ELSEVIER

Biophysical Chemistry 66 (1997) 145–158

Biophysical Chemistry

Spatially Resolved *in vitro* Molecular Ecology

John S. McCaskill

Institut für Molekulare Biotechnologie, Department of Molecular Information Processing, Beutenbergstr. 11, D-07745 Jena, Germany

Abstract

Sensitive CCD-based fluorescence detection has made spatially resolved studies of evolving cell-free molecular systems possible. In recent years our attention has focussed on making the transition to open and interacting spatially-resolved amplification systems using silicon microreactor technology and on providing a hardware platform for individual based simulation of such systems. Significant progress has been achieved in this direction. Open microflow reactors have been realized in zero (well-mixed), one and two dimensions with volumes small enough to allow long-time studies with limited biochemical materials. The primer directed 3SR reaction (amplifying DNA and RNA) has been used as a basis for constructing interacting model systems with both predator-prey and cooperative amplification character. Theoretical work has demonstrated the need for individual based modeling of such systems: a significant fraction of the population consists of distinct sequence polymers in any case. A massively parallel processor-configurable computer NGEN has been designed and constructed which allows the high speed simulation in hardware of relatively large populations of locally interacting individual strings of chosen length (e.g. up to 2000*2000 for 64 bases), in addition to its application as an evolvable hardware machine. Simulations show self-replicating spots to stabilize the cooperative amplification in evolving systems (a mechanism proposed by the author in 1994). Both oscillatory kinetics and pattern formation are expected in the experimental model systems under investigation which profoundly affect the course of evolution. Such *in vitro* model systems serve both to test current theories of cooperative evolution and provide clues for optimisation strategies in molecular biotechnology.

Keywords: *In vitro* amplification; molecular ecology; reaction-diffusion; evolutionary biotechnology; evolvable hardware; parallel computer; fluorescence detection; DNA; RNA; FPGA.

1. Introduction

Darwin's observations of natural populations leading to his proposal of the origin of species involved primarily the correlation between geographical proximity or accessibility and species similarity. Indeed spatial structuring of evolving species plays a major role and deserves careful experimental analysis in model systems. In this article, I seek to cover the attempts to extend to spatially resolved systems the precision and completeness of *in vitro* evolution studies, where evolution experiments can be performed on

0301-4622/97/$17.00
PII S0301-4622(97) 00073-2

laboratory timescales for large populations, and where populations can be sequenced and their kinetics analysed. There is a second motivation for this program: the role of spatial isolation in the emergence of cooperation. While the cell provides a clearly defined spatial unit, with the benefits of intracellular reactions being restricted to the cell concerned, the coordination of different reactions to the benefit of the participating genetic determinants is problematic in the absence of this rigidly controlled dividing spatial structure. Where biochemistry dictates the use of separate molecules, the problem of exploitation is acute. Spatially resolved studies of evolving populations which interact can help to shed light on this phenomenon.

In the absence of controlled cell division, spatial diffusion provides a weaker form of spatial correlation between reactants and products. The square root scaling of transport distance with time effectively defines a confined volume, on any chosen timescale, within which material exchange occurs. Differences in diffusion coefficients and nonlinearities in the reaction kinetics are the usual prerequisites for pattern formation. Traveling concentration waves form a basic dynamical structure incorporated into concentric rings, spirals and other structures. The kinetic properties of molecules selected in such patterns can then depend on the patterns involved and thereby modify the course of natural selection.

In vitro experiments in spatially resolved evolution began in 1987 and came to first fruition in the development of a one dimensional closed reaction media in disposable plastic capillaries (total length 30m) [1,2]. We review this work briefly in section three before turning to the construction of open microreactors which allow both an indefinite prolongation of amplification and open the path to complex pattern formation. Sensitive fluorescence detection is necessary to monitor the concentration of species online. A linkup with *in vivo* work was achieved [3], involving spatially resolved evolution of bacterial viral infections (plaques) on agar plates. The idea of a spatial phenotype via front velocity [1] has also born fruit in that area [4]. Complex spatial pattern formation in resource limited bacterial colonies on agar plates has also been found [5].

In evolving populations, the discrete nature of individuals plays a key role [6], with mutant changes amplifying initially from single copies. With the comparatively high error rates of *in vitro* systems, the populations consist typically to a large extent of unique individuals [7] and this sequence variety determines the rate of evolution. A modeling facility at the individual level is required to handle these systems stochastically, in contrast with the conventional partial differential equation approach to reaction-diffusion systems (see section four). Modeling at the individual molecule level has the additional advantage of being able to handle variations in amplification mechanisms within the same framework and offers the conceptual transparency of the programmable matter approach. The author initiated in 1992 a program to develop configurable hardware (massively parallel) with which the local properties of reaction and diffusion of individuals could be handled in a sequence dependent manner. The first stage of this development is now complete and the computer NGEN developed in our lab now presents a major simulation facility for this research field [8,9].

Two basic *in vitro* ecological model systems have been developed in our lab and are presented in sections five and six: a molecular predator-prey system and a cooperative amplification system. They are based on one of the isothermal amplification systems reviewed in section two. The paper concludes with an outlook for the field of *in vitro* molecular ecology. While this review concentrates on the perspective developed in our group, interested readers should find pointers to other relevant literature in the references.

2. Isothermal amplification systems as *in vitro* models

2.1 Serial transfer and Qß replicase

The discovery of the *in vitro* amplification of genetic material of the *E.Coli* virus Qß opened the way for *in vitro* evolution studies [10]. Prolonged amplification under constant conditions were achieved using the serial transfer technique, which involves the repeated transfer of a small sample to a fresh solution which then supports further

amplification. Using this discrete approximation to a chemostat flow reactor, natural selection could be shown to operate: shortening the genetic material to remove the coding of now superfluous proteins and enhance the amplification rate and introducing mutations to confer resistance to chemicals such as ethidium bromide [11]. The purification and physical characterisation of the cellular extract showed a single four subunit enzyme (mwt 215,000), with one subunit encoded by the native virus, to be responsible for the amplification of the single stranded RNA. Apart from enzyme binding sequence motives, the kinetics of refolding of the complementary template and replicate during the copy process, avoiding double strand formation, is also considered crucial to the sequence dependence of the amplification [12]. The overall kinetics of the amplification process requires a three component mechanism in order to account for the experimental concentration dependence on template (I) and enzyme (E) [13]. In the simplest case of reverse complementary palindromic sequences (identical plus and minus strands) the mechanism may be written:

$$I+E \rightarrow IE \rightarrow EI+I \qquad EI \rightarrow E+I$$

with template released last. When the RNA template concentration reaches that of the enzyme, the strong enzyme binding gives rise to a sharp titration transition from exponential to linear amplification which slows at still higher concentrations as a result of RNA double strand formation and product inhibition. The outcome of selection in the different growth phases is radically different: with sequences of different length, binding affinity and amplification rates emerging [14].

This system provided the experimental inspiration for a physico-chemical theory of natural selection formulated by Eigen [15]. Without a modern replicase enzyme, the amplification of information is error prone with accuracy limited by the chemical discrimination of base-pair interactions. The quasispecies theory established that optimization operates with such error prone replication processes, up to a maximum sequence length. The *in vitro* Qß system has been found to confirm this theory at error rates about a factor of a hundred lower than in the enzyme free case, so the Qß enzyme can be viewed as an environmental factor and the RNA as self-replicating in this special environment. The discovery of simple template chemistry [16] and catalytic RNA [17] has provided some substance to an early RNA-like world [18] without protein enzymes, and although splicing and cleavage processes probably played a key role there, the quasispecies model and the Qß system retain a central place in our understanding of physico-chemical optimization.

2.2 PCR and synchronisation

The discovery of thermostable polymerases in organisms living in hot springs enabled a controlled amplification via the process of single to double strand polymerisation of DNA and temperature cycling (the polymerase chain reaction, PCR) [19]. Since the polymerase requires a short piece of double strand to initiate polymerisation, the amplification may be made sequence specific by providing primer sequences complementary to the chosen target in the reaction mix. This property of primer direction has been found useful for the construction of the specific model systems described below. In PCR, the temperature cycling synchronizes the amplification of all the molecules involved so the only optimization of sequences during amplification involves variable rates of template loss. Although this reaction has been of great importance technologically, its non-isothermal nature and synchronization makes it unsuited for reaction-diffusion studies. The thermal expansion of water and convection lead to mixing.

2.3 Strand displacement amplification

The idea to replace the temperature dependent double strand separation by an enzymatic process involving a restriction enzyme (Hinc II) lead to the isothermal strand displacement amplification reaction (SDA) [20]. The restriction enzyme Hinc II nicks a double stranded DNA at a specific location, with one side protected by the use of a modified thiolated nucleotide and the other stemming from a DNA primer. This provides a starting point for displacement polymerisation with the Klenow fragment of DNA Pol 1 (Exo -). *In vitro* evolution using serial transfer was studied in Eigen's

laboratory [21]. The engineering of a thermostable DNA polymerase enzyme will improve the specificity of this system. No theoretical work has as yet addressed the specific potential of this system as a model system for evolution.

2.4 3SR (NASBA) isothermal amplification

One of the by-products of retroviral research, accompanying the research effort to deal with AIDS, has been the discovery of an isothermal amplification process involving reverse transcriptase. This 3SR cycle (for self sustained sequence replication) was discovered by groups in the USA [22] and Holland [23], where it is known as the NASBA reaction. Here the problem of double strand replication is solved via RNA transcription intermediates (the copying of double stranded DNA in the cell is a complicated process, in particular for the reverse strand). A promoter sequence in the double stranded DNA initiates the transcription (by RNA polymerase from the phage T7) of a single stranded RNA product in multiple copies. A DNA primer sequence anneals to these RNA strands and allows a reverse transcriptase enzyme (e.g. from HIV) to retranscribe the RNA to single stranded DNA (via an RNA-DNA hybrid). The digestion of the RNA strand can be further enhanced using the endonuclease RNAse H. A second DNA primer containing the promoter sequence anneals to this DNA template with a dangling end and reverse transcriptase completes the double strand in both directions, completing the amplification cycle.

From the point of view of *in vitro* evolution model systems it has several advantages. Firstly, the two different primers allow considerable flexibility in the choice of sequences being amplified. Although promoter requirements are quite strict and the remainder of the primer sequence must be chosen to minimize undesired secondary structure formation with reaction intermediates, there remains a considerable freedom of choice. Secondly, the occurence of both DNA and RNA intermediates allows the possibility of incorporating catalytic RNA in the amplification system and using DNA sequencing to analyse the population. Thirdly, the process of conversion with several steps involving primers opens the reaction to coupling variants involving more primers and templates as we shall see below. The disadvantage of the reaction compared with PCR is in the specificity of the amplification: generally PCR products exhibit the narrow product spectrum which results from high temperature annealing. A thermostable version of the 3SR reaction, involving the genetic engineering of reverse transcriptase, would be of advantage here. The restricted specificity of the 3SR reaction leaves it open also to radical short-circuit mutants. The T7 RNA polymerase enzyme has also been shown to support self-replicating RNA hairpins in the absence of primer [24,25], and minimonster mutants of the 3SR reaction have been characterized [26] which replace one or more of the primer annealing steps by intramolecular folding. These possibilities must be taken into account in the design of coupled systems.

3. Traveling waves and spatially resolved microreactors

3.1 Traveling wave experiments with Qβ and the 3SR reaction

The first experiments on spatially resolved *in vitro* amplification were performed with the QB replicase model system[1]. Translucent polyethylene capillaries with inner diameters between 0.4 and 1mm were used in these experiments. The initial apparatus involved shining a focussed laser beam along the axis of a 15 cm long capillary, mounted with a moveable cross-hair for measuring wavefront position based on manual observation of lateral fluorescence intensity. This rapidly gave place to a fully automated and parallel traveling wave reactor consisting of 30m of capillary strung on a metal frame in 15cm segments inside a thermostatted chamber with quartz windows and observed by a peltier-cooled CCD camera through interference filters.

Modelled as partial differential equations, the amplification of RNA with the Qß replicase may be viewed as a multicomponent generalisation of the Fisher equation [27] for the concentration u

$$\frac{\partial u}{\partial t} = D\nabla^2 u + f(u) \text{ with } f(u) = ku(u_0 - u),$$

with D the diffusion coefficient, k the amplification rate and u_0 the limiting concentration of template.

For the multicomponent mechanism reviewed in section two, a minimal front velocity may be calculated from the roots of a polynomial of degree six. As in the Fisher equation, convergence from localized initial concentrations to a traveling wave with this velocity occurs. The ability to reduce the front velocity problem to a solution of polynomial equations for the class of "pull" fronts, whose relative reaction rates f'(u) are monotonic decreasing, places this class of inhomogeneous amplification equations in a similar situation to the matrix solution of linear systems of kinetic equations for the homogeneous case. The situation when the relative amplification rate is highest at intermediate concentrations is relevant in section six. In the case of Qß, where enzyme binding is strong, a good approximation to the front velocity is the effective Fisher form $v = 2\sqrt{\kappa D}$ where D is the diffusion coefficient of the RNA-enzyme complex and κ the overall amplification rate computed as eigenvalue of the low concentration homogenous kinetic scheme.

Fronts are relatively sharp, decay lengths ca 100μm, and velocities in the range of 100μm/min with the replication times of the order of a minute. This length scale, shorter in more viscous media, has prompted the transition to microstructured reactors. Already in the capillary reactor, thousands of wavefronts could be followed simultaneously and space-time images of their evolution constructed and measured automatically to determine probability distributions of front velocities and their changes [2]. Mutation and selection take place in the traveling wavefront and result in sharp increases in front velocity. The traveling wavefront is thus a homeostatic evolving population as in an ideal flow reactor with a geometricaly measureable phenotype. *In vitro* evolution under conditions of limiting low concentration of UTP, needed in RNA elongation, have confirmed the separate evolutionary optimization of the elongation rate parameter as a systematic reproducible process in such wavefronts [28].

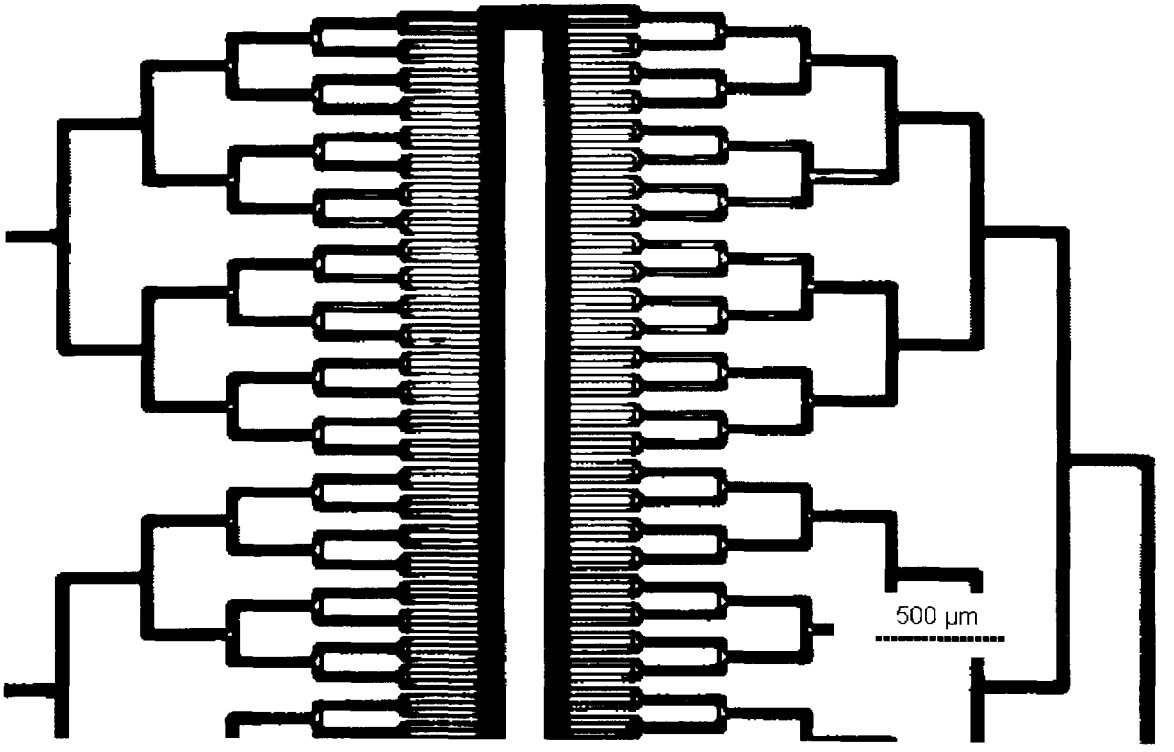


Fig. 1 Open linear flow reactor. A small vertical segment of a 7cm long reactor is shown. The long and narrow reaction chamber is etched aniostropically to a depth of 100μm and a shallow lateral tree inlet and outlet structure etched isotropically (depth 10μm). Concentration waves and other structures can be detected in the centre reaction chambers using a CCD camera and fluorescently labeled nucleic acids. The cross flow (from left to right) is too rapid in the inlet structure to allow back contamination and low in the chamber to allow diffusional mixing.

3.2 Open microflow reactor

The transition to microreactor technology first occurred with the etching of capillary meanders in silicon. Both isotropic and compensated anisotropic etching processes were employed. The etched silicon wafers were bonded anodically to pyrex to close the reaction vessels. After an initial period of filling reactors from the perimeter of the silicon wafers, an ultrasonic drilling technique was established with glued contacts to external capillaries. This resulted in an easily temperature controlled and optically cross-talk minimized microreactor version of the capillary apparatus. Microchannels with depths from 10 to 100μm and of varying widths could be closely packed on the silicon wafer. In a second step, a reaction chamber designed to hold back enzyme loaded beads was designed as an open flow reactor and a capillary meander used downstream to provide a spatial image of the time course of the reaction in the microstructured chamber [29]. 17 such reactors were packed in the first stage of integration on a single wafer; several hundred being achievable with the mm scale.

3.3 Linear open microreactor

Achieving an open spatially resolved reactor with cross flow, requires the development of homogeneous flow profiles. Arne Bochmann in our lab devised a symmetric hierarchical tree structure which guarantees an identical flow rate at all outlet branches. Corresponding etched silicon flow channels were produced by Dr. Schmidt and tested with fluorescent beads. Using two juxtaposed tree inlet structures a flow reaction chamber with laminar interface [30] (a microscopic implementation of the macroflow reactor [31]) in which lateral traveling waves could be initiated using stopped flow from a defined starting concentration profile was developed and tested. More recently, use has been made of rapid diffusional mixing on the length scale up to 100μm to construct a one dimensional open microreactor [32] as shown in Fig. 1. A factor of 100 difference in the flow velocities in the inlet and reactor prevents back contamination of the inflow while ensuring diffusional mixing in the chamber over a wide range of flow rates. With this reactor, continuous spatially resolved evolution can be explored, including the self-replicating structures described in section six.

3.4 Planar open microreactor

A multilayer silicon-pyrex sandwich structure is under development to implement an open microreactor in two dimensions [33]. The requirement of homogeneous sheet flow perpendicular to a planar reactor is a strict one and has required some construction tricks and compromises. Pressure barriers were made by etching vertical bars through thin silicon wafers to even the velocity flow and hydrodynamic calculation performed to optimize the reactor. This reactor will enable the interplay between two-dimensional pattern formation and evolution to be explored. Systems exhibiting such patterns are described below.

3.5 Fluorescence detection

The double strand intercalator ethidium bromide (EBr) was used in most of the traveling wave experiments, excited at 366nm with a transilluminator and observed at around 600nm in the red. Subsequently, thiazole orange has been employed extensively, excited by an argon ion laser at 488nm and observed at 530-580nm. The latter dye is now well known in its commercial variant (To-Pro) from Molecular Probes. It shows a binding constant of 3μM to MNV-11 (one of the self-replicating RNA variants of length 86 nucleotides) and its fluorescence enhancement on intercalation of up to a factor of 3000 is certainly superior to EBr (enhancement factor 14) enabling a dynamic range of over three orders of magnitude in RNA detection. Its use in other *in vitro* reactions such as the 3SR

reaction is limited by the background fluorescence of the DNA primers. Single molecule detection experiments in our laboratory have succeeded in detecting DNA of only 106 base pairs in free solution using To-Pro-1 [34] and this limit has recently been reduced to 61. A spatially resolved single molecule detector has been conceived [35] and is currently being tested. This research has pushed our detection to allow the measurement of DNA/RNA concentrations in the sub micromolar range in the thin films only 20-100μm deep in Si microreactors.

4. Simulating evolving reaction-diffusion systems

4.1 NGEN as programmable matter

A powerful simulation tool for such systems has developed as a by-product of research towards evolvable hardware via the programmable matter metaphor. Programmable matter can be achieved using field programmable gate arrays (FPGAs) to implement distributed processors at the level of digital logic. The observation that molecular evolution and diffusion could be formulated simply in terms of an array of simple two state, two input, two output switches each processing strings sequentially was the starting point for this development. The processes of diffusional interchange, copying, selection, mutation, recombination are captured in this abstraction. Sequence dependence of the states of the switch already allows a significant fraction of population genetics modeling to be covered. Following the successful implementation of a test card in Göttingen, a massively parallel 18 card machine, NGEN[8], was constructed in Jena allowing the rapid parallel simulation of populations of millions of strings of length of the order of 100 bases using distributed high speed memory chips. The results reported below were obtained with this variable architecture 162 FPGA machine.

A population is modelled as a two dimensional array of binary strings in diffusive motion, undergoing probabilistic reactions between neighbours. This is realized at high speed in hardware in a dataflow concept using a line of processing sites (x-coordinate) through which the population streams cyclically from and to high speed distributed memory in several thousand parallel threads each containing a thousand successive bit-serial individuals. Lateral diffusion (in x) is implemented by multiplexers between neighbouring strands whereas vertical diffusion (in y) employs multiplexers between successive strands in a thread, these being retained in the processors *via* small configurable RAMs [36]. Random numbers are generated in a linear array of simple interconnected elements at machine frequency[37].

4.2 Evolvable Hardware

The field of evolvable hardware seeks to construct hardware systems capable of evolving in complex real-world situations . Traditional FPGAs have not allowed rapid partial reconfiguration because of both the complexity of the design-binary mapping and the shift register basis of serial configuration. A new generation of FPGAs (such as the Xilinx 6200 series [38]) treats the configuration of the FPGA as a random access memory (SRAM) and allows fine-grained partial reconfiguration at MHz frequencies. This opens the door to incremental evolution of the hardware [39]. A second generation configurable computer POLYP is currently being developed to exploit these possibilities [40]. Embedded in a feedback loop with a real world environment, such as a two dimensional optical detector processing single photon events, and equipped with an interface to actuators, such systems become capable of autonomous evolution and intelligent control of real systems. One such target system is the biochemical microreactor discussed in the previous section. A hybrid evolution of detection, digital control, and biochemicals becomes possible. This link up of evolvable hardware with molecular systems is a major driving force behind the development of configurable computers in our group, opening as it does the door to an evolvable bioelectronics.

4.3 Interacting and Non-interacting Population Models

Trimolecular reactions are rare in aqueous solution and a rather general biochemical simulation facility can be developed on the basis of pairwise

reactions. In order to incorporate this in a data flow concept, reactions in which molecules are created are treated uniformly as reactions of the type A+B→C+D with reactants being overwritten by products. In this way the copy process becomes A+ϕ→2A where ϕ is either an empty string or another reactant which is overwritten by the copy process. In this way copy processes and selection are linked via population saturation as in the Moran model of population genetics. This is not a fixed feature of NGEN but has been found to be a useful modeling platform. Reaction rates are modelled as reaction probabilities and these are calculated in a sequence dependent manner.

In general two sequences A and B determine the outcome as to which is copied (or what recombination of both is produced in more complex models). Monadic phenotypic fitness of sequences, independent of the constitution of the rest of the population, emerge in the special case when the magnitude of reaction probabilities for copying of A or B induces a total ordering on the sequence space: a special case of more general interaction schemes. Simple models of fitness landscapes such as the N-K model [41] can be implemented very compactly in digital logic. More generally, the existence of reaction probabilities such that A>B>C>A from binary encounters enables interacting models with frequency dependent selection effects to be studied. For example, the predator-prey situation obeys R>B>ϕ>R where R is the predator, B is the prey and ϕ is an empty space. Processes such as diffusion and sequence dependent processes can be parameterised with probabilistic rates using efficient parallel random numbers distributed in digital logic. The initial criticism of digital logic modeling as being devoid of reaction rate information has not been substantiated. In the following sections we describe two interacting population models in more detail which have been implemented both experimentally *in vitro* and in hardware simulation.

5. Molecular predation

Since the discovery of oscillatory dynamics in predator prey systems, the Lotka-Volterra model has served as a central model for population ecology. Lotka's formulation was in chemical terms [42], so the absence until now of chemically evolvable examples is somewhat surprising. An autocatalytically amplifying species, the prey, acts as a chemical resource for the amplification of a second species, the predator. Spatially, predator-prey systems show patch formation and predators can lead to regional segregation of two competing prey species [43]. Saturation effects play a key role in the dynamical behaviour of the model, with predator saturation behaving analogously to the Michaelis-Menten enzyme kinetics [44]. Whether a system with the predator prey form of opposed quadratic interactions and saturation effects displays stable limit cycles or not depends on the details of the mechanism and parameters. Traveling wave velocity expressions have been derived for the predator-prey case by Dunbar [45].

5.1 Molecular predator-prey model in vitro

The 3SR reaction described in section two admits the synthesis of different independently replicating systems by the appropriate choice of the two primer sequences [46]. Coupling between different such cycles can be achieved using a replication intermediate of one cycle as a primer in the second. The second amplification cycle then acts as a predator of the first. Functionally, the template sequences involve the region encoded by primer 1, a promoter sequence, an intervening sequence and the complementary sequence to primer 2 which consists of a primer motif followed by a promoter sequence in the reverse orientation. The DNA single stranded template produced as a result of reverse transcription in the prey cycle can either be captured by the prey primer or by the predator template. In the latter case it acts as primer 2 of the predator cycle, providing the necessary promoter sequence without which no predator amplification is possible. This system has been constructed and parameters optimized and confirmed to operate according to the proposed mechanism. Details may be found in the original publication [46].

A kinetic mechanism for this reaction has been proposed involving the various primers, templates and enzyme complexes with reverse transcriptase and RNA polymerase. An analysis of the homogeneous kinetics shows that the system does

indeed oscillate for reasonable values of the parameters when placed in a flow reactor [47]. Experiments are in progress to examine this behaviour in the microflow reactors described in section three. The critical flow rate has been determined and evolution experiments are also underway to investigate the sequence stability of the system. As a second step, analysis of spatial pattern formation in the one and two dimensional microflow reactors is planned.

5.2 Predator-prey fitness and modeling

One basic feature of the sequence dependence of predator prey interactions concerns the molecular recognition between predator and prey. Since this involves competing DNA-DNA hybridisation of prey template to primer or predator in the predation reaction step, the kinetics of association plays a decisive role. Unfortunately, for a simple theory, the prediction of sequence dependence for this process where both partners already match is complicated by intramolecular folding of the single stranded DNA molecules. The presentation of a hairpin sequence unique to the complementary partner is expected to be optimal for nucleation of the hybrid. In the case of the discrimination of mismatched partners, the sequence dependence is theoretically more accessible. With a double strand hybridisation length of 20 nucleotides, single mutations generally destabilise the complex significantly and double mismatches are likely to be prohibitive for recognition. This motivates the following simple sequence dependent model to address the key behaviour of this evolutionary model system. Amplification rates of the prey cycle have a direct sequence constraint (modeling the restrictions of primers and promoter functions) and the unconstrained portions of the prey sequence are open to hybridisation by the predator. The predator cycle is dependent on the ability to match the prey at this subsequence. This situation is reminiscent of the problems of viral-host coevolution which give rise to the red queen phenomenon [48]. Depending on the way in which predator saturation effects are treated, various behaviours are possible. Work on NGEN to highlight the special features of the evolution of molecular predator-prey interactions is in progress and will be reported elsewhere.

6. Molecular Cooperation

Catalytic function generically involves the use of one molecule to enhance the rate of formation of a distinct reaction product, i.e. heterocatalysis. In the Qß system described in section two, the polymerase is an essential heterocatalyst and the RNA template (if self reverse-complementary) an autocatalyst. The optimization of heterocatalysts requires a coupling of the benefits of the product to the producing catalyst. In practical examples of RNA catalytic optimization *in vitro* [49], this coupling has been ensured using artificial separation technologies such as SELEX [50]. Using ribozymes which were preoptimized to a high rate of ligation by PCR amplification and off-line selection, Wright and Joyce [51] have recently succeeded in demonstrating continuous evolution of a ligase ribozyme via serial transfer alone in a 3SR system. Amplification requires the ligation of a primer containing a promoter for T7 and outcompetes minimonster parasites only at higher ligase rates. In this section, we explore the biotechnologically relevant question of whether an intrinsic spatial coupling can suffice.

6.1 Information threshold and hypercycles

The origin of molecular cooperation was first seen as a problem in the apparent giant step required to coordinate the biosynthesis of proteins. The modern translation apparatus employing the triplet code involves a complex system of self-assembling protein and RNA components including the various aminoacyl transferases, the ribosomal proteins and RNA, transfer and messenger RNAs and transcription factors. In Eigen's seminal work [15], both catalytic networks of proteins and nucleic acid directed synthesis were analysed based on the fundamental physico-chemical limits of molecular recognition of these molecules. Nucleic acid base-pairing provides a limited accuracy of template replication, while the self-reproduction of proteins in catalytic networks depends on a rare coincidence of catalytic closure and is not evolvable. Hypercycles were proposed as a solution to this dilemma [15,52]: providing generic self-replication based on nucleic

acids with increased total molecular information resulting from cyclic catalytic coupling.

Catalytic cycles are generically susceptible to parasitic couplings or side branches which do not contribute to the cyclic closure. Such parasitic couplings are also a problem with the catalytic coupling of inherent replicators such as nucleic acids, directly or indirectly via proteins. Furthermore, parasitic couplings are generically optimizable to dominate over the cyclic catalysis because the constraint of product catalytic activity is removed [53,54]. The solution to this dilemma was seen by all authors as lying in compartmentalization, Eigen arguing that compartments without internal hypercyclic coupling as proposed in the stochastic corrector model [55] were subject to the same low information threshold which the hypercycle model was designed to deal with.

6.2 Hypercycles and Spirals

Diffusion provides an alternative to compartments as a means to ensure that catalytic products contribute to the amplification of catalysts. A cellular automata model was first employed to show that hypercycles with more than five catalytic species with diffusion demonstrate spatial structuring in spirals and can persist in the presence of parasitic species which would kill a homogeneous system [56]. The novel feature is that the spiral structure, in contrast with a dividing compartment, self-organizes as a result of the underlying kinetics specifying the hypercycle. This work has been extended to survey a range of kinetic mechanisms including the phenomenon of spiral replication in concentration gradients [57].

The high number of species required for spiral formation and parasite stability leave open the issue of whether spirals are practical in stabilizing cooperation. In addition, a three dimensional analysis suggests that the spiral solution to contain parasites does not work [58]. In what follows we have identified an alternative spatial structure which is generated naturally by *in vitro* models with the simplest form of cooperative coupling and appears to provide a solution of the stable coupling problem even for two species.

6.3 Self-replicating Spots

The discovery of self-replicating spot structures [59] in the Scott-Gray cubic catalysis model of resource-limited spatial amplification

$$\frac{\partial a}{\partial t} = D_a \nabla^2 a + k_c\, a^2 b - (k+F)a$$

$$\frac{\partial b}{\partial t} = D_b \nabla^2 b - k_c\, a^2 b + F(b_0 - b)$$

where b is a faster diffusing resource buffered to concentration b_0 and a is an autocatalytic species in a flow reactor at rate F, lead the author [60] and others [61] to propose that such structures are of importance to the stabilization of cooperation. The idea is that the self-replicating spots of catalysts in which a parasitic species exploiting the coupling to amplify, but not contributing catalytic function, goes extinct whereas parasite-free spots proliferate. The mutation rate to parasites must not exceed the order of magnitude of spot replication when spot replication is by duplication [62], although this requirement may be relaxed with increasing spot fragmentation. Self-replicating spot patterns have to date only been found experimentally in an inorganic system [63]. The prediction of front velocities in such nonlinear amplification systems without renewable resources has been reviewed [64].

The author has recently described a second order activation mechanism which like the third order Scott-Gray model also generates self-replicating spot patterns and exhibits parasite suppression:

$A+A \rightarrow A'+A$; $A'+B \rightarrow A+A$; $\phi \rightarrow B$; $A,A',B \rightarrow \phi$

where A' is the activated form of A and B is the resource. This scheme has been simulated stochastically using NGEN at the individual molecule level in a two dimensional space. A phase diagram including the complex patterns of Pearson and some new stochastic ones can be constructed with populations in the millions. Kinetic parameters can be linked to the sequences of the molecules involved and the open evolution studied. Such analysis is in progress and the indications are that self-organization to the appropriate region of pattern space is possible. The model is closely related to a reduced model of an *in vitro* coupled system CATCH [65] showing such cooperative behaviour, for which we show an example of a simulation result below.

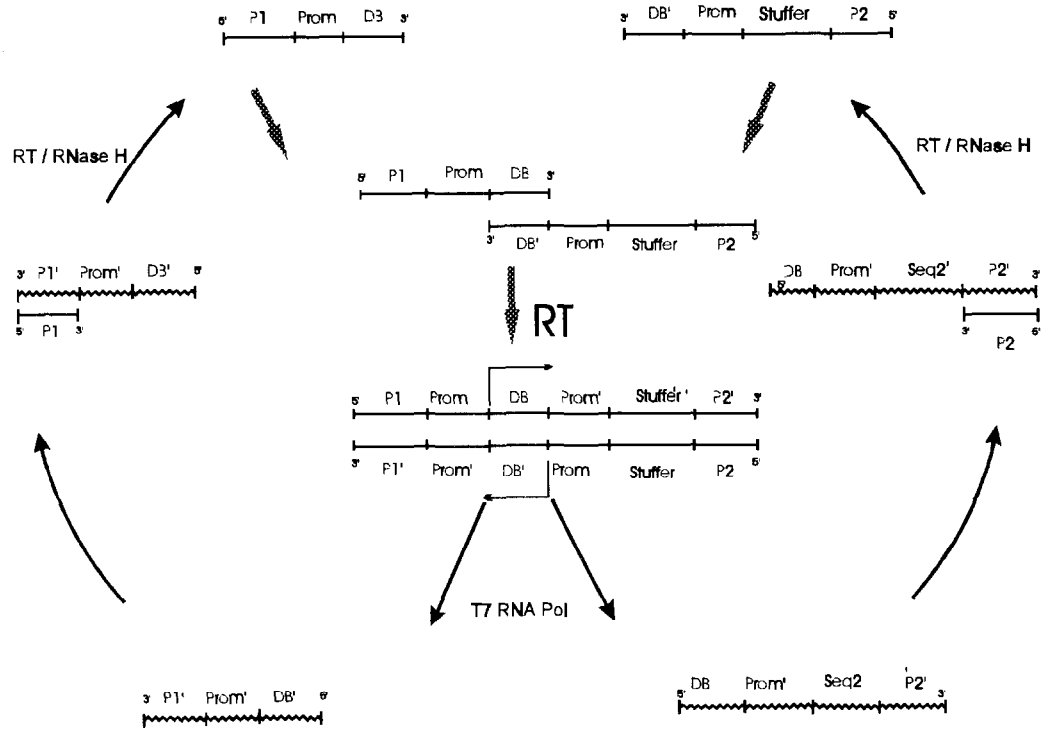


Fig. 2 CATCH mechanism of cooperative amplification of two templates (see text and [65]). Two DNA templates (at top) hybridize at the DB subsequence and are completed to a double stranded structure (centre) from which RNA transcipts starting after the promotor subsequence (prom) are formed. With the aid of two primers, P1 and P2, these RNA transcripts are retranscribed to DNA. The bimolecular template hybridization step results in hyperexponential amplification at low concentrations. Since each template contains the promotor for transcription of the other, the sequence information is cooperative.

6.4 CATCH

The search for a realizable cooperatively coupled *in vitro* system for quantitative evolution studies has given rise to the 3SR based system CATCH (cooperative amplification of templates by cross-hybridization)[65]. The coupling scheme of this reaction is shown in Fig. 2. Two single stranded DNA templates hybridize with dangling ends by a second order reaction step. Reverse transcriptase (RT) completes the double strand to the combined length, and two promoter sequences trigger the transcription of multiple copies of the RNA templates which are converted by RT using added DNA primers to the starting DNA single strands.

This system is a model system for cooperation because each template encodes the promoter sequence which initiate the transcription leading to the other. It is open to exploitation by sequence variants which carry a defective promoter and have used this additional degree of freedom to increase their own synthesis rate. The bimolecular step means that the reaction kinetics are no longer linear in template concentration at low concentration: amplification should be greater than exponential at these concentrations.

A simplified kinetic model distinguishes only the double stranded species A' and the two templates A (written here as identical, to make the connection with the bimolecular spot model above clear) and a resource species B (e.g. rNTP used in the transcription):

$$A+A\rightarrow A';\quad A'+B\rightarrow A'+A;\quad \phi\rightarrow B;\quad A,A',B\rightarrow\phi.$$

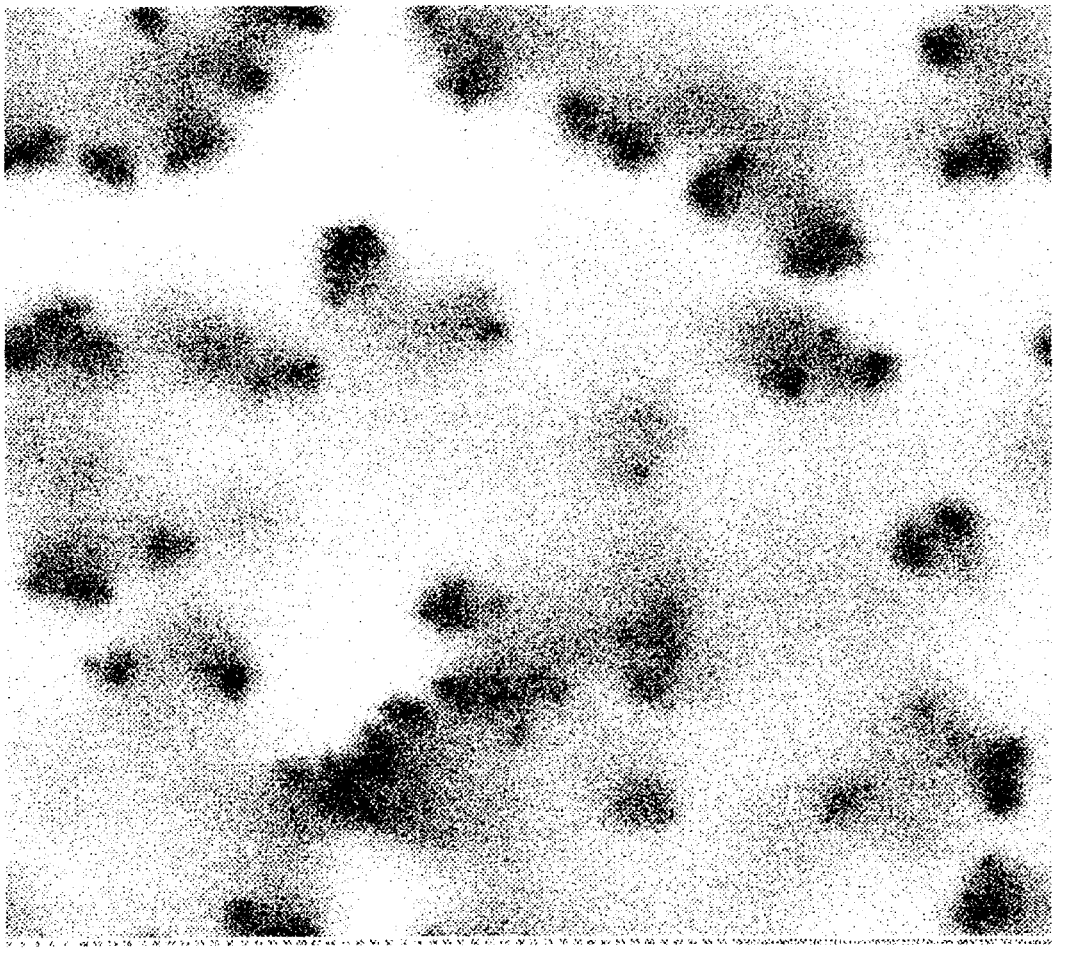
Indcpendently of whether separate templates A_1 and A_2 are used, this system shows self-replicating spots (in a somewhat different portion of phase space from the previous model). An example of a large

individual based simulation of this system along the lines discussed in section four is shown in Fig. 3.

6.5 Evolvable hybridisation model

A mechanistically more realistic and correspondingly more intensive computationally model of this coupled amplification system has been developed [66]. Sequence dependent hybridization of templates and primers is treated in arbitrary relative registration, providing an effective sequence dependent modeling of the process of multistrand complex formation. Promoter dependent transcription on a base by base level allows detailed mutation modeling. The model formulation at the individual level has meant that no additional overhead on NGEN is involved in allowing evolution in addition to the reaction-diffusion.

Fig. 3 Individual based simulation of CATCH system in two dimensions according to the simplified mechanism of section 6.4. In this black and white rendering, all departures from homogeneous resource a shaded dark. Self-replicating spots persist at long times in this biochemically realisable model of *in vitro* amplification. A 2400x2100 array of strings of length 64 bits was employed in the simulation, which was run for 320.000 iterations in several minutes on NGEN. The bit serial processing allows evolution of individuals to be studied with no extra computation time cost. Already this system has been shown to suppress parasitic species in the appropriate parameter regime. This work will be reported in detail elsewhere.



Simulations of this type have demonstrated that spatial structuring together with certain sequence space effects radically alters the susceptibility of the CATCH system to parasites. Extensions of the original catch system to include additional templates can result in spiral structures (with equal diffusion coefficients) if enough species are involved. With a more rapidly diffusing resource species such as NTP, the system demonstrates self-replicating spots and the key question concerns the open sequence coevolution of catch templates and exploiters to determine what behaviour emerges in an evolving population where novelty is introduced by mutation and the derived new possibilities of hybridization. This work is currently being pursued and will be reported elsewhere.

7. Perspectives

The fields of reaction-diffusion systems and of molecular evolution are well established and I have had no space to review the exciting separate developments in recent years. Rather, I have endeavoured to show how spatially resolved studies on coupled *in vitro* amplification systems provide an experimentally accessible and theoretically challenging testing ground both for our understanding of the evolution of cooperative systems and for engineering ideas concerning catalytic optimization.

In coming years, the engineering goals will shift increasingly to active microreactors with intelligent on-line data processing and flow-control. We are working on an optical-magnetic interface to microprocessor control. Hybrid evolving systems consisting of evolvable electronics and parallel microcontrollable biochemistry will emerge. The control of microenvironments will allow combinatorially complex selection conditions for coupled systems.

Spatial experiments in open microreactors have begun on *in vitro* amplification systems which have been kinetically characterized. Once the teething problems of gas bubble formation and surface binding have been resolved, which adversely affect the velocity homogeneity and detection respectively, we expect spatial pattern formation to be prevalent under appropriate resource and flow conditions. Template coupled systems can be augmented by catalytically coupled ones, first on a ribozyme and then protein basis. The interplay between molecular evolution and pattern formation provides a new dimension for studies of self-organization and quantitative studies will foster our understanding of physico-chemical models of the central biological phenomenom of molecular cooperation.

Acknowledgment:

Part of this research was funded by a project grant of the German Government (BMBF 0310799).

The author wishes to express his congratulations to Manfred Eigen on the occasion of his 70th birthday. His scientific vision has brought physical chemistry and evolutionary biology onto a common level of inquiry, providing the necessary step for the research presented here. Personally, I would also like to honour Manfred Eigen's uncompromising stand for a caring and democratic scientific leadership which has been a beacon on a sometimes obscure landscape.

References

[1] G. Bauer, H. Otten and J.S. McCaskill. Proc. Natl. Acad. Sci. USA 86 (1989) 7937.
[2] J.S. McCaskill and G. Bauer. Proc. Natl. Acad. Sci. USA 90 (1993) 4191.
[3] J. Yin and J.S. McCaskill. Biophys. J. 61 (1991) 1540.
[4] Y. Lee and J. Yin. Nature Biotechnology 14 (1996) 491.
[5] E. Ben-Jacob, O. Schochet, A. Tenenbaum, I. Cohen, A. Czirók and T. Vicsek. Nature 368 (1994) 46.
[6] W.G. Wilson, A.M. de Roos and E. McCauley. Theor. Pop. Biol. 43 (1993) 91.
[7] M. Eigen, J.S. McCaskill and P. Schuster. Adv. Chem. Phys. 75 (1989) 149.
[8] J.S. McCaskill, T. Maeke, U. Gemm, L. Schulte and U.Tangen. Lect. Not. Comp. Sci. (1997) in press.
[9] U. Tangen, J.S. McCaskill and J. Ackermann. Proc. European Conference on Artifical Life (1997) (MIT Press) in press.
[10] S. Spiegelmann. Q. Rev. Biophys. 4 (1971) 213.
[11] D.R. Mills, R.L. Peterson and S. Spiegelman. Proc. Natl. Acad. Sci. USA 58 (1967) 217.
[12] C.K. Biebricher. Chemica Scripta 26B (1986) 51.
[13] C.K. Biebricher, M. Eigen and W.C. Gardiner Jr. Biochemistry 22 (1983) 2544.
[14] C.K. Biebricher, M. Eigen and W.C. Gardiner Jr. Biochemistry 23 (1984) 3186.
[15] M. Eigen. Naturwissenschaften 58 (1971) 465.
[16] E.A.Wintner, M.M. Conn and J. Rebek, Jr. J. Am. Chem. Soc. 116 (1994) 8877.
[17] T.R. Cech. Science 236 (1987) 1532.

[18] G. Joyce. Nature 338 (1989) 217.
[19] E. Fahy, D.Y. Kwoh and T.R. Gingeras. PCR Methods and Applications 1 (1991) 25.
[20] G.T. Walker, M.C. Little, J.G. Nadeau and D.D. Shank. Proc. Natl. Acad. Sci. USA 89 (1992) 392.
[21] N. Walter and G. Strunk. Proc. Natl. Acad. Sci. USA 91 (1994) 7937.
[22] J.C. Guatelli, KM. Whitfield, D.Y.Kwoh, K.J. Barringer, D.D. Richman and T. R. Gingeras. Proc. Natl. Acad. Sci. USA 87 (1990) 1874.
[23] J. Compton. Nature 350 (1991) 91.
[24] M.M. Konarska and P.A. Sharp. Cell 57 (1989) 423.
[25] C.K. Biebricher and R. Luce. EMBO J. 15 (1996) 3458.
[26] R.R. Breaker and G. Joyce. Proc. Natl. Acad. Sci. USA 91 (1994) 6093.
[27] R.M. Fisher. Ann. Eugenics 7 (1937) 355.
[28] J.S. McCaskill. IMB Annual Report (1994) 118.
[29] K. Schmidt, P. Foerster, A. Bochmann and J.S. McCaskill. Proc. IBC Conference on microfabrication and microfluidic technologies, SanFrancisco, 12-13 Aug. (1996).
[30] A. Bochmann. Diploma Thesis, Friedrich-Schiller-Universität Jena (1997).
[31] P. Foerster, Y. Zhang and J. Ross. J. Phys. Chem. 97 (1993) 4708.
[32] J.S. McCaskill and K. Schmidt. *A linear cross-flow microreactor with local mixing.* In preparation (1997).
[33] K. Schmidt, P. Foerster, A. Bochmann and J.S. McCaskill. Proc. 1st International Conference on Microreaction Technology, 23-25 Feb. 1997, Dechema e.V. Frankfurt, Germany.
[34] B.Wagner, H. Mathis, K. Schmidt, G. Kalusche and J.S.McCaskill. Nucleosides and Nucleotides (1997) in press.
[35] M. Köllner, P.Fischer, J.S.McCaskill. S.P.I.E. Proc. Biomedical Optics Lille (1994)112.
[36] J.S McCaskill, J. Ackermann and J. Breyer. *Efficient dataflow architectures for diffusion in chemistry and population dynamics.* In preparation (1997).
[37] J. Ackermann, B. Böddeker, J. Breyer, U.Tangen and J.S. McCaskill. *A minimum logic parallel pseudo-random number generator for configurable hardware.* IEEE Transactions on Parallel and Distributed Systems, submitted (1996).
[38] Xilinx, Inc. (1995). *The XC6200 FPGA family advance product information sheet.*
[39] A. Thompson, *Silicon Evolution.* In J.R.Koza et.al., eds., Proc. of Genetic Programming (1996) MIT Press, 444.
[40] U. Tangen, L. Schulte and J.S. McCaskill. A parallel hardware evolvable computer POLYP. IEEE Transactions (1997) in press.
[41] S.A.Kauffman and E. Weinberger, J. Theor. Biol. 141 (1989) 211.
[42] A.J. Lotka (1925) *Elements of physical biology.* (Dover, New York, 1956)
[43] T.Ikeda and M.Mimura, J. Math.Biol. 31 (1993) 215.
[44] J. Emlen. *Population Biology - The Coevolution of Population Dynamics and Behaviour.* (MacMillan, NY, Collier, London, 1984).
[45] S. Dunbar, J. Math. Biol. 17 (1983) 11.
[46] B. Wlotzka and J.S. McCaskill. Chemistry and Biology 4 (1997) 25.
[47] J. Ackermann, B. Wlotzka and J.S. McCaskill. In vitro DNA-based Predator-Prey System with Oscillatory Kinetics. Bull. Math. Biol. submitted (1996).
[48] M.C. Boerlisjt and P. Hogeweg. Proc. R. Soc. Lond. 253 (1993) 15.
[49] G.F. Joyce. Science 257 (1992) 635.
[50] C. Tuerk and L. Gold. Science 249 (1990) 505.
[51] M.C. Wright and G.F. Joyce. Science 276 (1997) 614.
[52] M. Eigen and P. Schuster. Naturwissenschaften 64 (1977) 541 *ibid* 64 (1978) 7 *ibid* 65 (1978) 341.
[53] J. Maynard Smith. Nature 280 (1979) 445.
[54] C. Bresch, U. Niesert and D. Harnasch, J.Theor. Biol.85 (1980) 399.
[55] E. Szathmary and L. Demeter. J. Theor. Biol. 128 (1987) 463.
[56] M.C. Boerljist and P. Hogeweg. Physica D 48 (1991) 17.
[57] M.C. Boerlijst and P. Hogeweg Physica D 88 (1995) 29.
[58] M.B. Cronhjort and A. M. Nyberg. Physica D 88 (1995) 1.
[59] J.E. Pearson. Science 261 (1993) 189.
[60] J.S. McCaskill, *Origins of Molecular Cooperation: Theory and Experiment,* Inaugural lecture at the Friedrich Schiller University, Jena, Germany, 31st May (1994) IMB Press.
[61] M.B. Cronhjort and C. Blomberg. Physica D 101 (1997) 289.
[62] B. Böddeker and J.S. McCaskill. J. Theor. Biol. submitted (1996).
[63] K-J. Lee, W.D. McCormick, J.E. Pearson and H.L. Swinney, Nature 369 (1994) 215.
[64] S.K.Scott and K. Showalter, J. Phys. Chem. 96 (1992) 8702.
[65] R. Ehricht, T. Ellinger and J.S.McCaskill, Eur. J. Biochem. 243 (1997) 358.
[66] J. Breyer, J. Ackermann, U. Tangen, and J.S. McCaskill. *Modeling the evolution in cooperatively coupled in vitro amplification systems.* Poster submitted to ECAL97.