sequence that was specified by chance.

Recombination not only results in the appearance of new species but also in a cooperative system of increased complexity. A simplified amplification scheme is shown in Figure 7. Each of the four single-stranded DNA templates that exist after recombination (i.e. the D1 or D2 core sequence each fused with DB or DB′, respectively) can

Figure 7



Recombination potentially leads to increased complexity of cooperation during CATCH amplification. Recombination events result in the formation of four different single-stranded intermediates consisting of the D1 (blue) or D2 (cyan) core each fused either to DB (D1+, D2+) or DB′ (D1−, D2−). Each of those templates can cooperate with two of the others during amplification, potentially leading to four cycles of cooperative CATCH-like amplification (elliptical arrows) that are coupled by the exchange of single-stranded intermediates. The mechanism shown is hypothetical in as much as only three of the four double-stranded intermediates were identified in the experiment.

cooperate during amplification with two of the others. The whole amplification cycle is consequently expected to consist of four subsystems that amplify according to the CATCH mechanism and that are connected by exchanging single-stranded intermediates. Sequence analysis, however, revealed double-stranded intermediates for only three of the amplification cycles (the input structure and the D1/D1 and D2/D2 recombination products). The fourth possible double-stranded structure (which is similar to the original CATCH but has an inverted DB sequence) was not found, although it might be formed from single-stranded intermediates of the D1/D1 and D2/D2 templates. Additionally, the distribution of the sequences among the three double-stranded intermediates identified is rather asymmetric with the majority of the sequences representing D1/D1 templates. This asymmetry could partially be attributed to the faster amplification rate of the shorter D1/D1 templates. On the other hand the complex system has only a short time to form an equilibrium between the four amplification cycles as it is rapidly overrun by the third and most interesting class of replicators which dominate the system in the last round of transfer.

Although the pathway of their formation is not clear, the third class of molecules are undoubtedly derivatives of the input templates. In contrast to the latter, however, they exhibit an RNA-Z-like structure that, due to its self-priming capability, forms the basis of one of the most efficient and fastest 3SR replicators [20,21]. It was well documented earlier that such molecules frequently infect 3SR amplification systems and destroy them by competitive amplification. In contrast to this process, the behavior of the RNA-Z-like replicators found in this study is much more interesting. They not only structurally but also functionally conserve the most important characteristics of the input CATCH system which is cooperation.

The dual concentration mechanism for stabilization of cooperation

In this paper, cooperative evolution has been investigated in serial transfer under batch conditions. Each single round of the transfer started at a low concentration of the replicator with an excess of resources and proceeded towards resource limitation. These conditions are complicated in a kinetic sense, as amplification proceeds through several phases with different kinetic coefficients limiting the reaction rate. At the beginning of the batch reaction, template concentrations are very low. Noncooperative hairpin replication, which is based on an intramolecular priming mechanism, will be fast under these conditions. Consequently, the two templates (D1 and D2) in this phase are expected to reproduce independently and exponentially via hairpin replication according to an autocatalytic first-order mechanism (Figure 6, cycles A and B). At the end of the reaction, template concentrations will be high. Mutual priming of the product templates (heterocatalytic phase) will no longer be slow and will give the system an advantage over hairpin replication, including a potentially hyperexponential growth rate (Figure 6, cycle C).

Moreover, facultative cooperation is the key mechanism that stabilizes the evolved CATCH system against parasites that amplify by obtaining *trans*-cooperative support from other species but have lost the *trans*-cooperative functions to support others. Alternating periods of high and low template concentration periodically switches the transcription process between heterocatalytic and autocatalytic phases. Promotorless templates, the typical class of CATCH parasites, will be overgrown in the low concentration autocatalytic phase. Facultative cooperation is a simple alternative to compartmentation in the stabilization of complex systems.

The evolution of the input CATCH system to a system showing facultative cooperation proceeded very quickly. The results presented here do not reveal whether this system, once it has developed, will be in an evolutionary stable situation under the applied conditions. Serial transfer experiments that start with two cooperating RNA Z

templates (D1-AT and D2-AT, see above) showed no general changes either in the template structure or in the dual amplification mechanism after five additional rounds of serial transfer (data not shown) suggesting that the system will continue to be stable.

CATCH as a model system for the evolution of cooperation

The CATCH system, the evolution of which has been investigated in this work, exhibits the first cooperatively coupled *in vitro* system based on nucleic acid amplification. The term cooperative is applied generally in the evolutionary ecology literature to an organismic trait that is either neutral or deleterious to the organism carrying it but that enhances the fitness of other individuals. In contrast with systems like Qb [5] and the basic 3SR reaction [1,2], where complementary copies of the same template are interconverted during amplification (equivalent to the various phases of the life-cycle of an organism), in CATCH the amplification of two templates with different information is cooperatively coupled.

A simplest mechanistic characterization of the form of cooperation in CATCH is:

$$\begin{aligned} &X + Y \rightarrow XY \\ &XY \rightarrow XY + X \\ &XY \rightarrow XY+Y \end{aligned} \qquad (1)$$

with exploiting species X′ and Y′ not supporting the third and second reactions, respectively. Here all interconverting single stranded DNA and RNA species have been lumped together for the two template sequences X and Y (corresponding to the sequences D1 and D2, respectively); XY is the double stranded DNA species. Detailed modeling of the homogeneous and spatially resolved kinetics of the CATCH system will be presented elsewhere. The present discussion focuses on the fundamental evolutionary properties.

The major novel result of the evolution of CATCH after four rounds of serial transfer can be most simply understood as a transition to species X and Y, which in addition to equation 1 also support independent replication:

$$\begin{aligned} &X + Y \rightarrow XY \\ &XY \rightarrow XY + X \qquad X \rightarrow 2X \\ &XY \rightarrow XY+Y \qquad Y \rightarrow 2Y \end{aligned} \qquad (2)$$

Here the species X and Y correspond to templates (D1 and D2) which can amplify either independently via an RNA-Z like mechanism or cooperatively via the CATCH scheme. While the kinetics of this evolved CATCH scheme does not fit the replicator equation format [24], since the coupling kinetics are not proportional to the product of concentrations, it does meet the central criterion and essence of a hypercycle as a cyclic catalytic coupling of self-replicating molecules. This is in contrast with the original and starting CATCH system in which the templates X and Y are not capable of independent replication. The kinetics of both the evolved and original CATCH amplification also differ from the coupled RNA and protein hyperexponential amplification established with an *in vivo* viral system by Eigen *et al.* [25]. The evolutionary stability of the latter system could not be addressed independently of host factors, viral packaging, cell lysis and so on, and was not investigated.

An alternative peptide-based cooperative *in vitro* amplification system has recently been established by Lee *et al.* [26]. The authors have retracted the claim that this peptide system is hypercyclic [27]. To avoid confusion it must be stressed that in Lee *et al.*'s system [26] no added enzymes are required in the environment, which grants their molecules a stronger claim to self-replication. As in Kauffmann's random autocatalytic sets [10,11] however, the cyclically closed catalysis loops are distinct separated cases with limited evolvability. It will be interesting to see to what extent the peptide system can be pushed towards evolution by increasing the probability of catalysis.

The evolution during serial transfer of a system that replicates via a dual amplification mechanism involving facultative cooperation is the most important finding of this paper. Its stabilization during evolution suggests that such a system is adapted to changing template concentrations during serial transfer. A key mechanism to conserve the complexity of the CATCH system was that parasites of the cooperative cycle were overgrown in the low concentration phases of independent amplification. This mechanism could be regarded and potentially be applied as a general principle to stabilize replication systems on the way to complexity.

Significance

The investigation of evolutionary processes on the molecular level by means of simple biochemical *in vitro* systems based on nucleic acid amplification has led to insights into the basic mechanisms of evolution. It has also made the selection of perfectly adapted functional biomolecules possible. To date, *in vitro* evolution experiments have focused on issues of mutation and selection of independent molecular species. The investigation of simple ecosystems, however, contributes both to our understanding of how complex systems might have evolved and stabilized in natural evolution and to our attempts to design technologies that allow the evolutionary optimization of heterocatalysts.

The work presented here contributes to the field of *in vitro* molecular ecology, in which detailed information about the relationship between sequence changes and molecular interaction is obtained. The paper focuses on

the evolution of cooperation as one of the fundamental schemes of molecular interaction found in complex evolving systems.

The CATCH (cooperative amplification of templates by cross hybridization) system, a recently established self-sustained sequence replication (3SR) based on an *in vitro* model system in which the amplification of two nucleic acids is cooperatively coupled, was subjected to four rounds of repeated amplification and dilution during which the reaction conditions for cooperative amplification alternately become favorable and unfavorable. These oscillatory conditions led to the evolution of a facultative cooperatively coupled system consisting of sequences perfectly adapted to both phases of the reaction (i.e. the sequences are capable of both cooperatively coupled and independent amplification).

The mechanism of facultative cooperation preserved the complexity of the system investigated in this paper, and might also be relevant to the evolution and stabilization of complex systems both in natural evolution and for the setup of *in vitro* systems for the continuous evolution of heterocatalysts.

Materials and methods

Enzymes

AMV RT was purchased from USB (Cleveland, Ohio, USA). *E. coli* Ribonuclease H (RNase H), T4 ligase, the Klenow fragment of *E. coli* DNA polymerase I and restriction enzymes were purchased from MBI Fermentas (Lithuania). Vent DNA polymerase was purchased from New England Biolabs (Schwalbach, Germany). HIV-1 RT and T7 RNAP both fused to a 6-Histidine-tag were produced from recombinant *E. coli* strains and purified according to Le Grice *et al.* [19] and Ellinger and Ehricht [28], respectively.

Nucleic acids

Oligonucleotides were purchased from MWG Biotech (Ebersberg, Germany) and purified using high performance liquid chromatography (HPLC) or polyacrylamide gel electrophoresis (PAGE).

The sequences of D1 and D2 (58 and 68 nucleotides in length) were: D1: 5′-CCTCTGCAGACTACTATTAC*ATAATACGACTCACTATA***GGGATCTGCACGTATTAGCC**-3′ (P1, *T7 promoter*, **DB**); D2: 5′-CCTGAATTCTTGCTGTGACGACAGACTGCC*TTAATACGACTCACTATA***GGCTAATACGTGCAGATCCC**-3′ (P2, stuffer, *T7 promoter*, **DB′**).

The sequence of the CATCH 106-mer with randomized DB module was: 5′-CCTCTGCAGACTACTATTAC*ATAATACGACTCACTATA***GGNNNNNNNNNNNNNNCC***TATAGTGAGTCGTATTAA*GGCAGTCTGTCGTCACAGCAAGAATTCAGG-3′ (P1, *T7 promoter*, **random DB**, *T7 promoter′*, stuffer′, P2′).

The sequence of the synthetic RNA Z replicon was: 5′-CCTCTGCAGACTACTATTACAAATATCATCTTTGGTGTTTCCTA*ATAATACGACTCACTATA***GAAATTATTATTTTTTATTTATGAAAAATAAAAATTTC**-3′ (P1, stuffer, *T7 promoter*, **stem–loop structure**).

The sequences of the evolved cooperating hairpin replicators were: T1AT: 5′-CCTCTGCAGACTACTATTAC*ATAATACGACTCACTATA***GAGATTTTTATTTATGAATTAAATAAATATCCC**-3′ (P1, *T7-promoter*, **AT-stem–loop**); T2AT: 5′-CCTGAATTCTTGCTGTGACGTTAT*TTAATACGACTCACTATA***GGGATATTTATTTAATTCATAAATAAAAATCTC**-3′ (P2, stuffer, *T7 promoter*, **AT′-stem–loop**).

Apparatus and reagents

The thermocycler and gel apparatus for agarose gel electrophoresis were purchased from Biometra (Göttingen, Germany). The LI-COR automated sequencing device was purchased from MWG Biotech (Ebersberg, Germany). The gel apparatus for polyacrylamide gels was purchased from Sigma (Deisenhofen, Germany). Nucleoside triphosphates were purchased from Pharmacia (Freiburg, Germany). Mobispin S-300 columns were purchased from Mobitec (Göttingen, Germany). Qiavac plasmid kits and manifold were from Qiagen (Hilden, Germany). All standard chemicals were reagent grade and were purchased from Sigma (Deisenhofen, Germany).

3SR reactions and serial transfer

3SR amplification was carried out in a total volume of 50 µl. Using AMV RT the reaction mix contained 10 nM D1 and D2 or 1 nM of the double-stranded 106 base-pair DNA intermediate, 2 µM P1 and P2, 40 mM Tris/HCl pH 8.0, 10 mM KCl, 5 mM dithiothreitol, 2 mM spermidine, 20 mM $MgCl_2$, 200 µM of each dNTP and 2 mM of each NTP. Enzyme concentrations were 1.0 u/µl AMV RT, 1.0 u/µl T7 RNAP, 0.006 u/µl *E. coli* RNase H. Using HIV-1 RT, the reaction mix contained 10 nM D1 and D2 or 1 nM of the double-stranded 106 base-pair DNA intermediate, 2 µM P1 and P2, 40 mM Tris/HCl pH 8.1, 5 mM KCl, 5 mM dithiothreitol, 2 mM spermidine, 30 mM $MgCl_2$, 1 mM each dNTP and 2 mM each NTP. Enzyme concentrations were 2.0 u/µl HIV-1 RT, 0.8 u/µl T7 RNAP, 0.16 u/µl *E coli* RNase H. The reactions were incubated for 2 h at 42°C. For fluorescence labeling 5′IRD-41 labeled primers P1 or P2 were added to a final concentration of 80 nM. The products of the reactions were desalted on Mobispin S-300 columns before an aliquot (1/4 to 1/50) of the reaction was analyzed by denaturing 10% PAGE on the LI-COR sequencing device [29].

PCR

To create the double-stranded 106 base-pair DNA intermediate with either fixed or randomized DB region, PCR was carried out over 25 cycles with 0.01 u/µl Vent DNA polymerase, 1 nM D1 and D2 (for fixed DB) or 100 nM of the single-stranded 106 base-pair synthetic CATCH template with randomized DB, 400 µM dNTPs, 4 mM Mg^{2+} and 2.5 µM P1 and P2 in a final volume of 100 µl. Temperature conditions were 1 min at 95°C, 1 min at 45°C and 1 min at 72°C. After ethanol precipitation, the double-stranded 106 base-pair DNA product was purified by gel electrophoresis and either used as template for 3SR reactions or subcloned as described below.

Cloning and sequencing

After ethanol precipitation, aliquots of the 3SR reactions were subjected to a single or combined digestion by *Eco*RI and *Pst*I. The double-digested fragments were cloned in the *Eco*RI/*Pst*I treated vector pTEC2 [7]. Single digested fragments were cloned in pUC118 linearized either with *Eco*RI or *Pst*I. Blunt-ended fragments were obtained from the 3SR amplification either directly or after Klenow treatment and were cloned in *Hinc*II linearized pUC118. For sequencing, plasmids were isolated from *E coli* cells grown in TY broth using a Qiavac 8 Kit on a Qiavac manifold (Qiagen, Hilden, Germany) and subjected to a sequencing reaction using the thermo sequenase kit (USB, Cleveland, Ohio, USA). Sequences were analyzed on a LI-COR automated sequencing device (MWG-Biotech, Ebersberg, Germany).

Acknowledgements

We thank Stuart F. L. Le Grice for the generous gift of the *E. coli* clone producing the His-tagged HIV-1 RT, Annette Wagenhaus and Angelika Heller for technical assistance, R. Breitling for help with the multiple alignment and all members of our department for helpful discussions and ideas. This work was supported by the German Ministry of Science and Education (BMBF Grant No: 0310799).

References

1. Guatelli, J.C., Whitfield, K.M., Kwoh, D.Y., Barringer, K.J., Richman, D.D. & Gingeras, T.R. (1990). Isothermal *in vitro* amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl Acad. Sci. USA* **87**, 1874-1878.

2. Compton, J. (1991). Nucleic acid sequence-based amplification. *Nature* **350**, 91-92.
3. 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.
4. Blanco, L., Lazaro, J.M., de Vega, M., Bonnin, A. & Salas, M. (1994).. Terminal protein primed DNA amplification. *Proc. Natl Acad. Sci. USA* **91**, 12198-12202.
5. Spiegelmann, S. (1971). An approach to the experimental analysis of precellular evolution. *Quart. Rev. Biophys.* **4**, 213.
6. Wlotzka, B. & McCaskill, J.S. (1997). A molecular predator and its prey: coupled isothermal amplification of nucleic acids. *Chem. Biol.* **4**, 25-33.
7. Ehricht, R., Ellinger, T. & McCaskill, J.S. (1997). Cooperative amplification of templates by cross-hybridization (CATCH). *Eur. J. Biochem.* **243**, 358-364.
8. Eigen, M. (1971). Self-organisation of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465-523.
9. Eigen, M. & Schuster, P. (1979). The hypercycle - a principle of natural self-organization. Springer, Heidelberg.
10. Kauffman, S.A. (1969).Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **22**, 437.
11. Kauffman, S.A. (1971). Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. *J. Cybernetics* **1**, 71-96.
12. Bresch, C., Niesert, U. & Harnasch, D. (1980). Hypercycles, parasites and packages. *J. Theor. Biol.* **85**, 399-405.
13. Dyson, F.J. (1982). A model for the origin of life. *J. Mol. Evol.* **18**, 344-350.
14. Szathmary, E. (1989). The integration of the earliest genetic information. *Trends Ecol. Evol.* **4**, 200-204.
15. McCaskill, J.S. (1994) Origin of molecular cooperation: theory and experiment. In *Inaugural lectures of the Friedrich Schiller University*, P. Schuser (ed), Friedrich-Schiller-Universität, Jena.
16. Boerlijst, M.C. & Hogeweg, P. (1991) Spiral wave structure in prebiotic evolution: hypercycles stable against parasites. *Physica D* **48**, 17-28.
17. McCaskill, J.S. (1997). Spatially resolved *in vitro* molecular ecology. *Biophys. Chem.* **66**, 145-158.
18. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51**, 263-273.
19. Le Grice, S.F.L., Cameron, G.E. & Benkovic, S.J. (1995). Purification and characterization of human immunodeficiency virus type 1 reverse transcriptase. *Methods Enzymol.* **262**, 130-144.
20. 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.
21. Breaker, R.R., Banerji, A. & Joyce, G.F. (1994). Continuous *in vitro* evolution of bacteriophage RNA polymerase promoters. *Biochemistry* **33**, 11980-11986.
22. Ehricht, R. (1998). Design, verwirklichung und funktional-evolutive. Charakterisierung kooperativ gekoppelter Amplifikation *in vitro*. (Design, setup and functional evolutionary characterisation of cooperatively coupled *in vitro* amplification, *Ph.D. thesis.* Friedrich-Schiller-Universität Jena.
23. Ikeda, R.A., Warshamana, G.S. & Chang, L.L. (1992) *In vivo* and *in vitro* activities of point mutants of the bacteriophage T7 RNA polymerase promoter. *Biochemistry* **31**, 9073-9080.
24. Hofbauer, J. & Sigmund, K. (1988). The theory of evolution and dynamical systems. Cambridge University Press, Cambridge, UK.
25. Eigen, M., Biebricher, C.K., Gebinoga, M. & Gardiner, W.C. (1991). The hypercycle. Coupling of RNA and protein biosynthesis in the infection cycle of an RNA bacteriophage. *Biochemistry* **30**, 11005-11018.
26. Lee, D.H., Severin, K., Yokobayashi, Y. & Ghadiri, M.R. (1997). Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature* **390**, 591-594.
27. Lee, D.H., Severin, K., Yokobayashi, Y. & Ghadiri, M.R. (1998). Emergence of symbiosis in peptide self-replication through a hypercyclic network [Corrigendum]. *Nature* **394**, 101.
28. Ellinger, T. & Ehricht, R. (1998). Single step purification of T7-RNA polymerase with a 6-histidine tag. *BioTechniques* **24**, 718-720.
29. Ehricht, R., Kirner, T., Ellinger, T., Foerster, P., & McCaskill, J.S. (1997). Monitoring the amplification of CATCH, a 3SR based cooperatively coupled isothermal amplification system, by fluorimetric methods. *Nucleic Acids Res.* **25**, 4697-4699.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.