Modeling Evolutionary Processes:

Major and Minor Transitions in Evolution

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

Steps in Evolution: Perspectives from Physics, Biochemistry and Cell Biology – 150 Years after Darwin

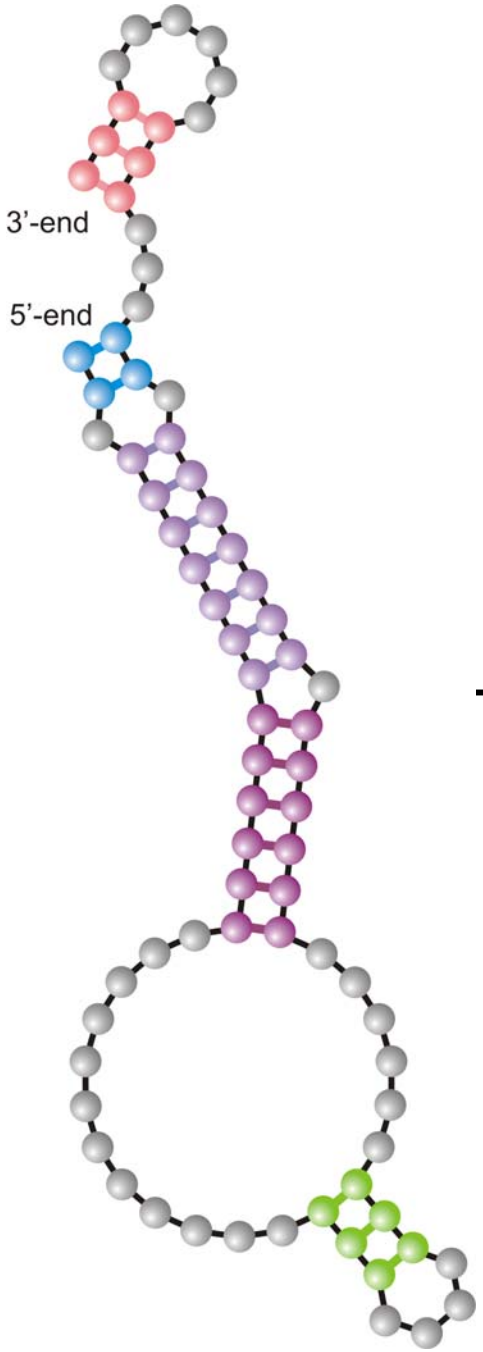
Bremen, 28.06.– 05.07.2009

Web-Page for further information:

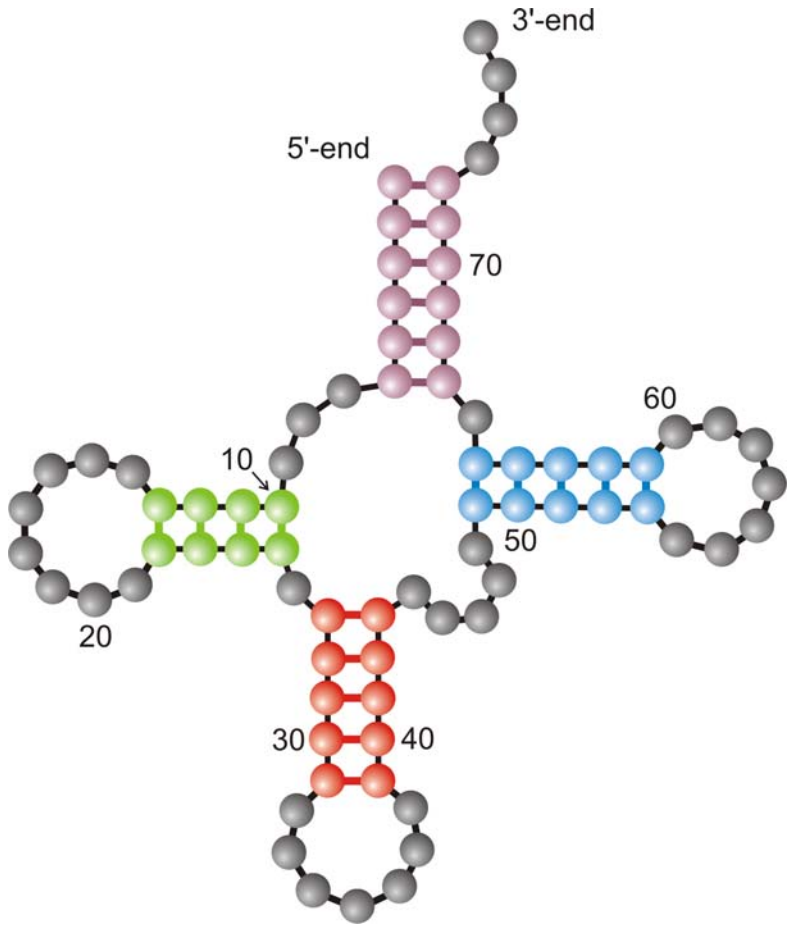
http://www.tbi.univie.ac.at/~pks

1. Reconstruction of evolutionary processes
2. Diffusion in sequence space and shape space
3. Continuous and discontinuous transitions
4. Mechanism of RNA optimization
5. Major transitions in evolution

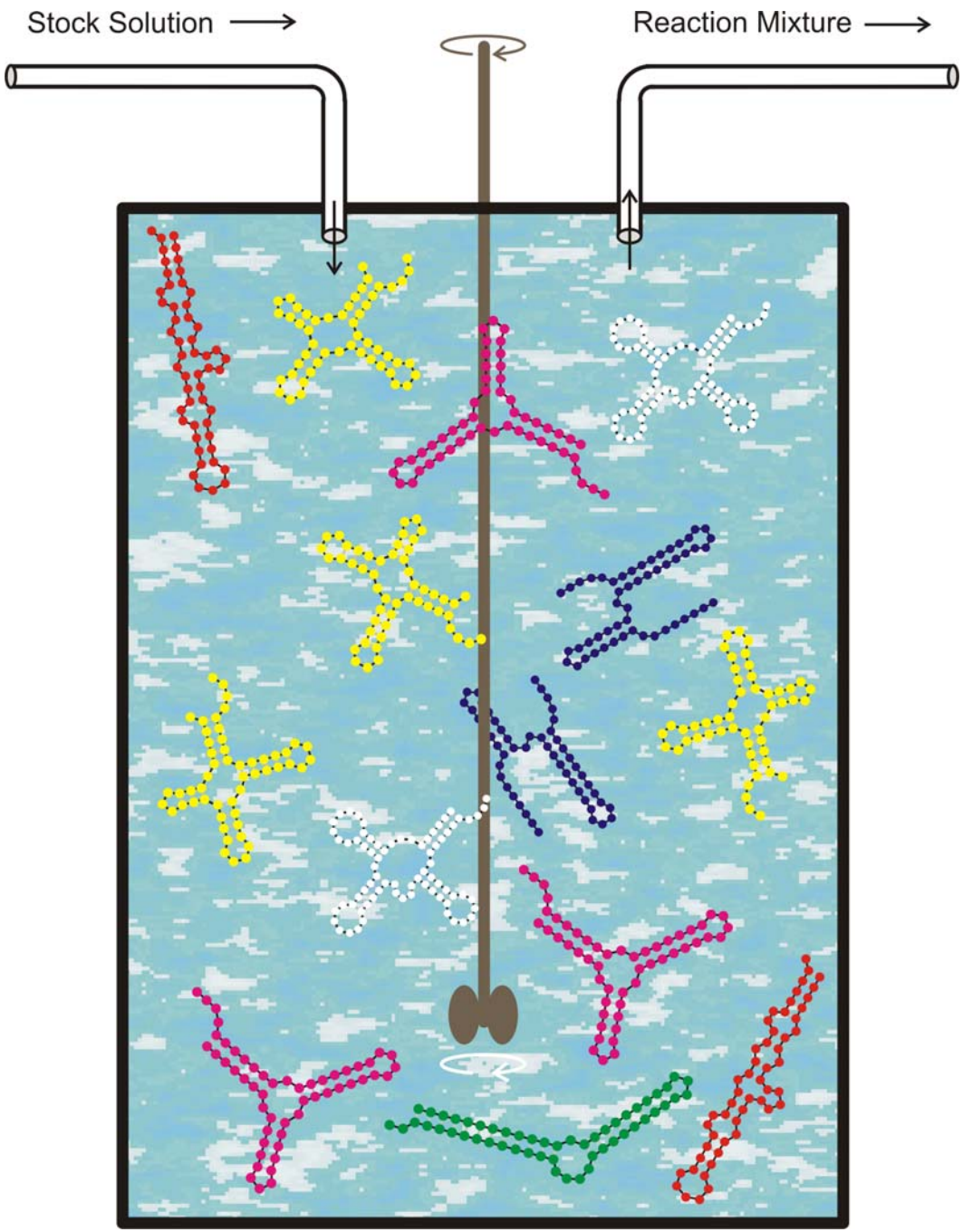
1. **Reconstruction of evolutionary processes**
2. Diffusion in sequence space and shape space
3. Continuous and discontinuous transitions
4. Mechanism of RNA optimization
5. Major transitions in evolution



Structure of randomly chosen initial sequence



Phenylalanyl-tRNA as target structure



Replication rate constant:

$$f_k = \gamma / [\alpha + \Delta d_S^{(k)}]$$

$$\Delta d_S^{(k)} = d_H(S_k, S_\tau)$$

Selection constraint:

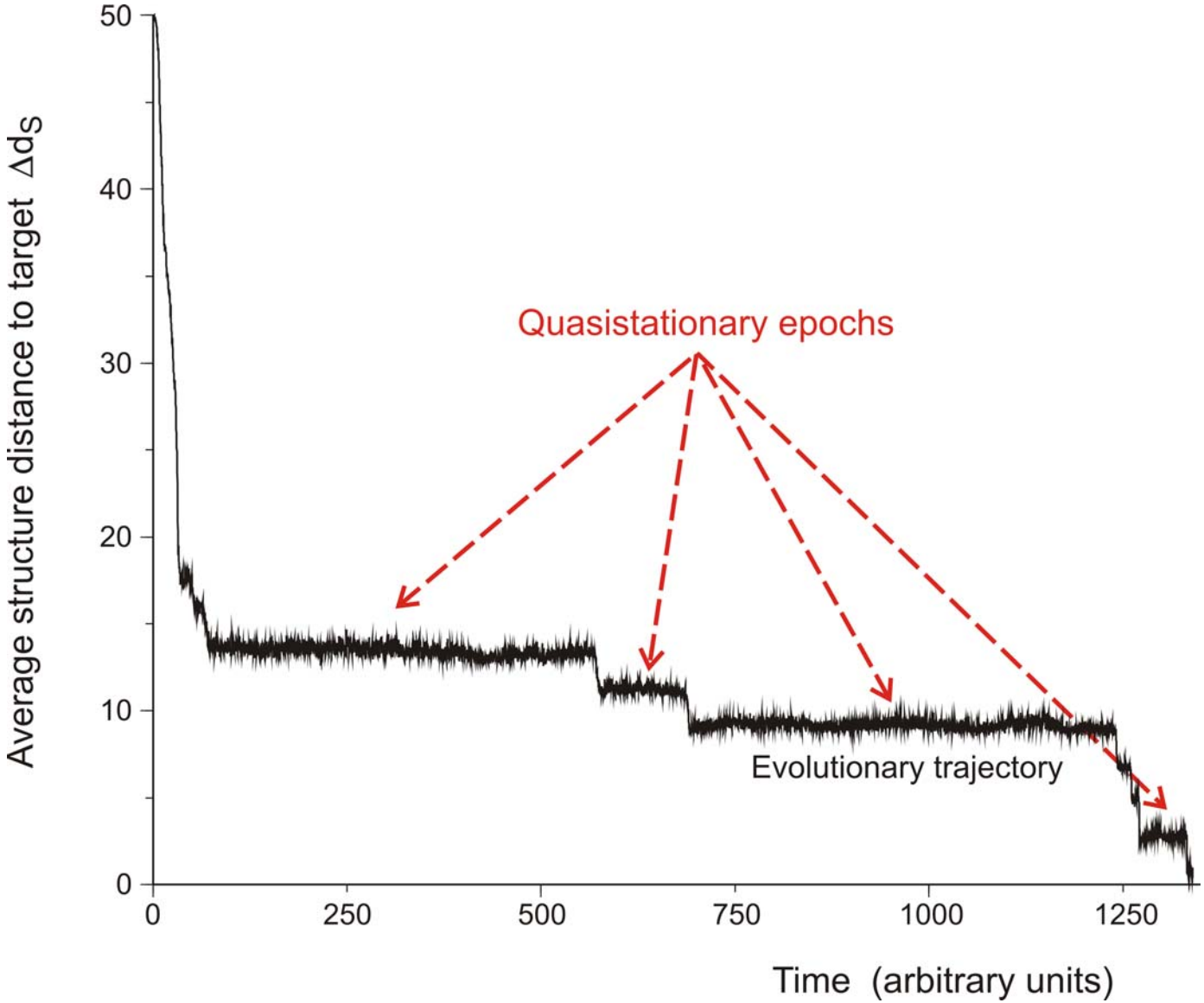
Population size, N = # RNA molecules, is controlled by the flow

$$N(t) \approx \overline{N} \pm \sqrt{\overline{N}}$$

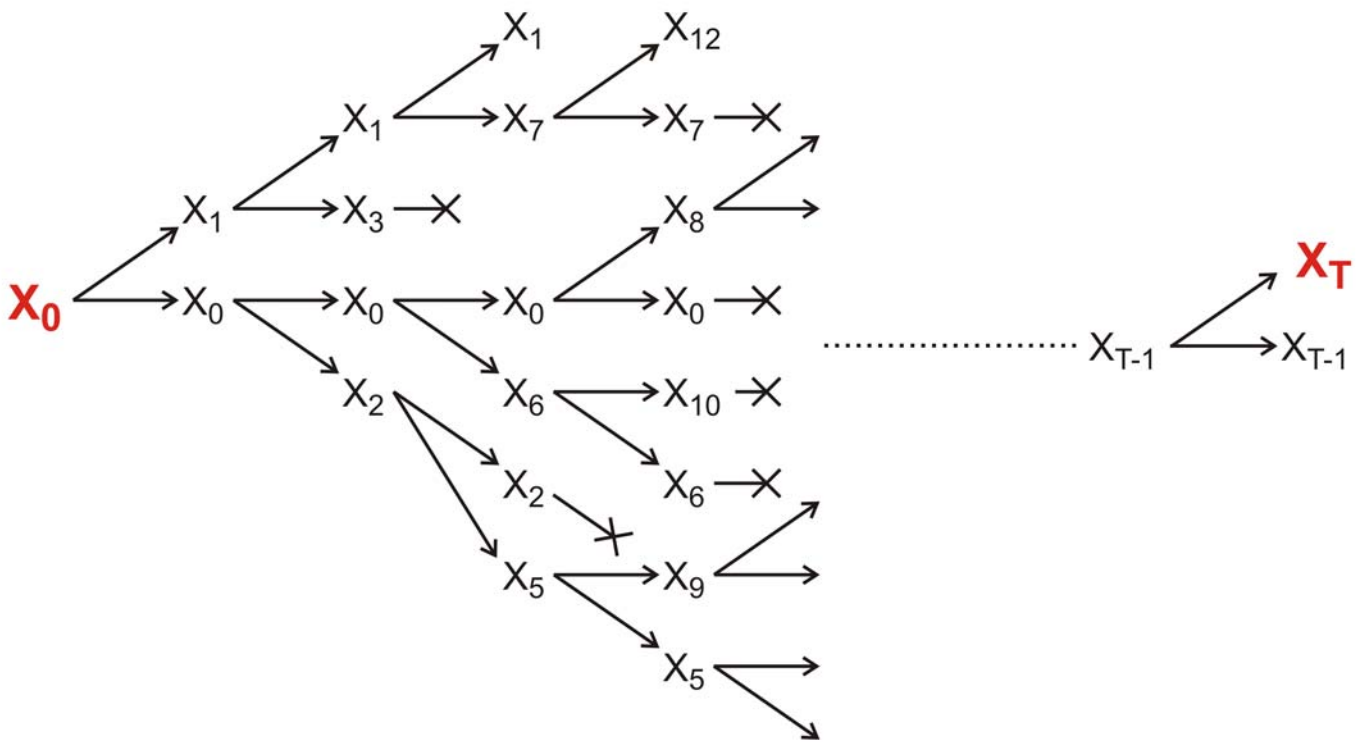
Mutation rate:

p = 0.001 / site × replication

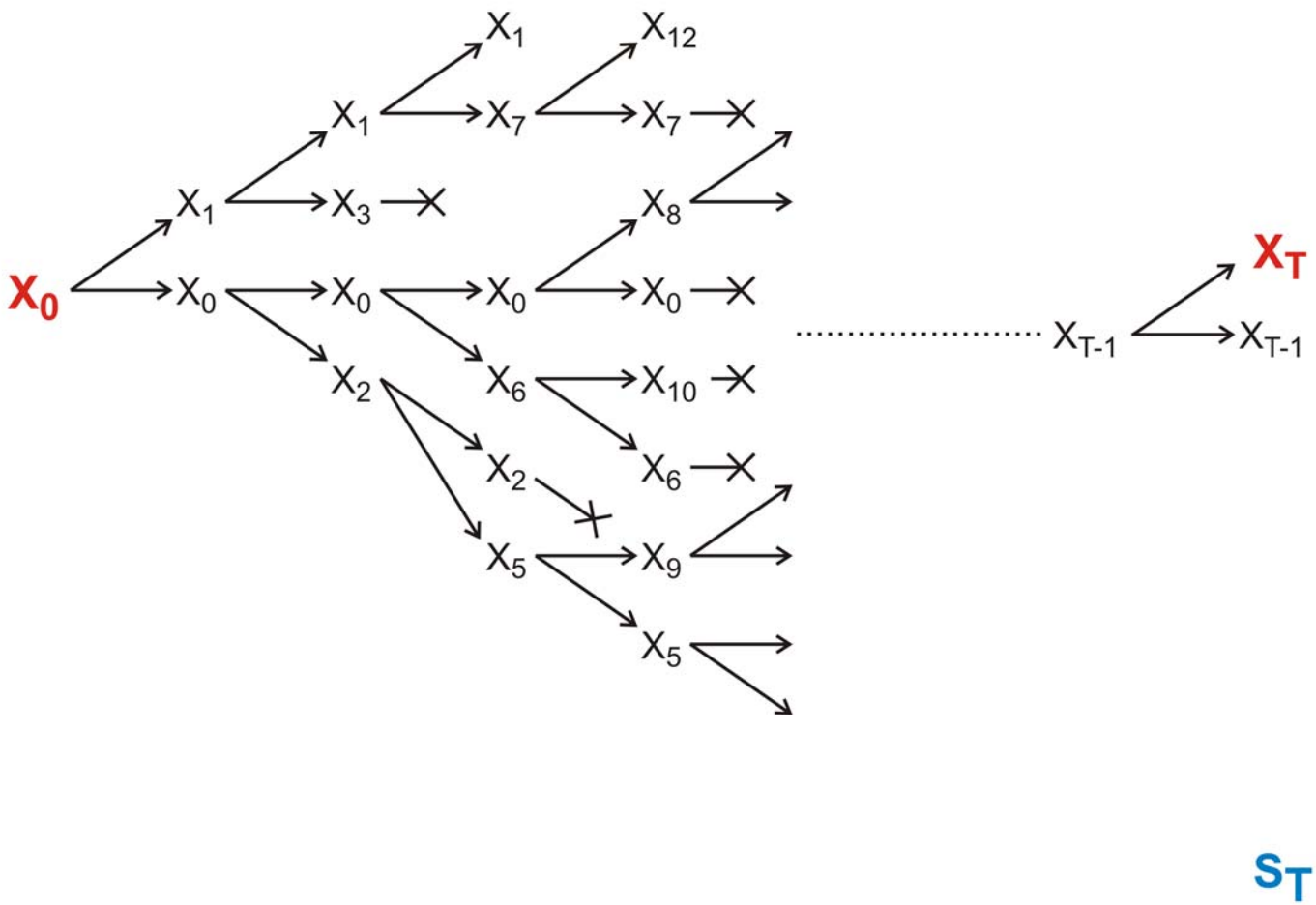
The flowreactor as a device for studies of evolution ***in vitro*** and ***in silico***



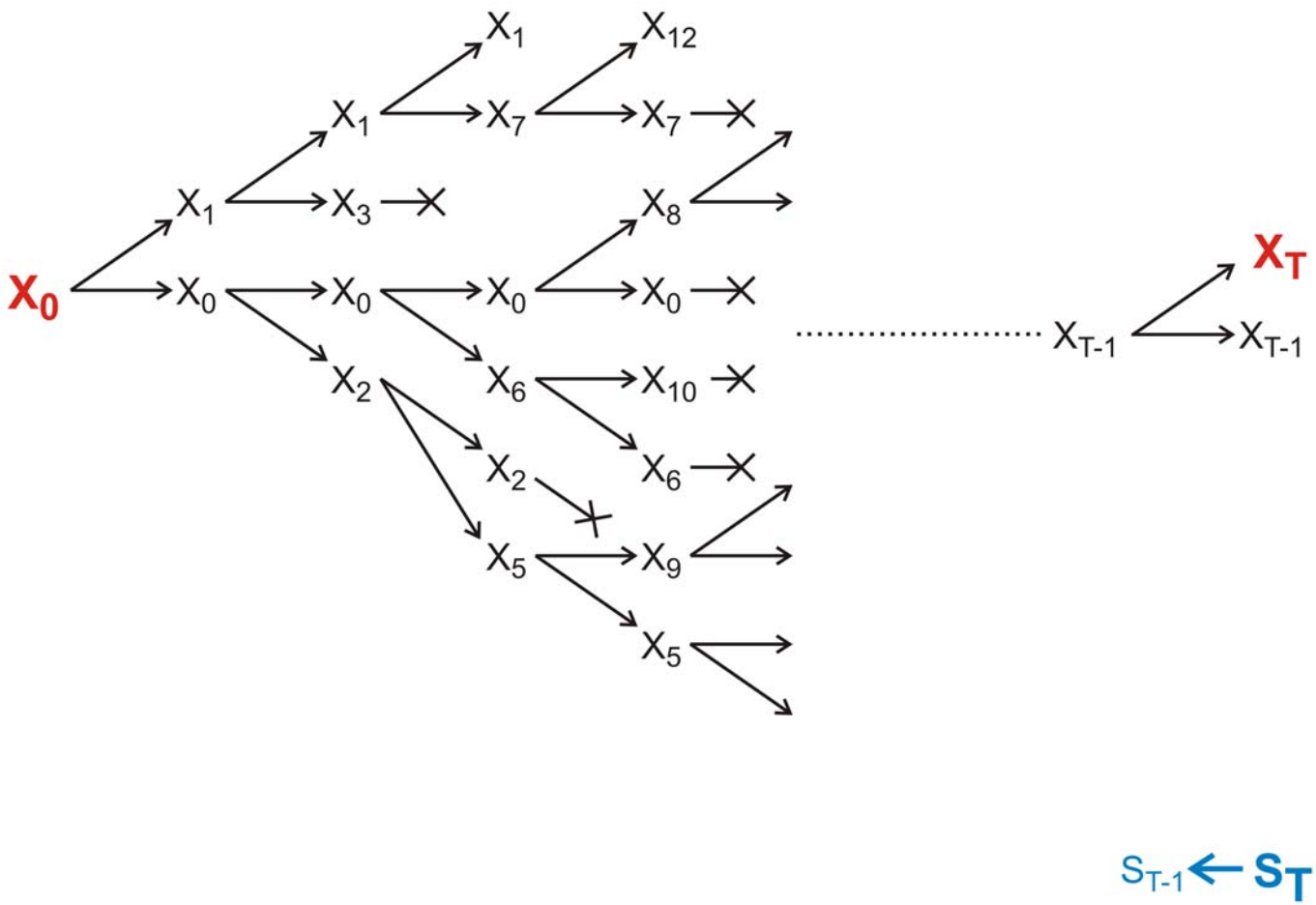
In silico optimization in the flow reactor: Evolutionary Trajectory



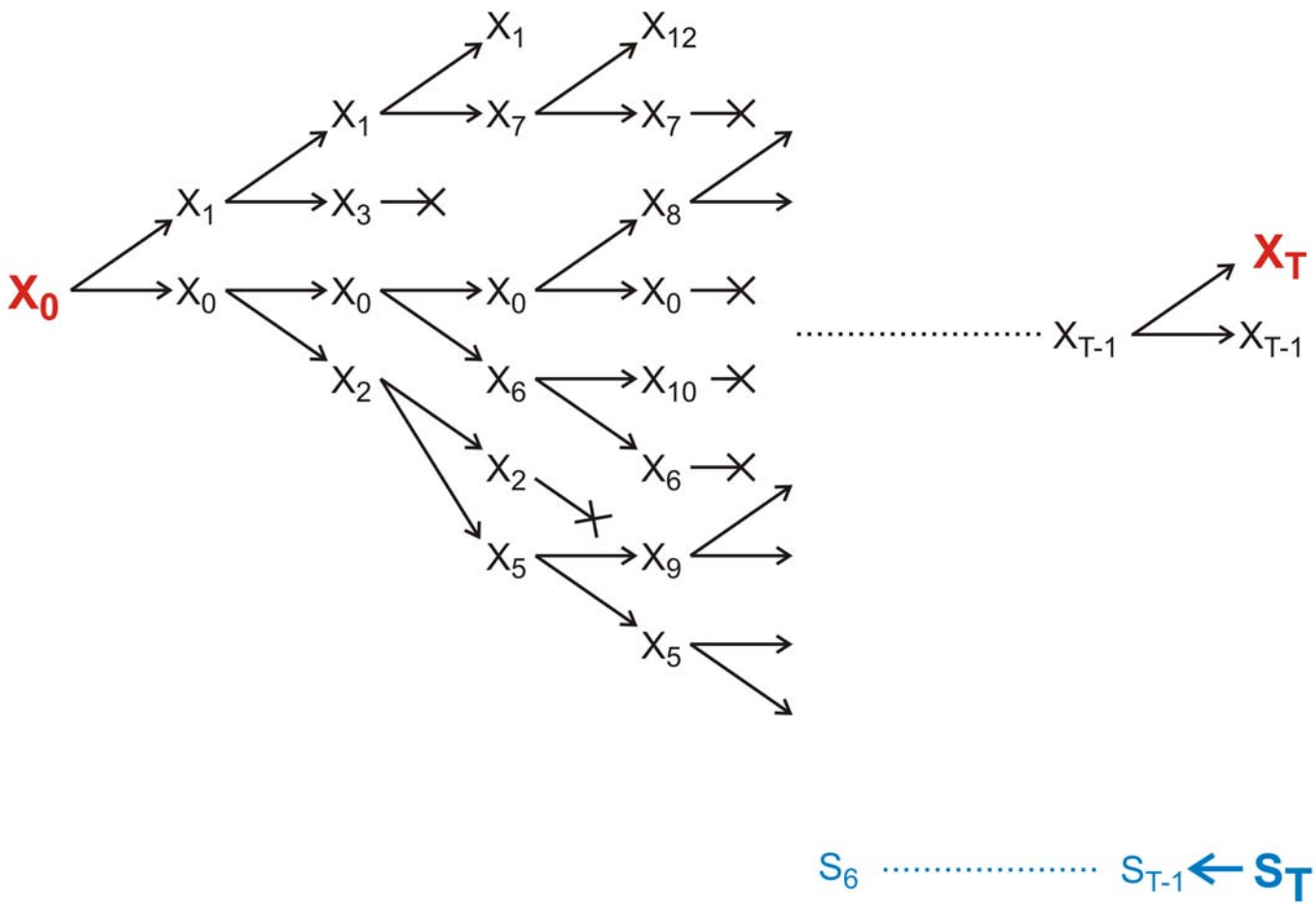
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



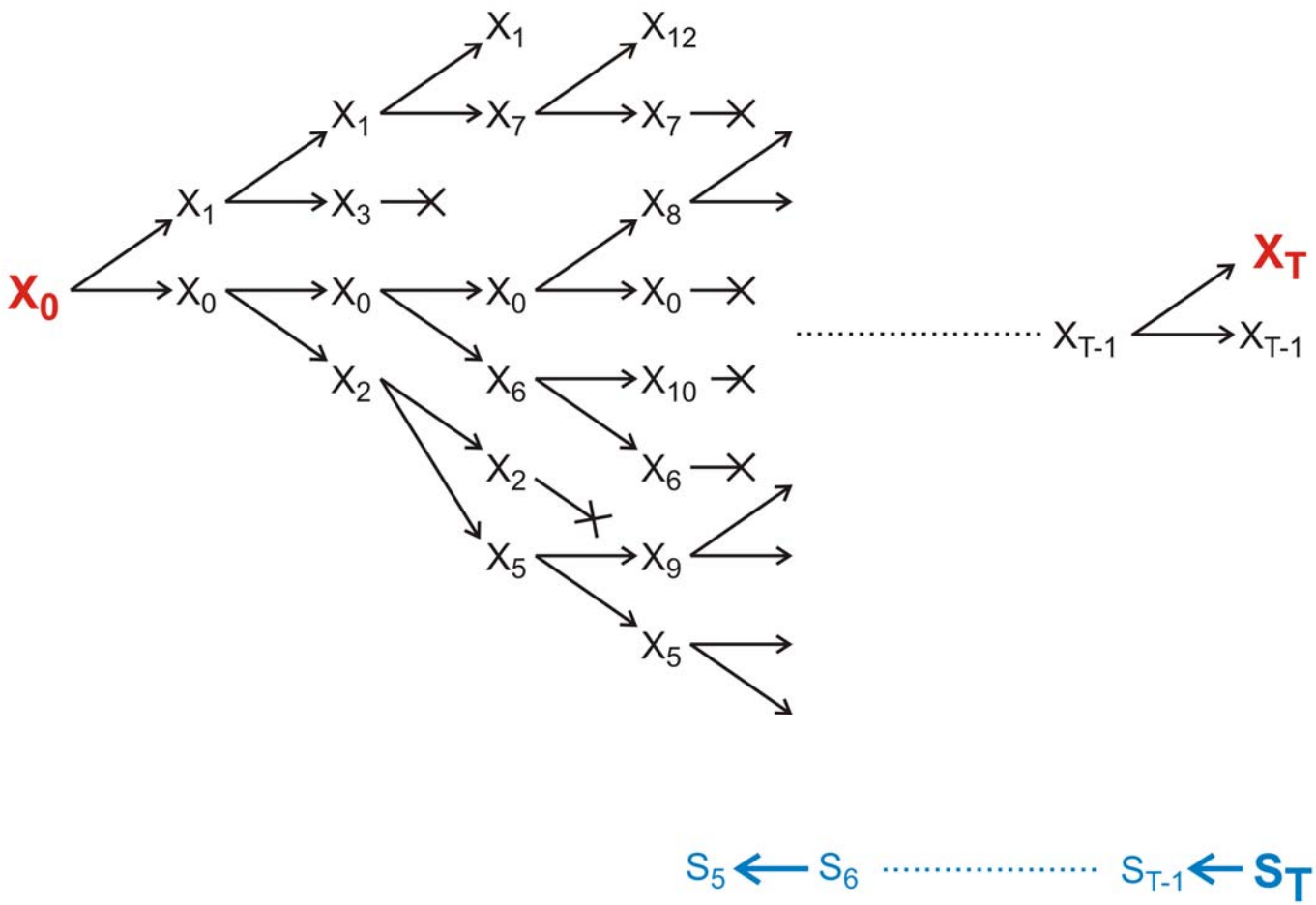
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



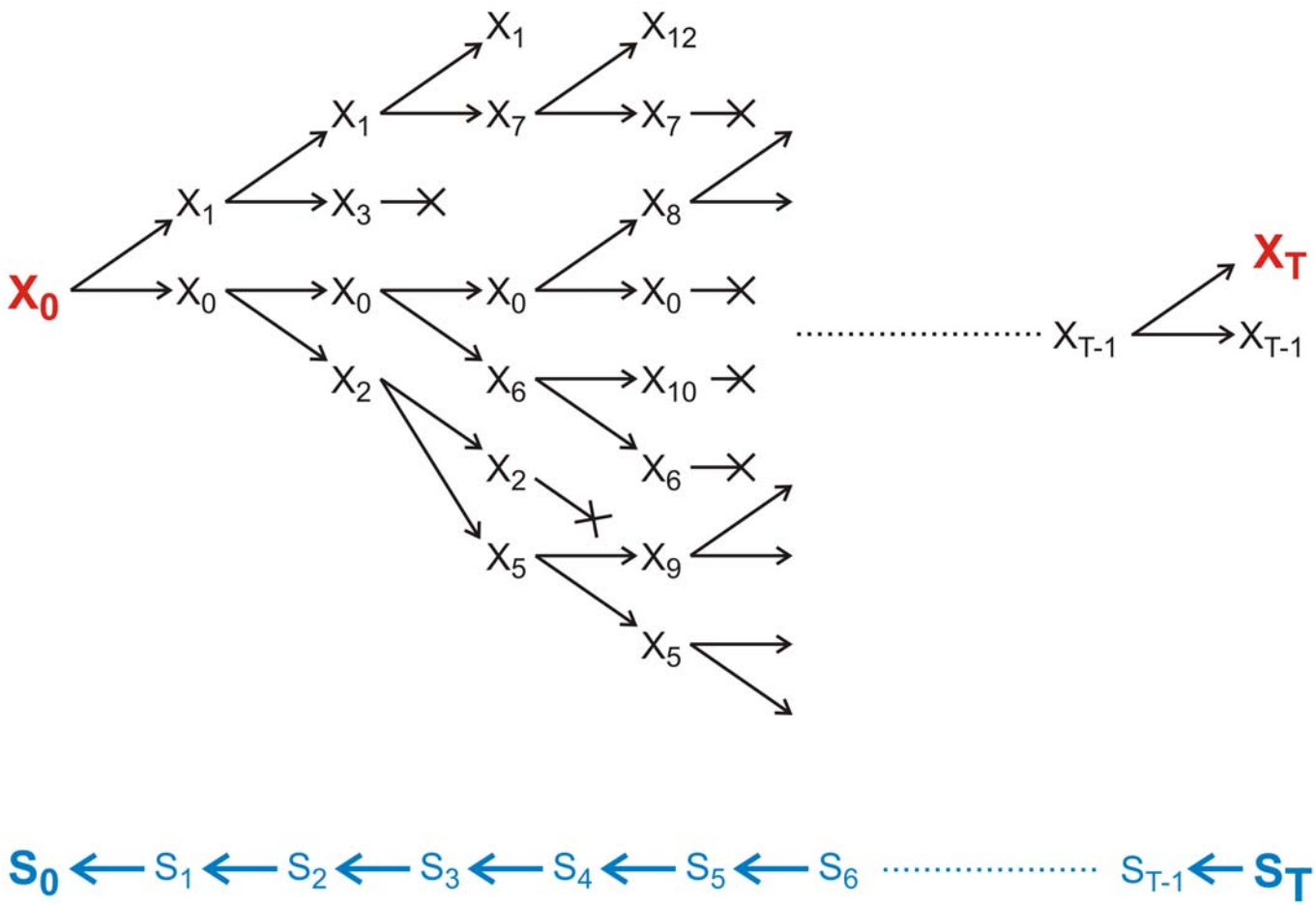
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



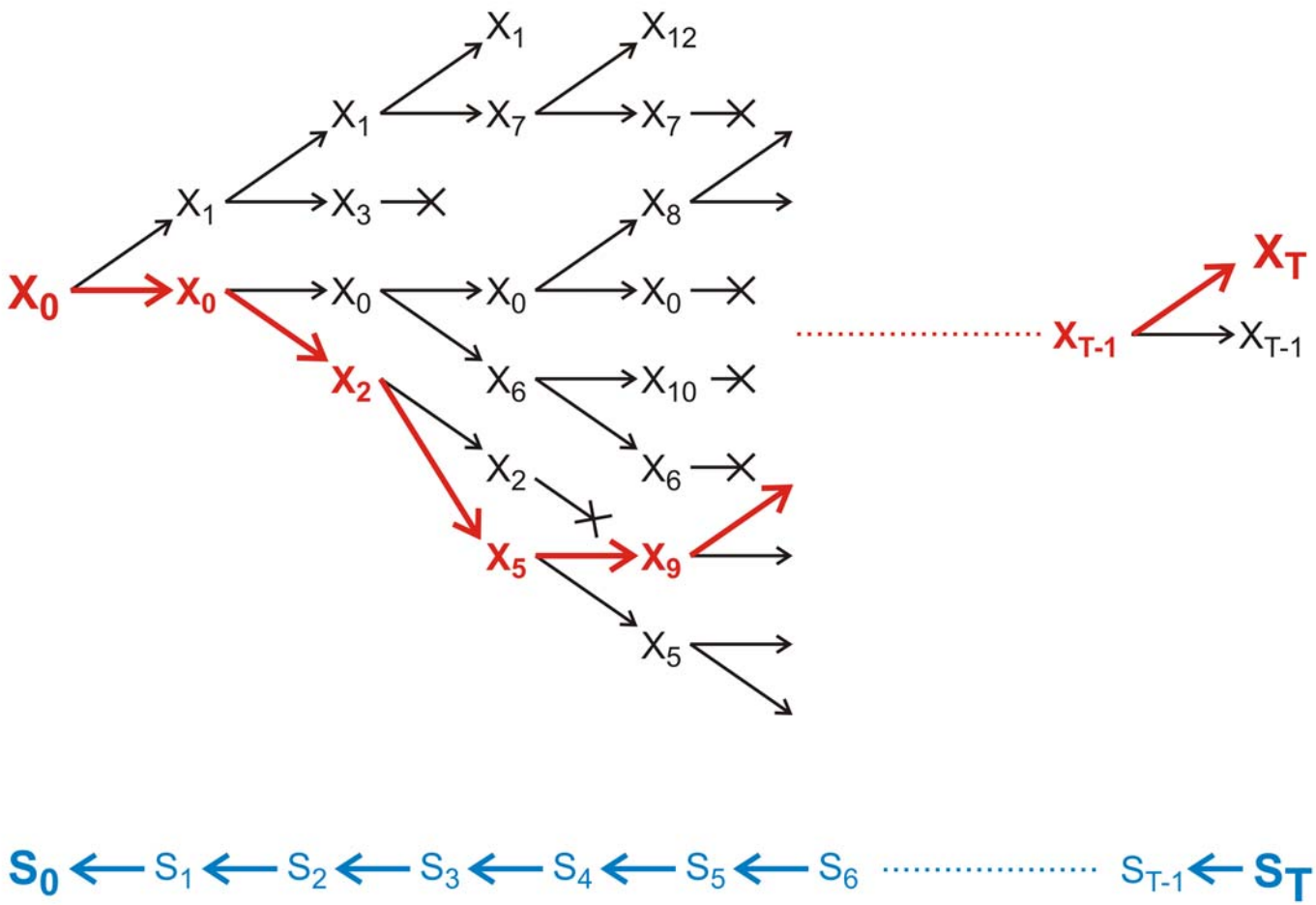
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



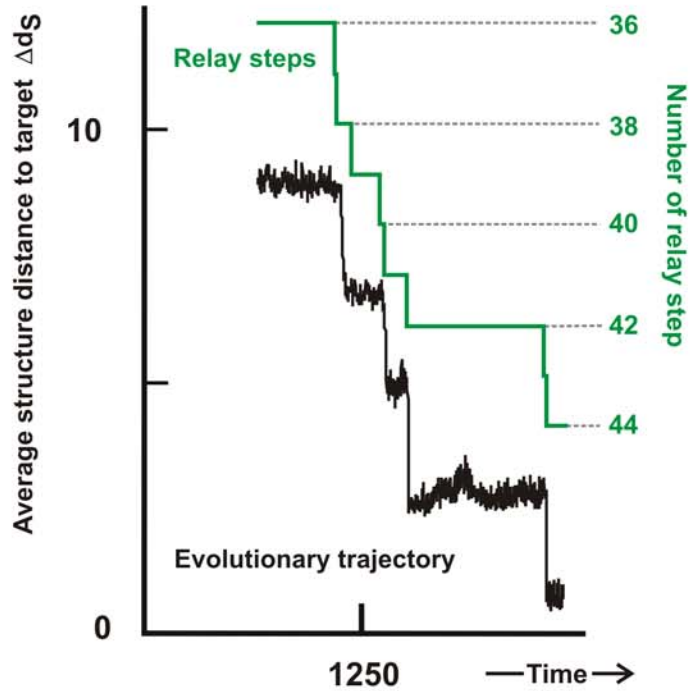
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



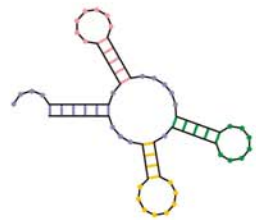
Evolution of RNA molecules as a Markow process and its analysis by means of the relay series



Evolution of RNA molecules as a Markow process and its analysis by means of the relay series

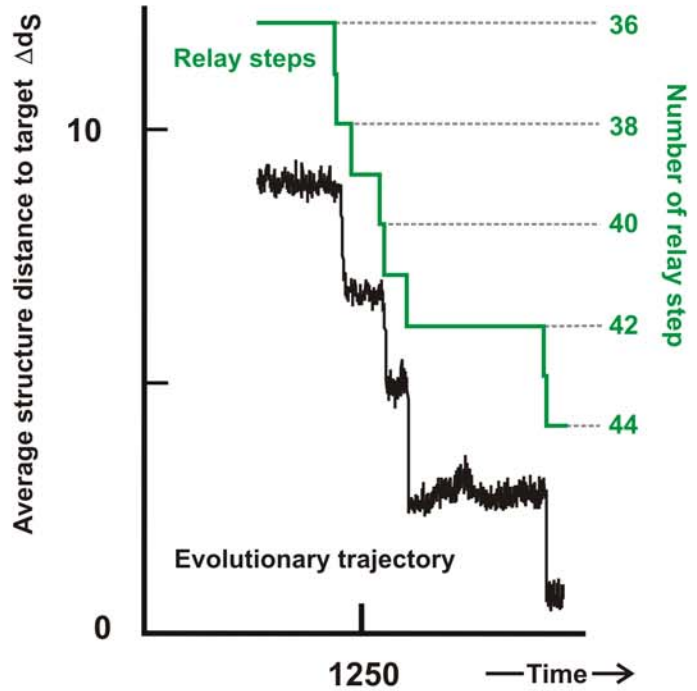


Evolutionary process



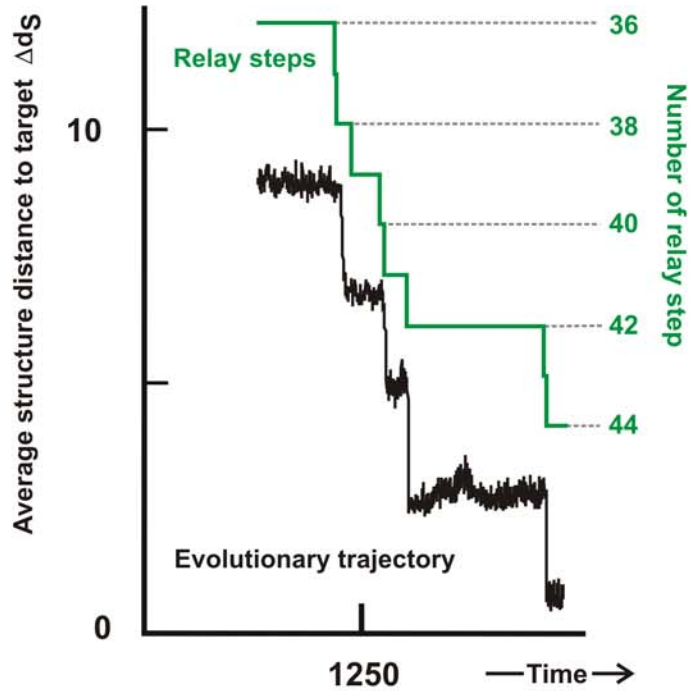
44

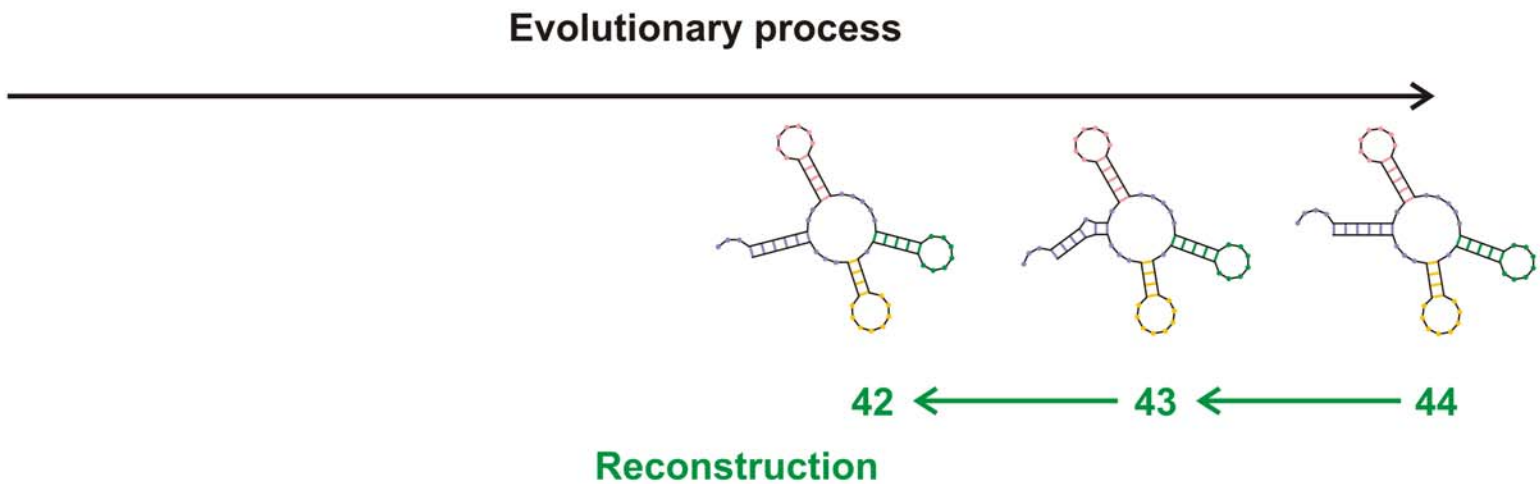
Reconstruction

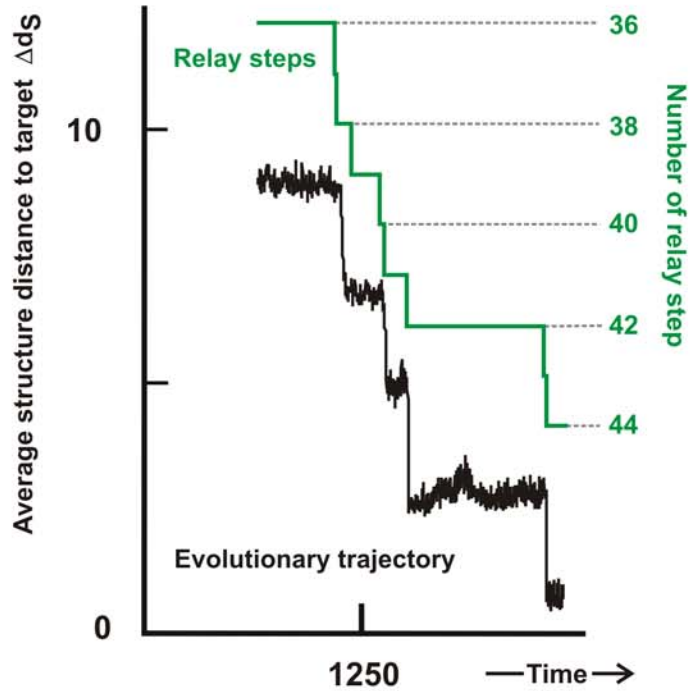


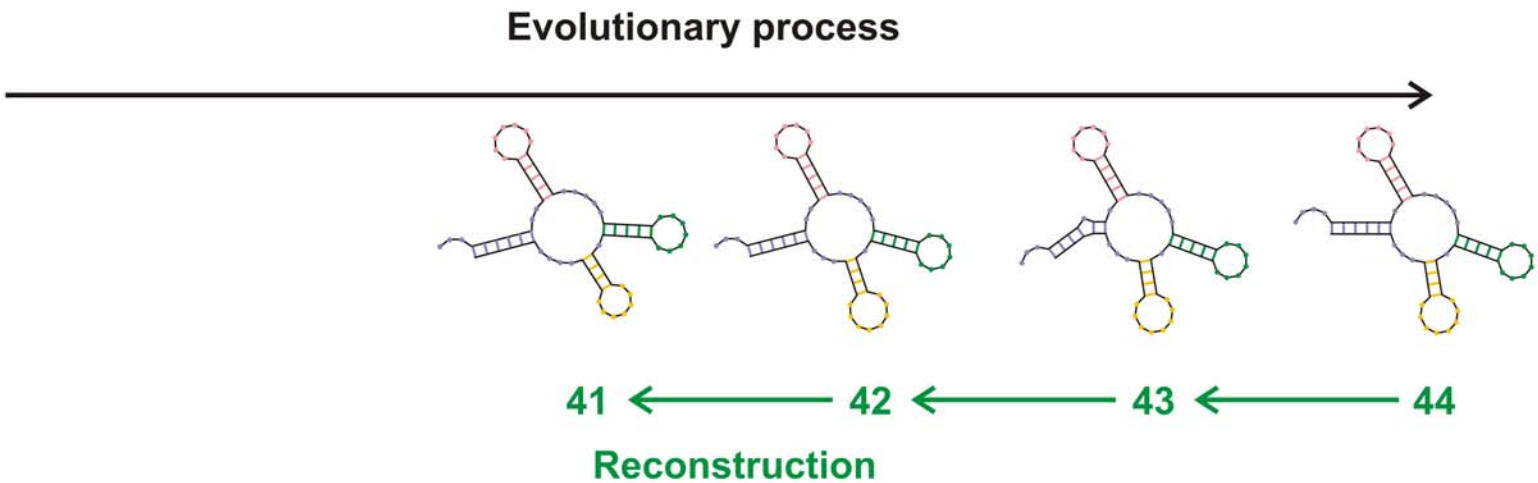
Evolutionary process

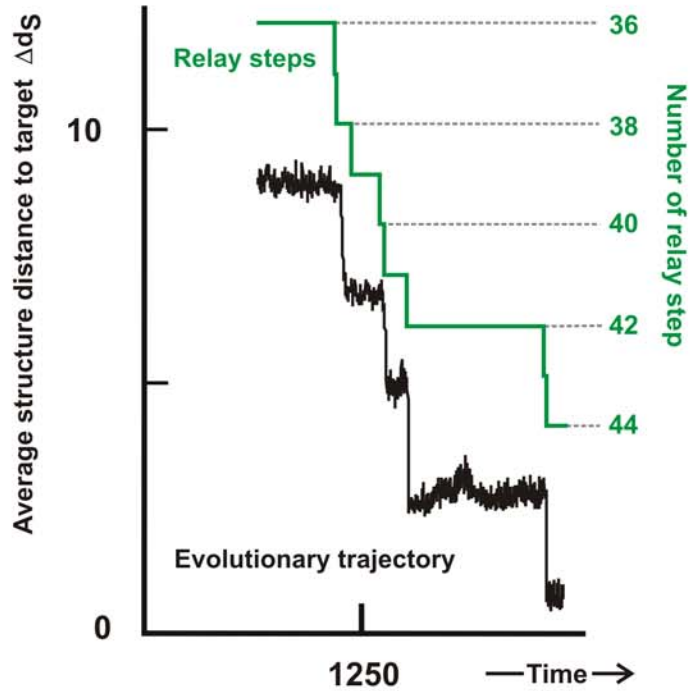
Reconstruction

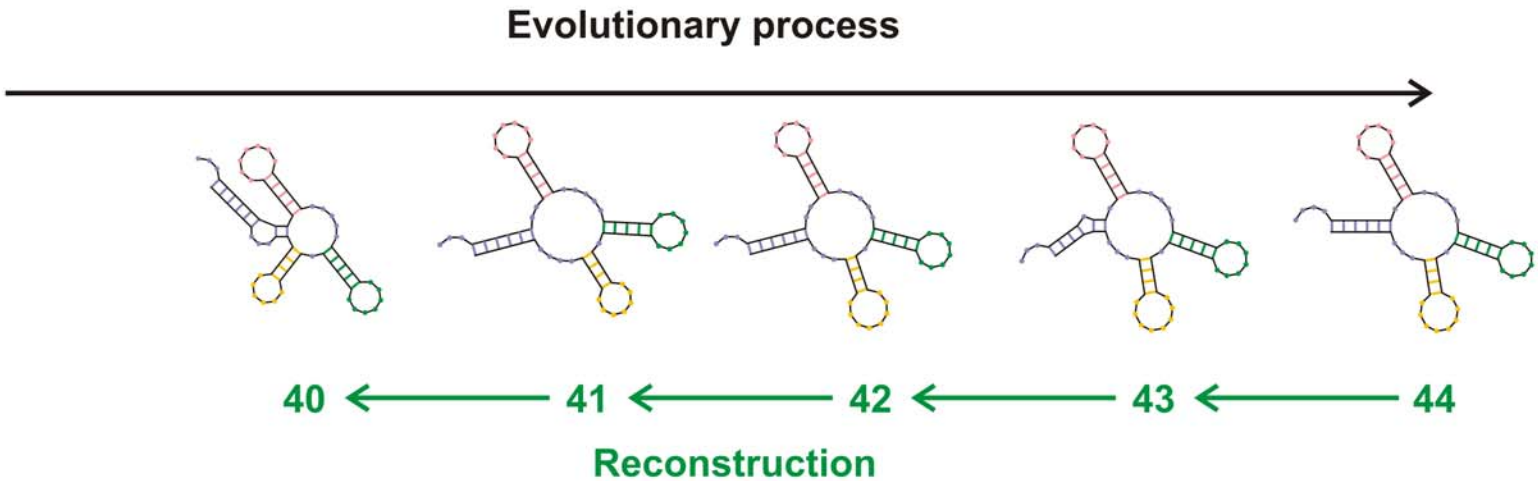

Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time

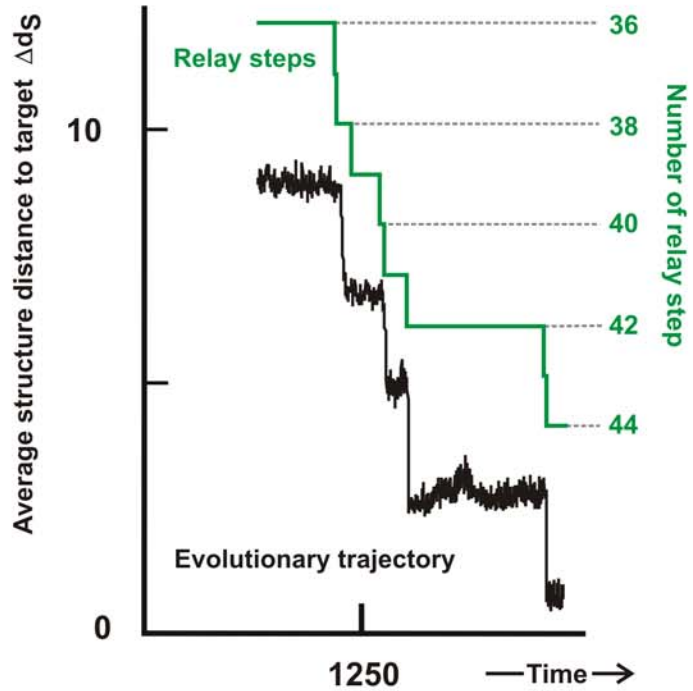

Evolutionary process
42
43
44
Reconstruction

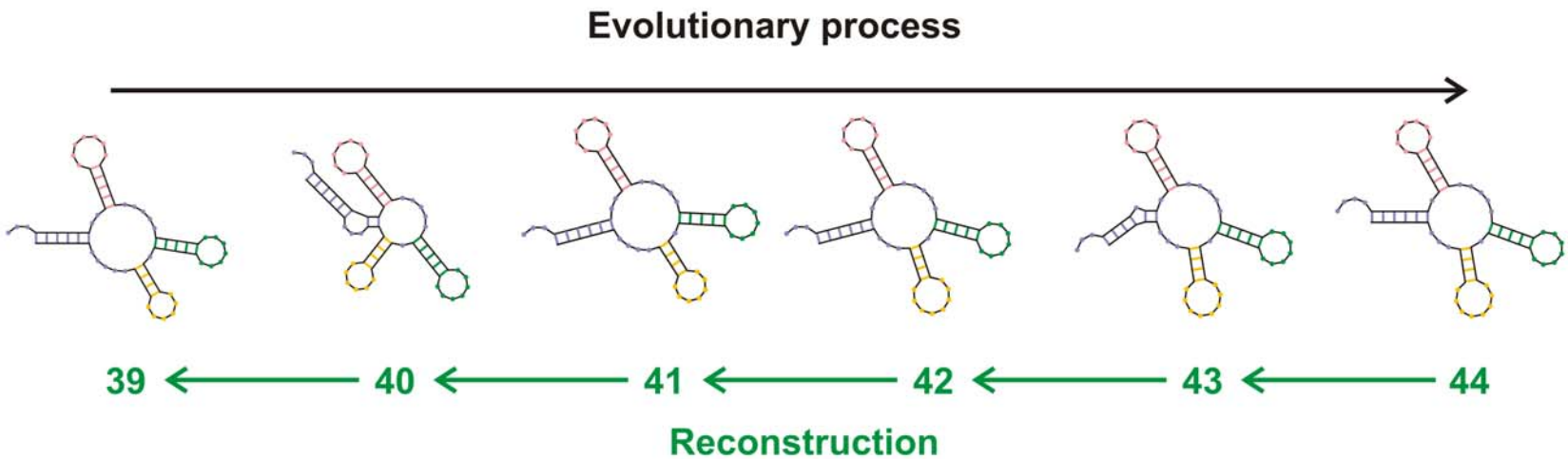

Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time

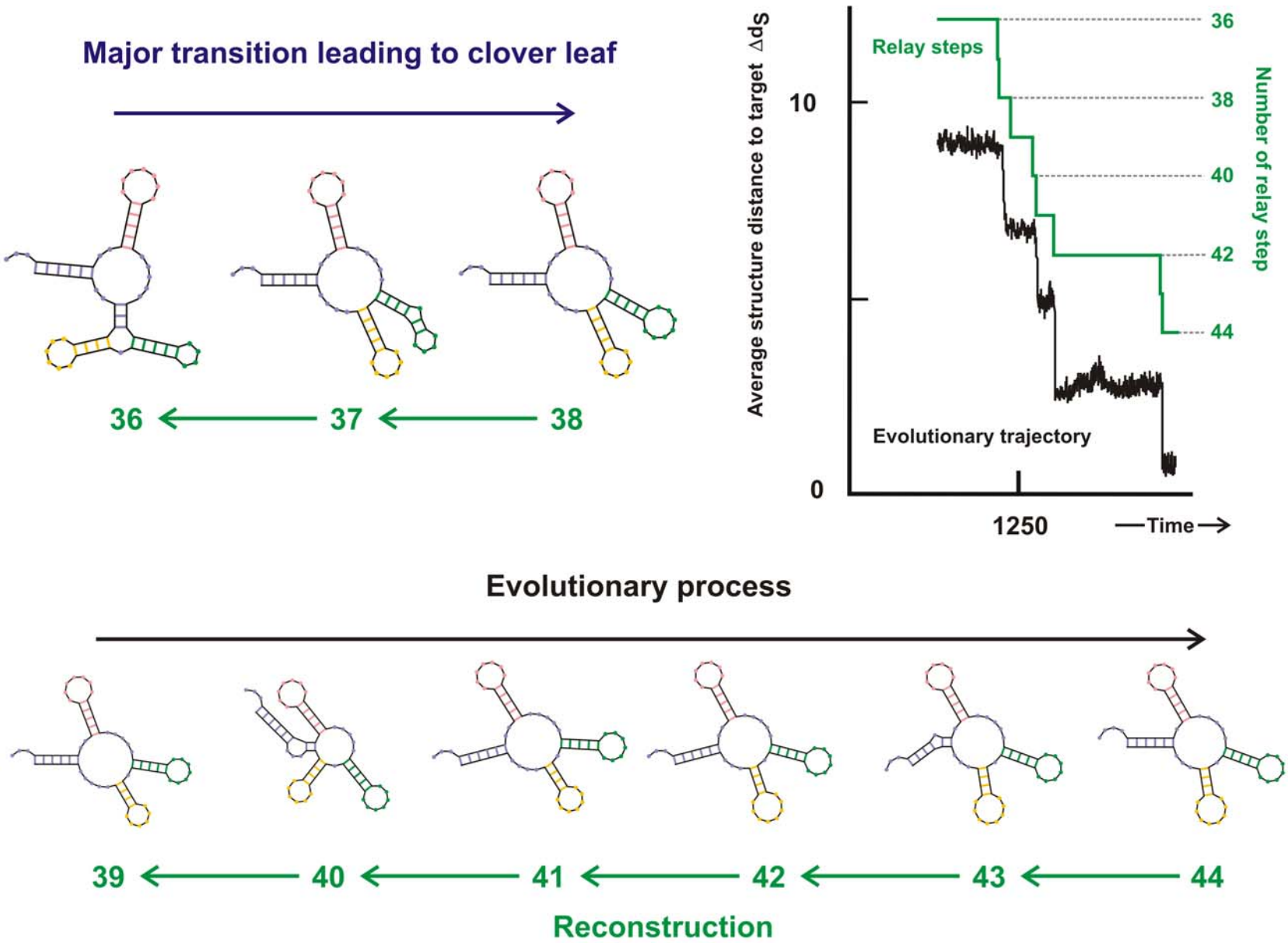

Evolutionary process
41
42
43
44
Reconstruction


Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time


Evolutionary process
40
41
42
43
44
Reconstruction

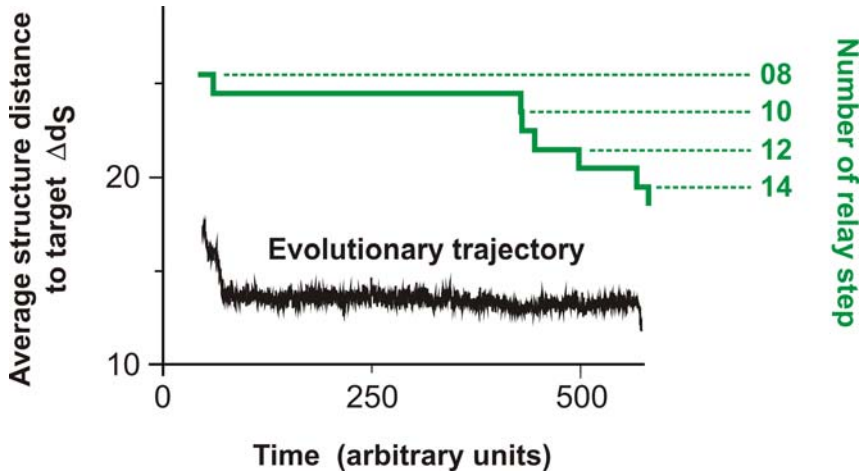

Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time


Evolutionary process
39
40
41
42
43
44
Reconstruction


Major transition leading to clover leaf
36
37
38
Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time
Evolutionary process
39
40
41
42
43
44
Reconstruction

1. Reconstruction of evolutionary processes
2. **Diffusion in sequence space and shape space**
3. Continuous and discontinuous transitions
4. Mechanism of RNA optimization
5. Major transitions in evolution

28 neutral point mutations during a long quasi-stationary epoch

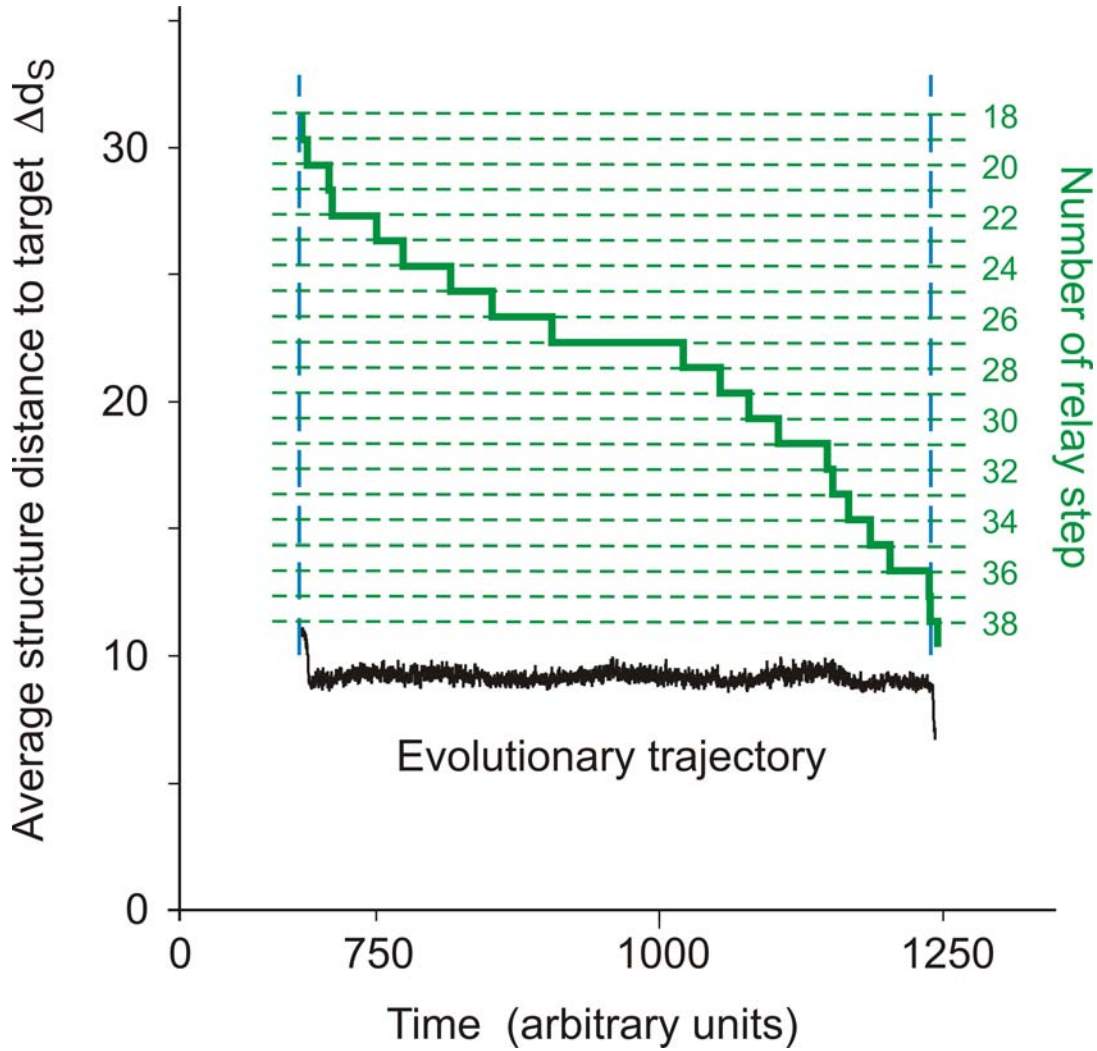


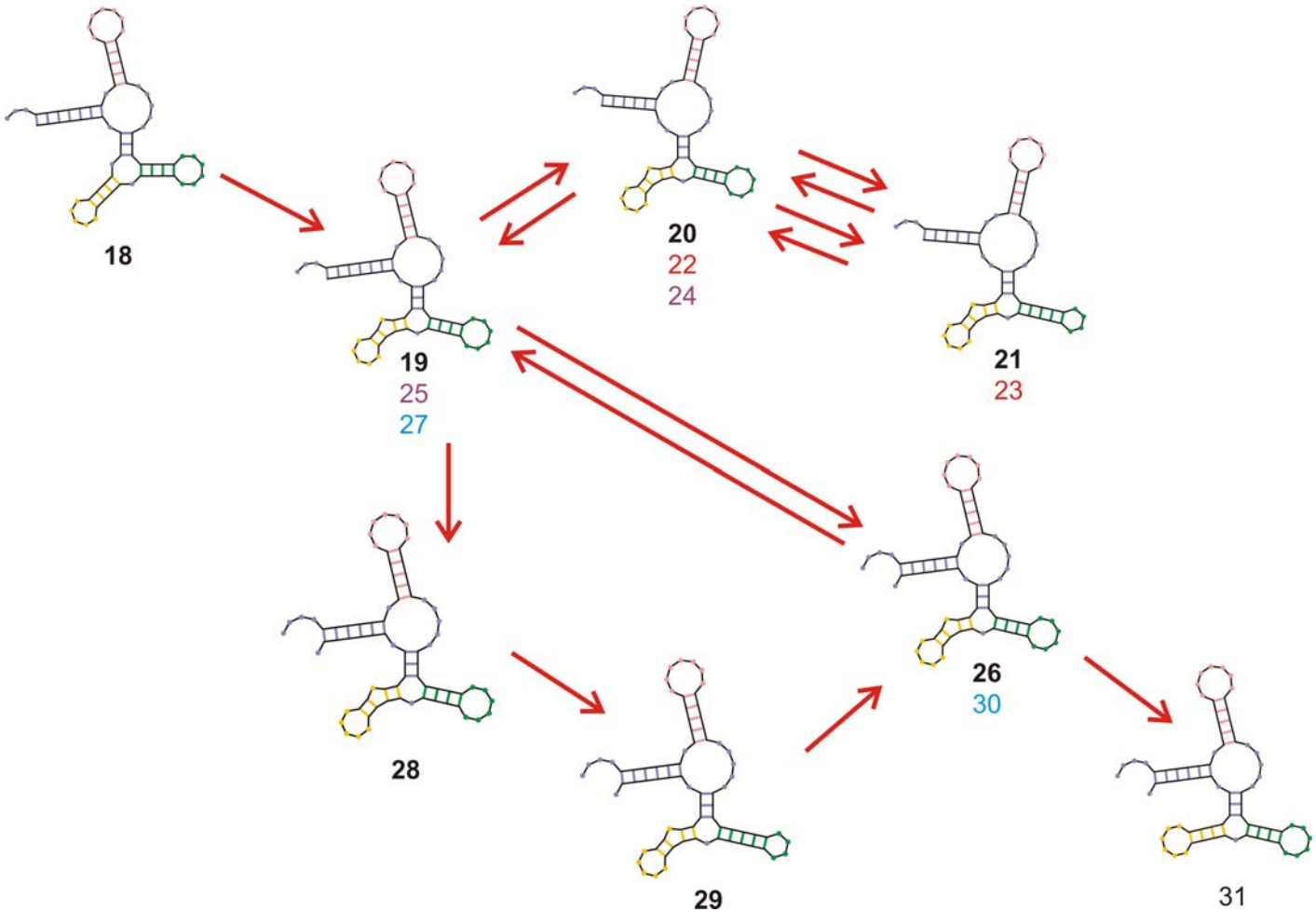

```
entry   GGUAUGGGCGUUGAAUAGUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCGUACAGAA
8       .((((((((((((.........(((....)))).......)))))).....((((.......))))))))))....
exit    GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCAUACAGAA
entry   GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUACCAUACAGAA
9       .((((((.(((((.........(((....)))).......)))))).....((((.......))))).))))))....
exit    UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACACCGUCCCAAG
entry   UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCCAAG
10      .(((((..(((((.........(((....)))).......)))))).....((((.......))))).))))))....
exit    UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCCAAG
```

Transition inducing point mutations change the molecular structure

Neutral point mutations leave the molecular structure unchanged

Neutral genotype evolution during phenotypic stasis

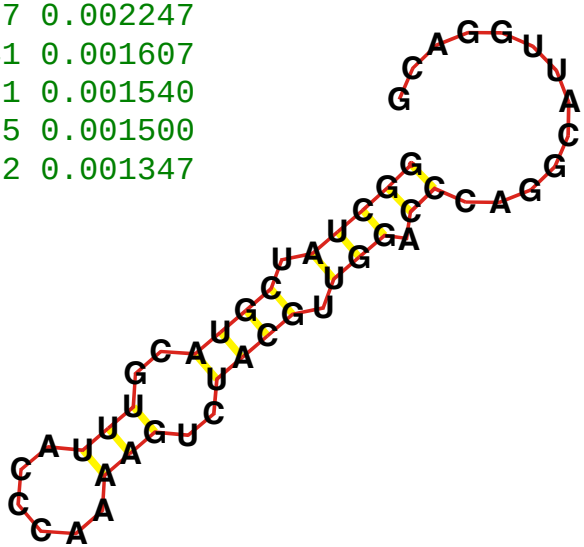

Average structure distance to target Δd_S
30
20
10
0
0
750
1000
1250
Time (arbitrary units)
Evolutionary trajectory
18
20
22
24
26
28
30
32
34
36
38
Number of relay step


18
19
25
27
20
22
24
21
23
28
29
26
30
31

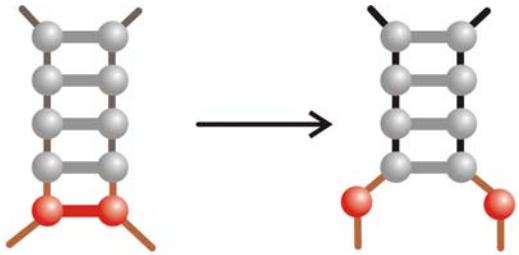
1. Reconstruction of evolutionary processes
2. Diffusion in sequence space and shape space
3. **Continuous and discontinuous transitions**
4. Mechanism of RNA optimization
5. Major transitions in evolution

	Number	Mean Value	Variance	Std.Dev.
Total Hamming Distance:	150000	11.647973	23.140715	4.810480
Nonzero Hamming Distance:	99875	16.949991	30.757651	5.545958
Degree of Neutrality:	50125	**0.334167**	0.006961	**0.083434**
Number of Structures:	**1000**	**52.31**	85.30	**9.24**

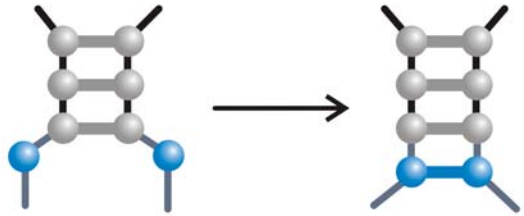
1	(((((.((((..(((.....))))..)))))..))).).....	**50125**	**0.334167**
2	..(((.((((..(((.....))))..)))))..))).....	**2856**	**0.019040**
3	((((((((((..(((.....))))..)))))))).)).....	**2799**	**0.018660**
4	(((((.((((..((((....))))..)))))..))).)).....	**2417**	**0.016113**
5	(((((.((((.((((.....))))).)))))..))).)).....	**2265**	**0.015100**
6	(((((.(((((.(((.....))).)))))..))).)).....	**2233**	**0.014887**
7	(((((..(((..(((.....)))..)))..)))..)).....	1442	0.009613
8	(((((.((((..((.......))..)))))..))).)).....	1081	0.007207
9	((((..((((..(((.....)))..)))))..)).)).....	1025	0.006833
10	(((((.((((..(((.....)))..)))))..))))).....	1003	0.006687
11	.((((.((((..(((.....)))..)))))..)))).....	963	0.006420
12	(((((.(((...(((.....)))...)))..))).)).....	860	0.005733
13	(((((.((((..(((.....)))..)))))..)).))).....	800	0.005333
14	(((((.((((...((.....))...)))))..))).)).....	548	0.003653
15	(((((.((((..............))))..))).)).....	362	0.002413
16	((.((.((((..(((.....)))..)))))..))..)).....	337	0.002247
17	(.(((.((((..(((.....)))..)))))..))).).....	241	0.001607
18	(((((.(((((((((.....)))))))))..))).)).....	231	0.001540
19	((((..((((..(((.....)))..)))))...)))).....	225	0.001500
20	((....((((..(((.....)))..))))).....)).....	202	0.001347



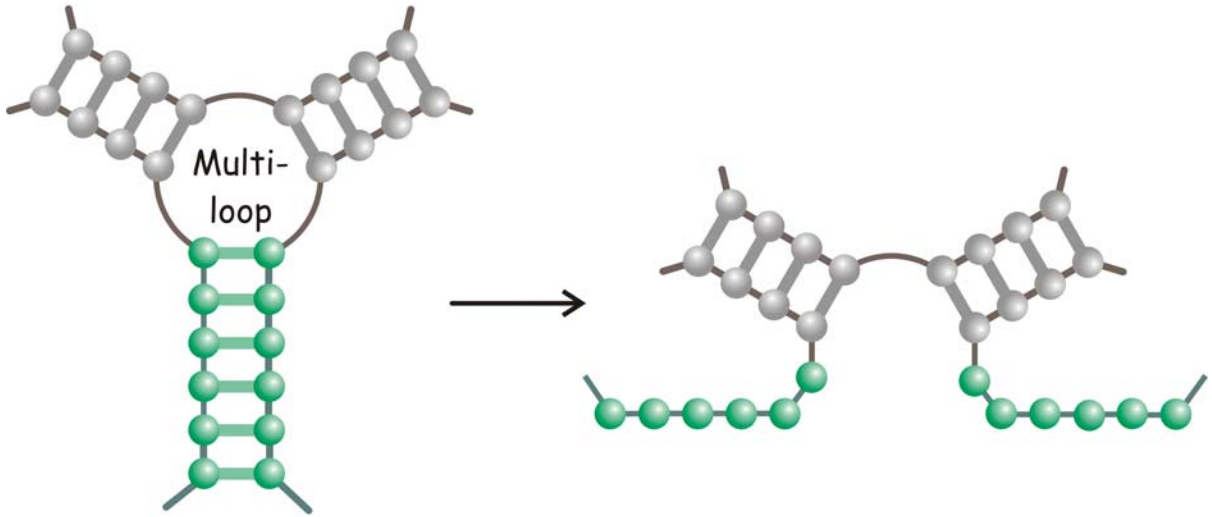
Shadow – Surrounding of an RNA structure in shape space:
AUGC alphabet, chain length n=50



Shortening of Stacks

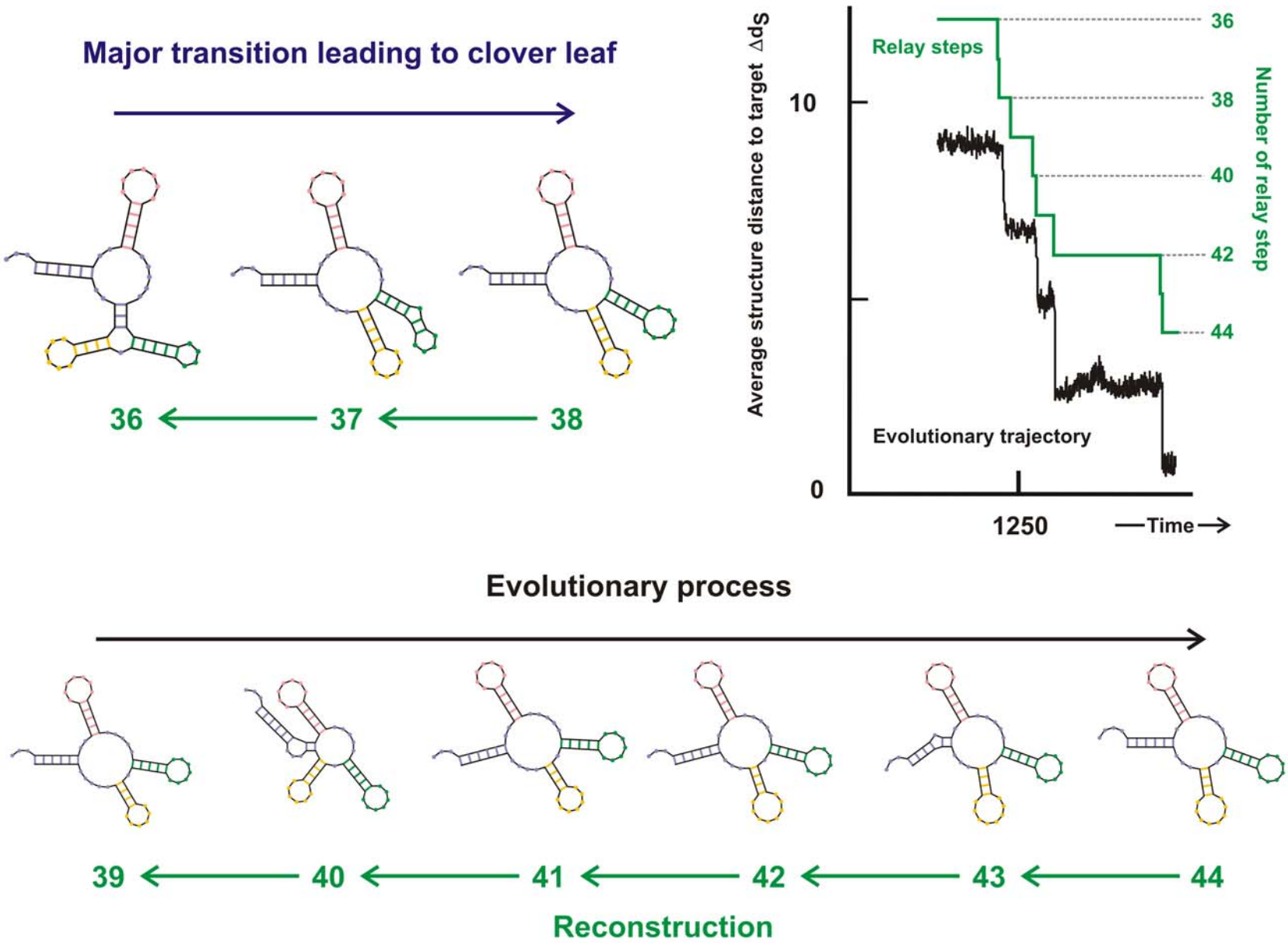


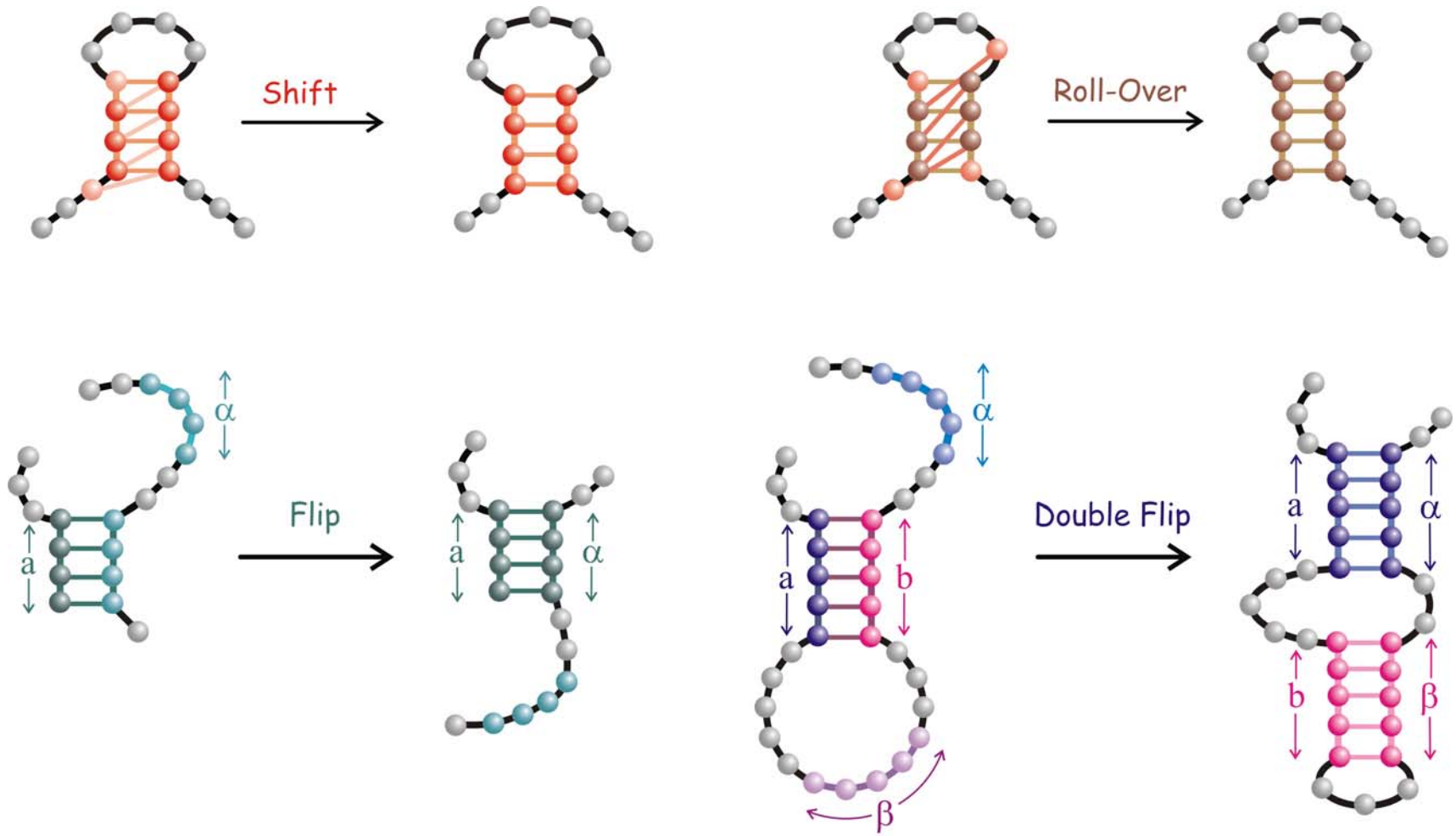
Elongation of Stacks



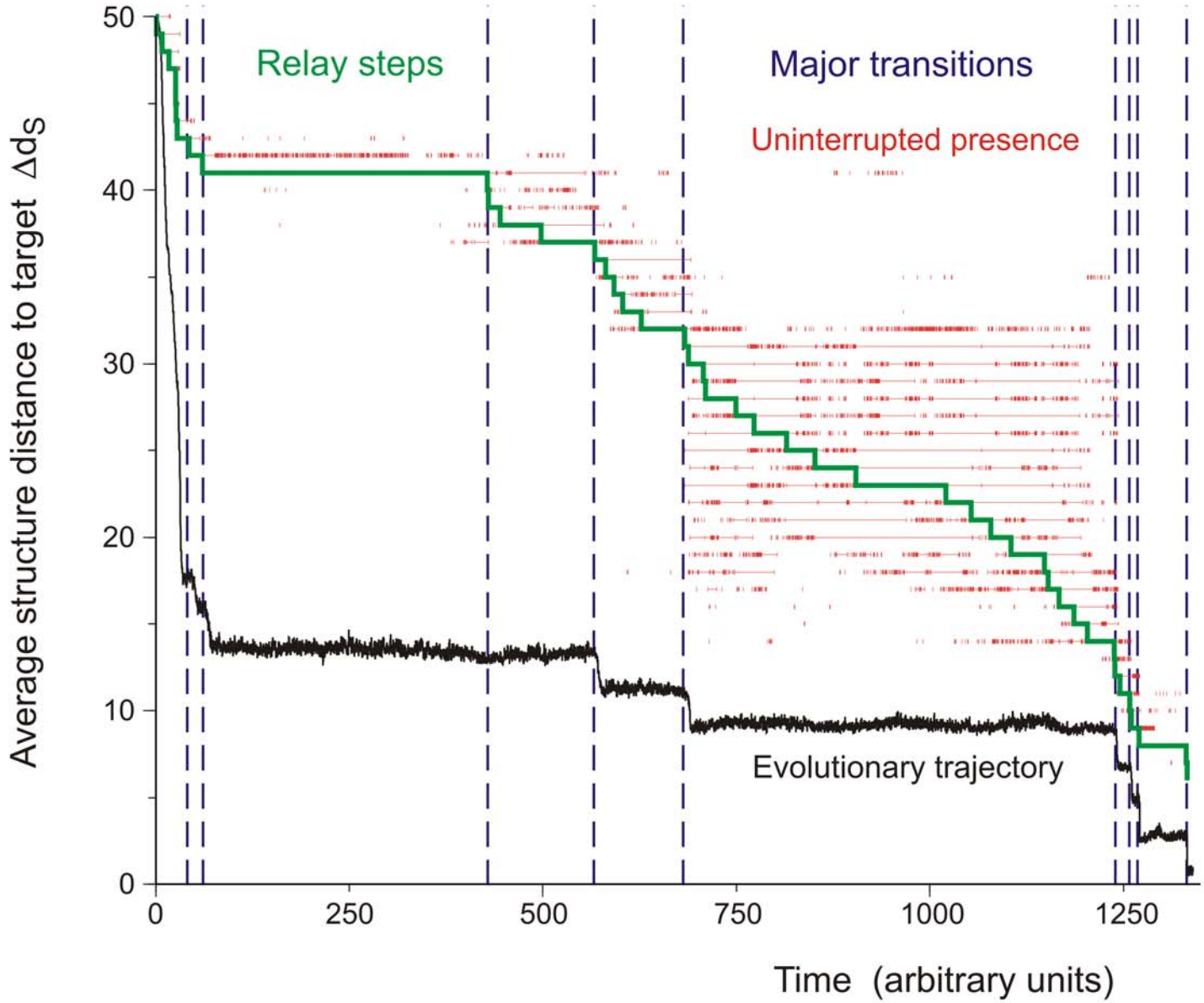
Opening of Constrained Stacks

Continuous Transitions in RNA Secondary Structures

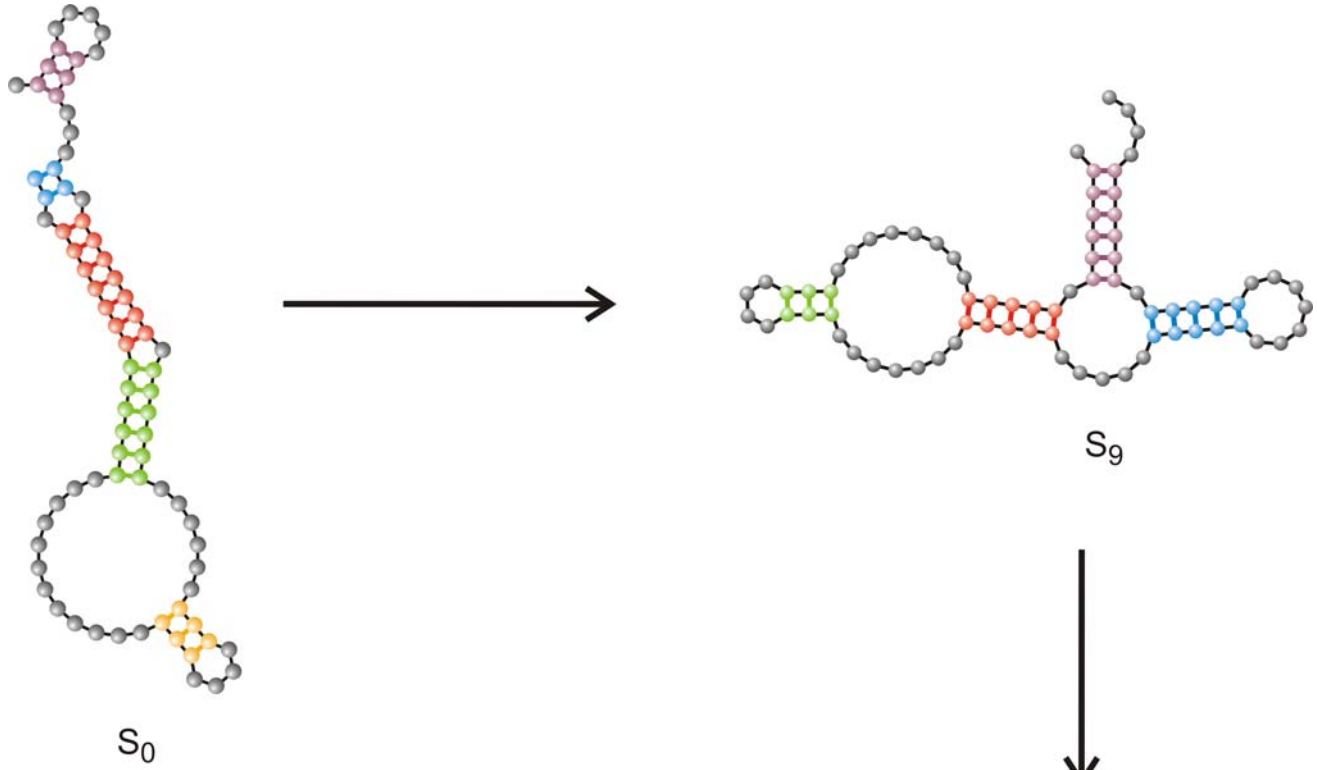

Major transition leading to clover leaf
36
37
38
Average structure distance to target ΔdS
10
0
Relay steps
36
38
40
42
44
Number of relay step
Evolutionary trajectory
1250
Time
Evolutionary process
39
40
41
42
43
44
Reconstruction



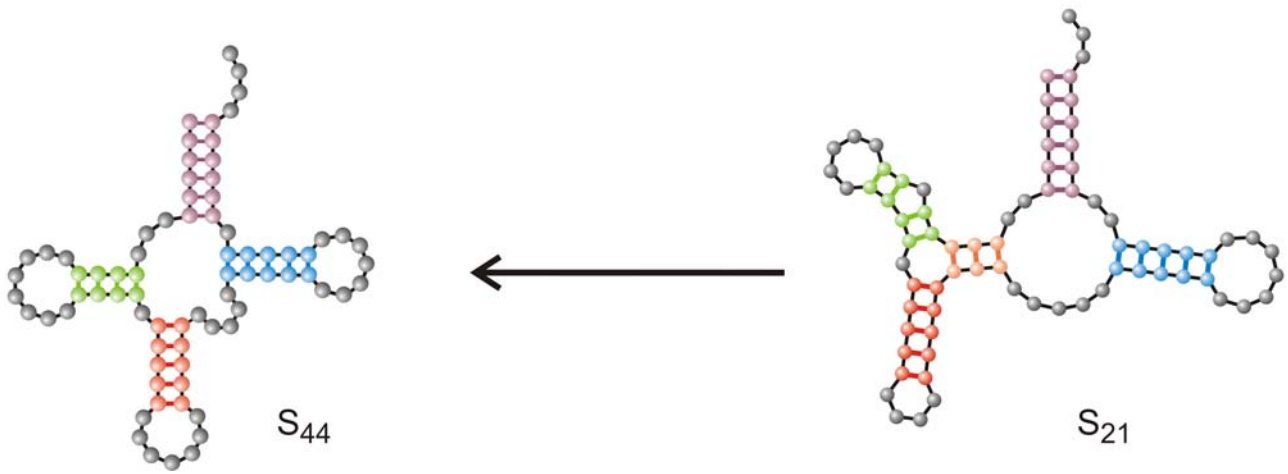
Discontinuous Transitions in RNA Secondary Structures


Relay steps
Major transitions
Uninterrupted presence
Evolutionary trajectory
Average structure distance to target Δds
Time (arbitrary units)
0
10
20
30
40
50
0
250
500
750
1000
1250

Randomly chosen
initial structure



Phenylalanyl-tRNA
as target structure

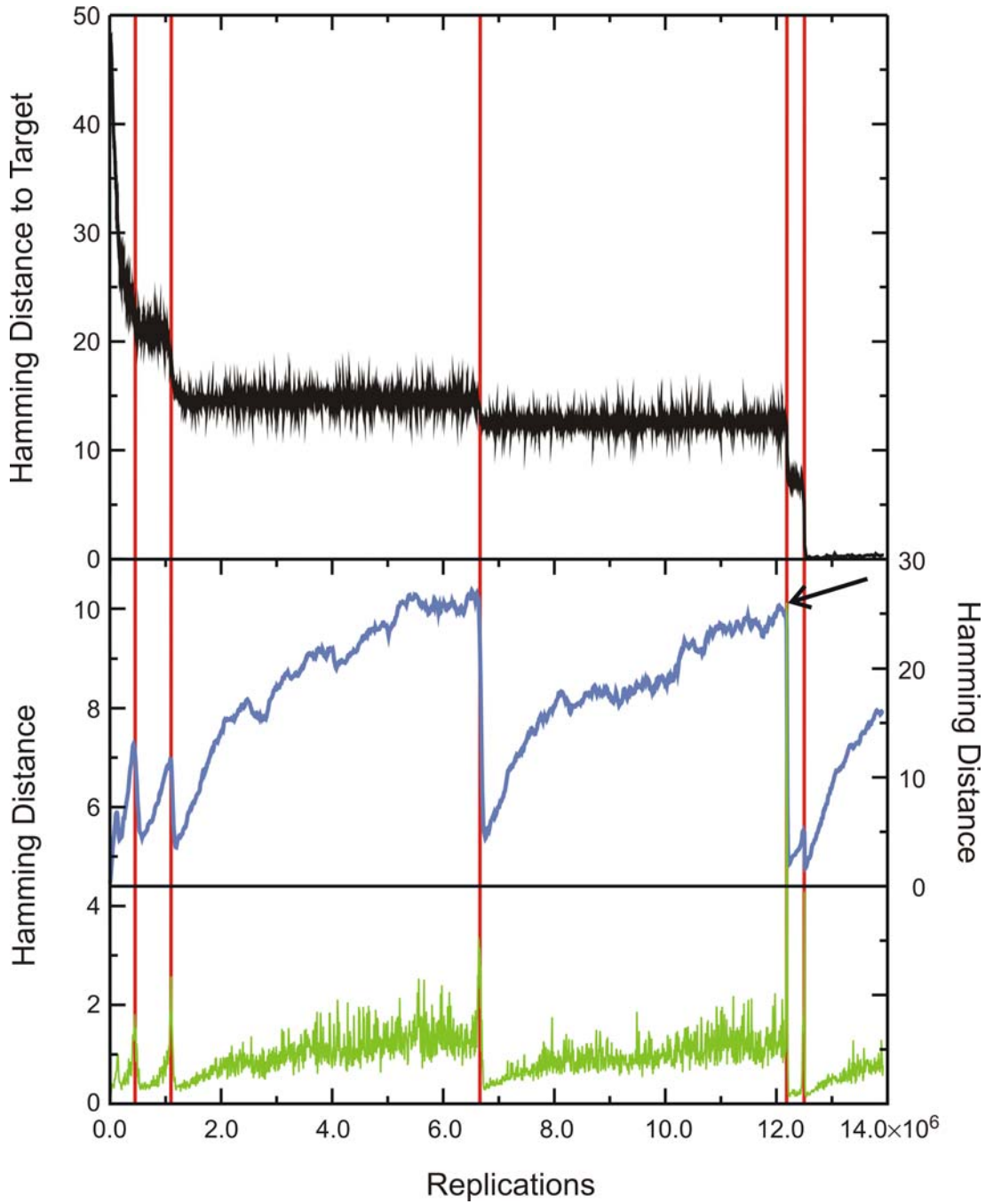


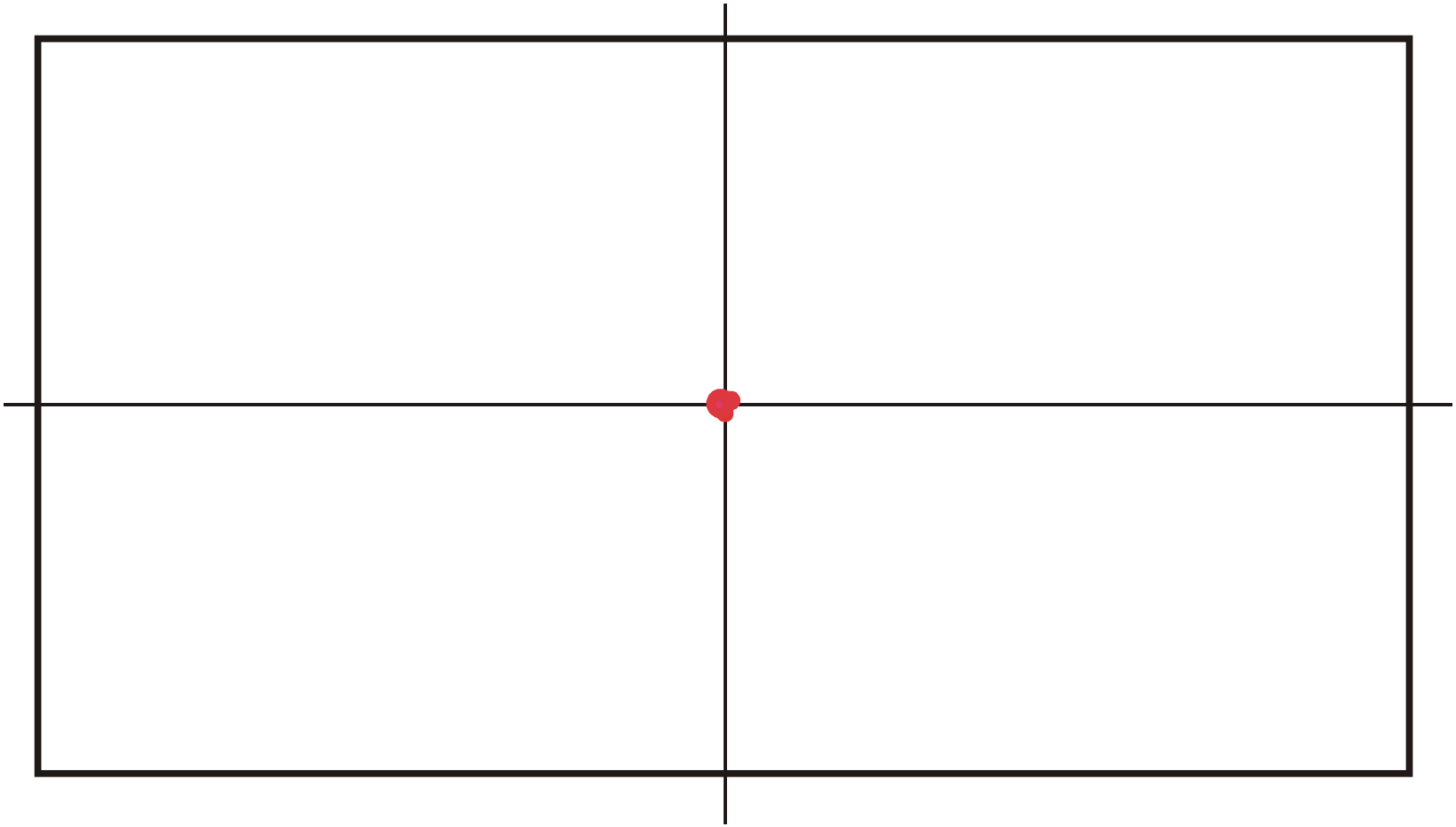
1. Reconstruction of evolutionary processes
2. Diffusion in sequence space and shape space
3. Continuous and discontinuous transitions
4. **Mechanism of RNA optimization**
5. Major transitions in evolution

Evolutionary trajectory

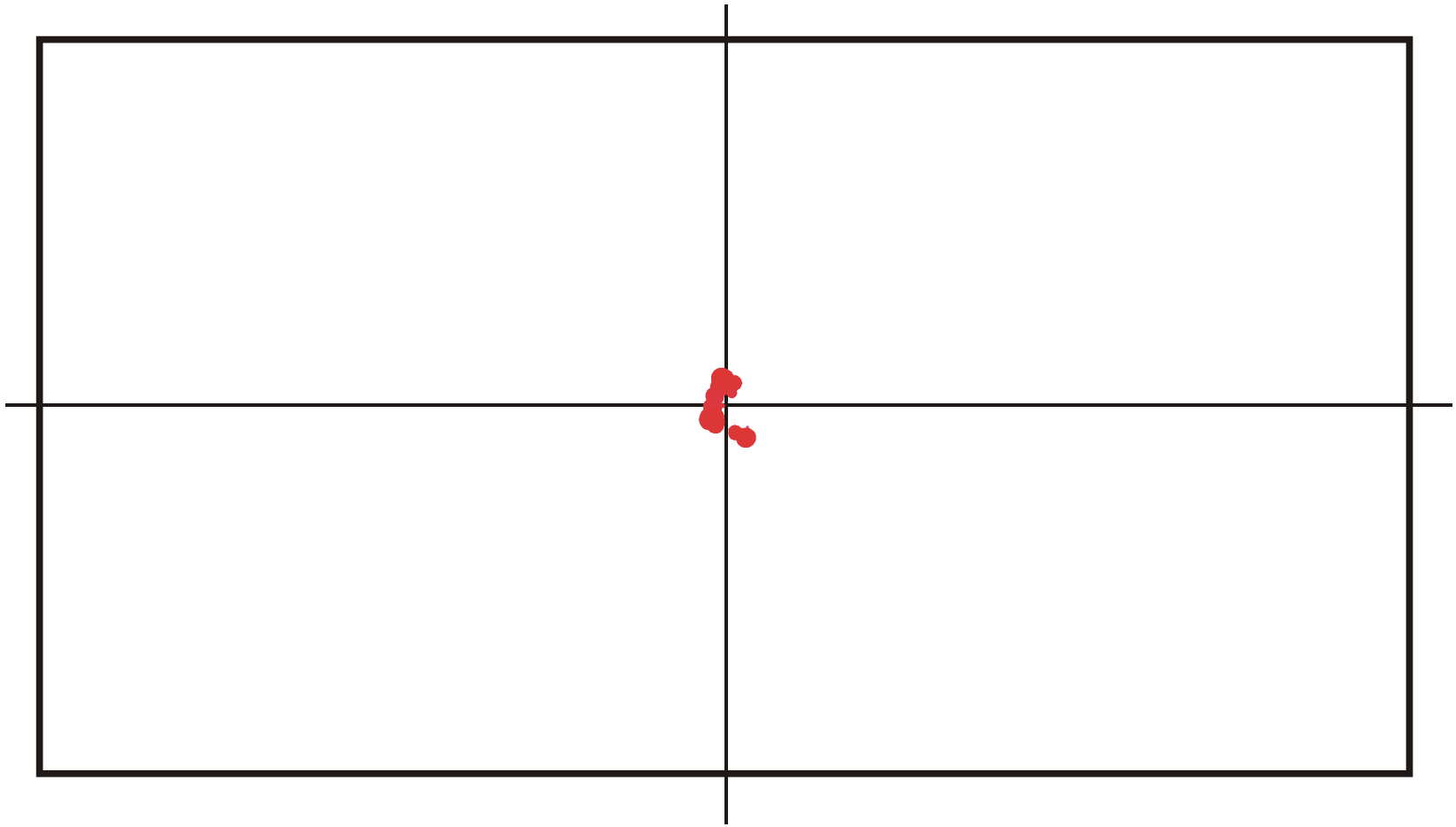
Spreading of the population on neutral networks

Drift of the population center in sequence space

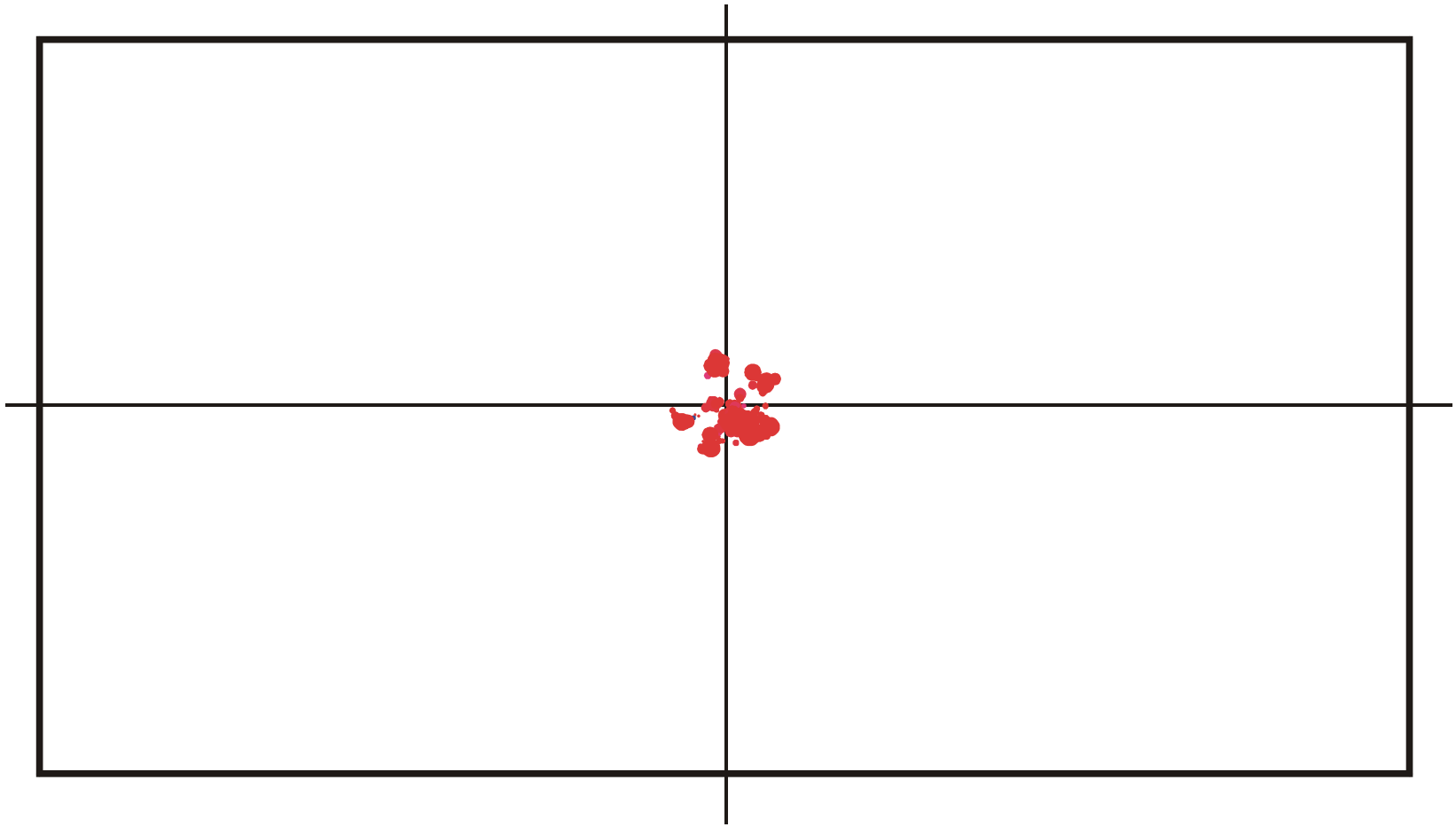




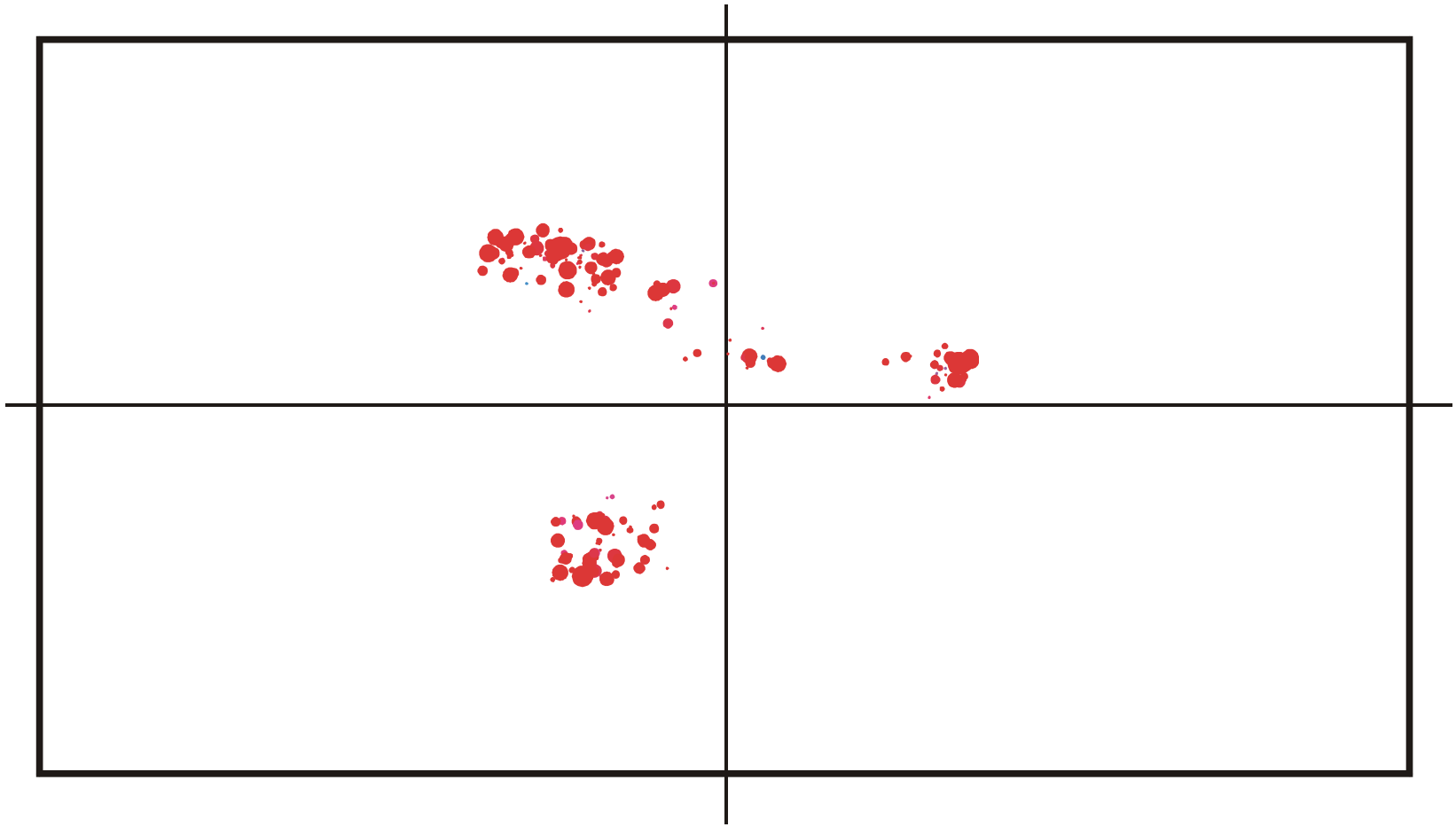
Spreading and evolution of a population on a neutral network: t = 150



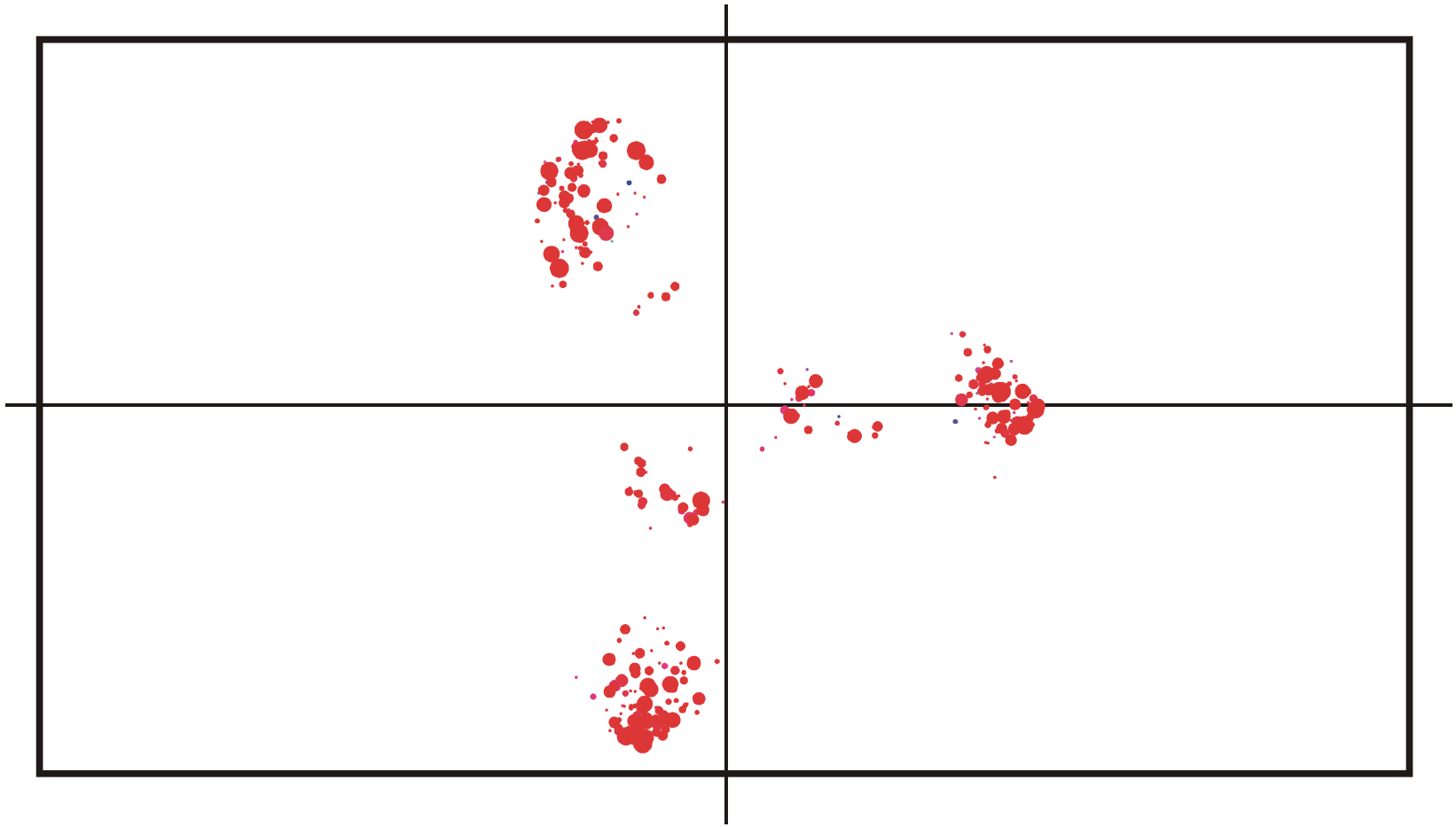
Spreading and evolution of a population on a neutral network : t = 170



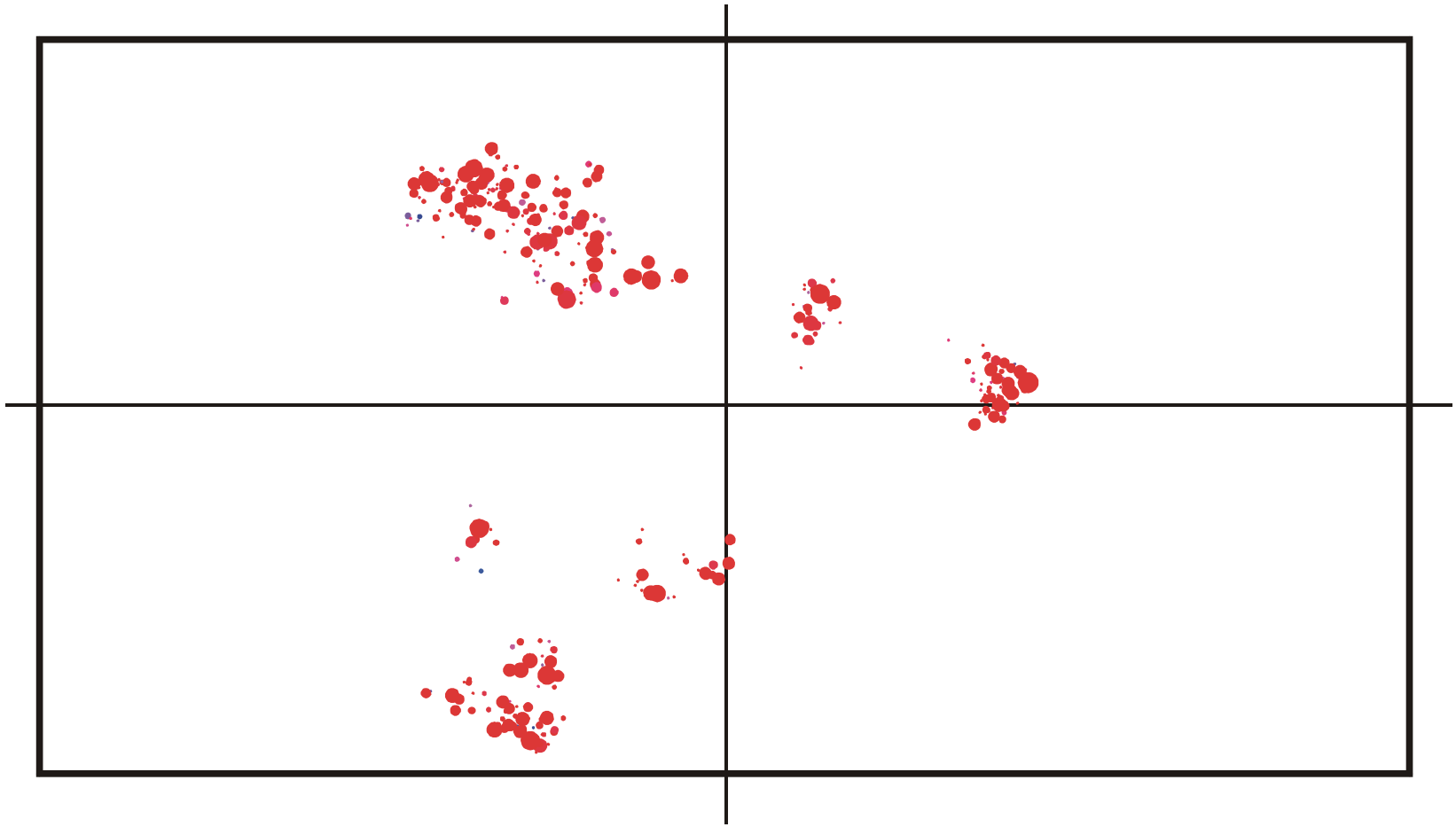
Spreading and evolution of a population on a neutral network : t = 200



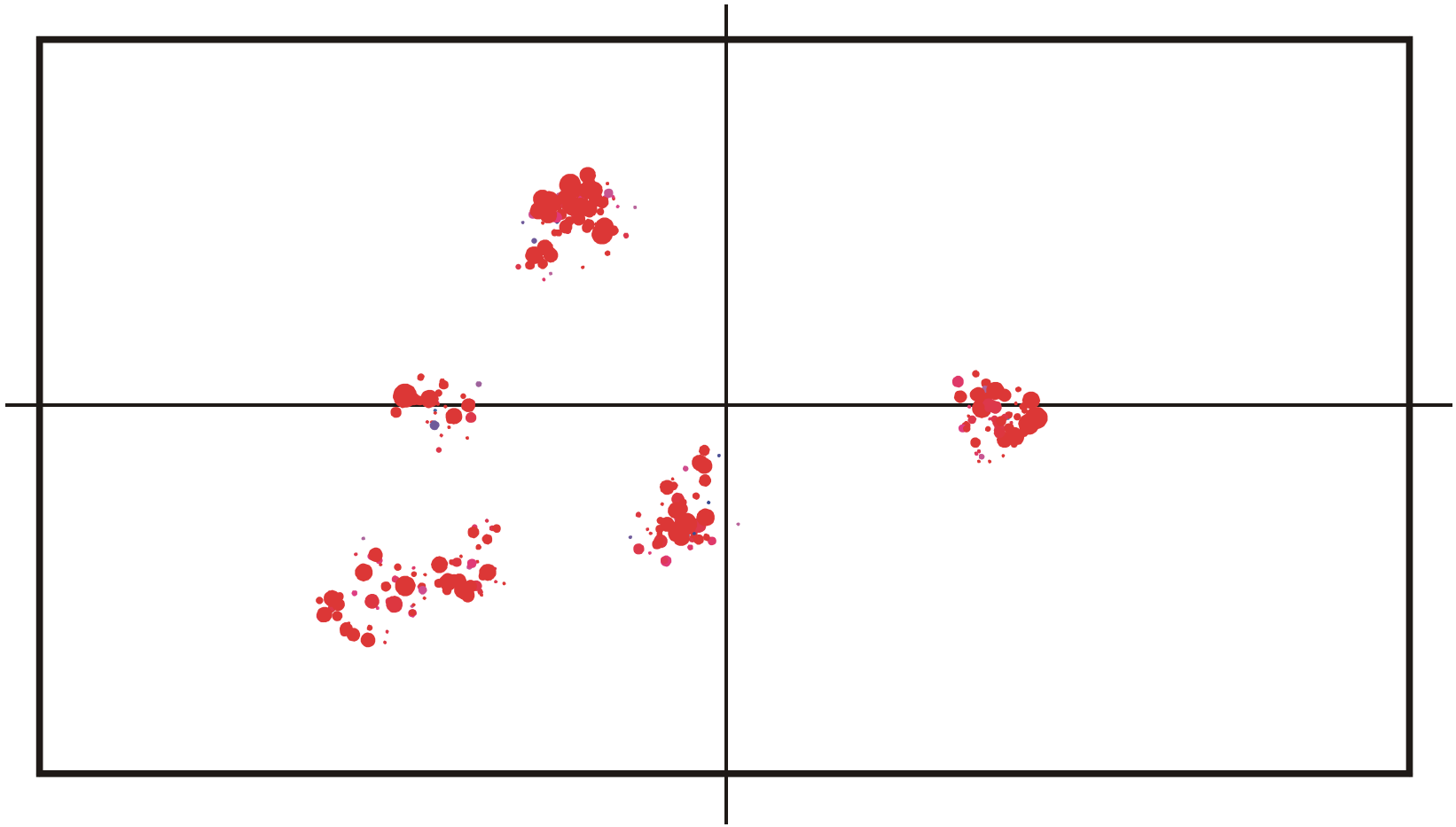
Spreading and evolution of a population on a neutral network : t = 350



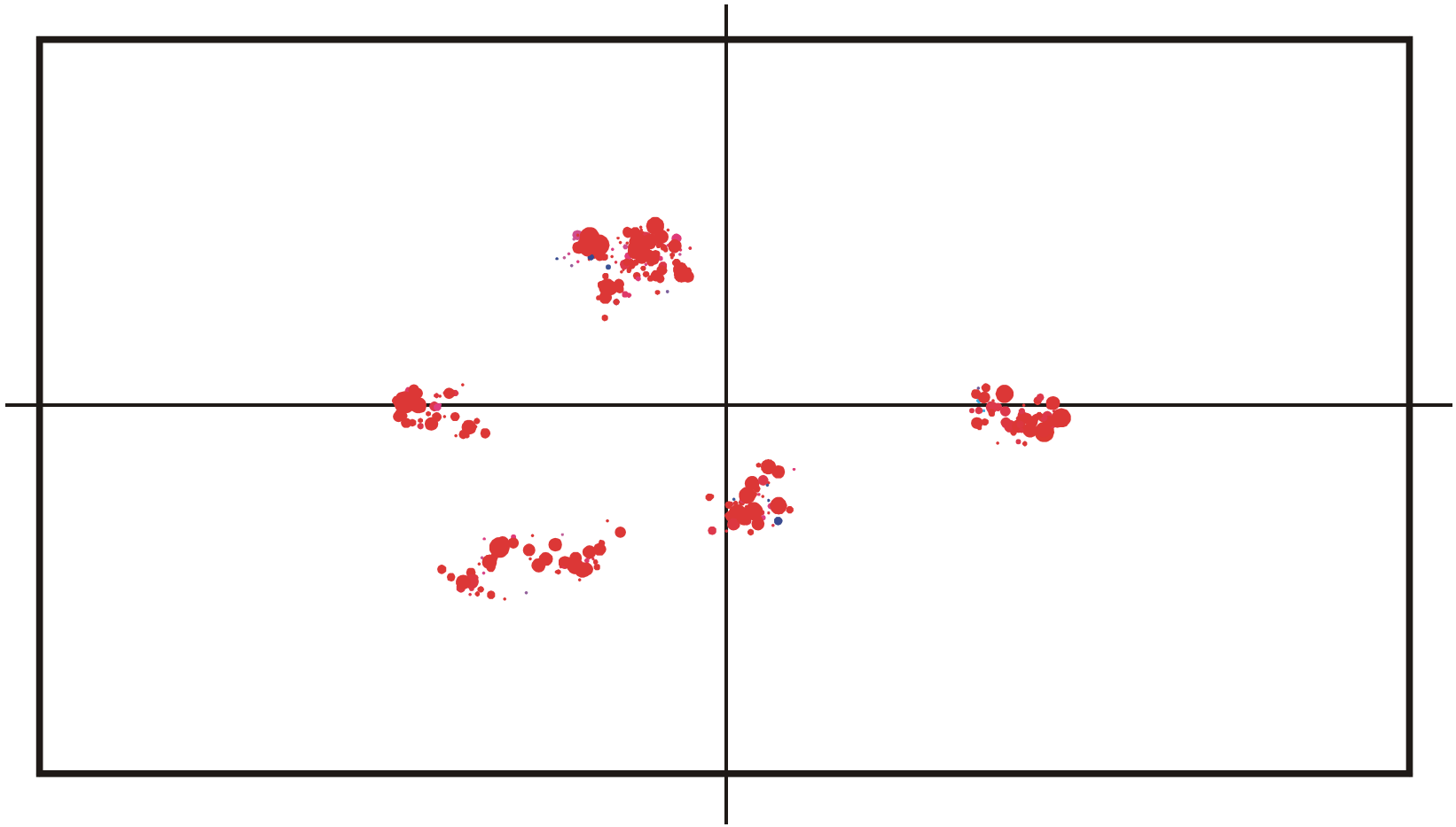
Spreading and evolution of a population on a neutral network : t = 500



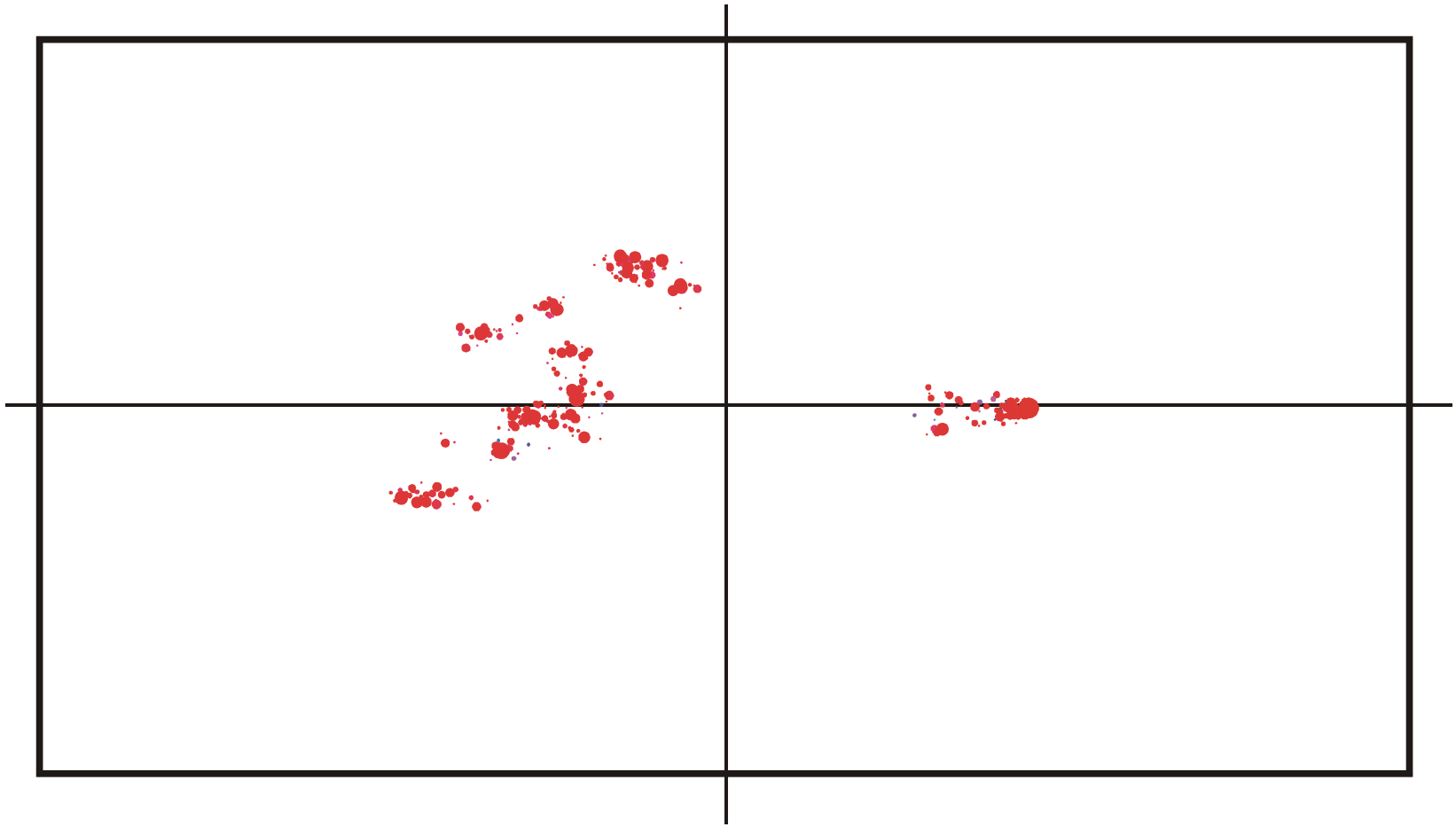
Spreading and evolution of a population on a neutral network : t = 650



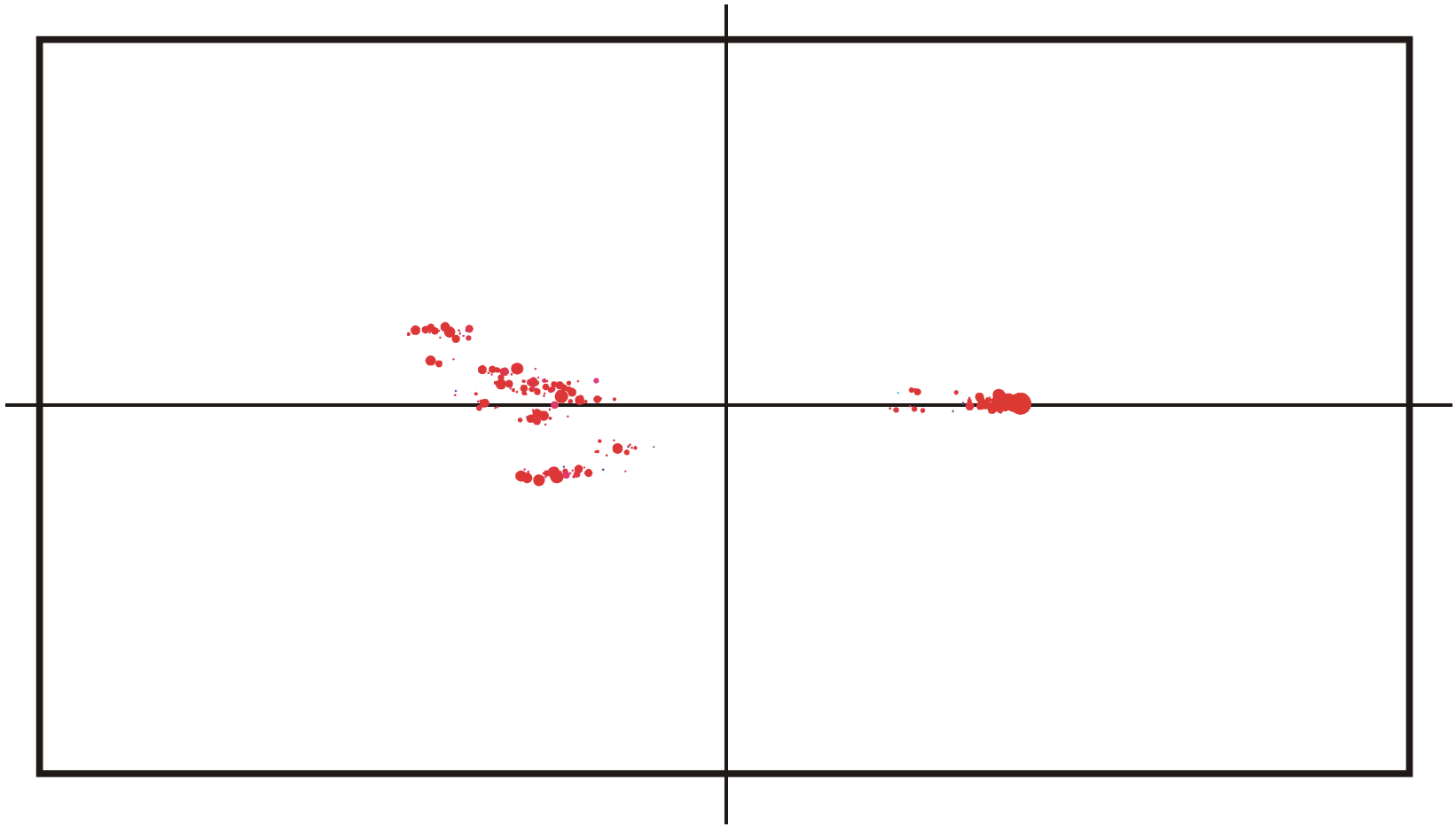
Spreading and evolution of a population on a neutral network : t = 820



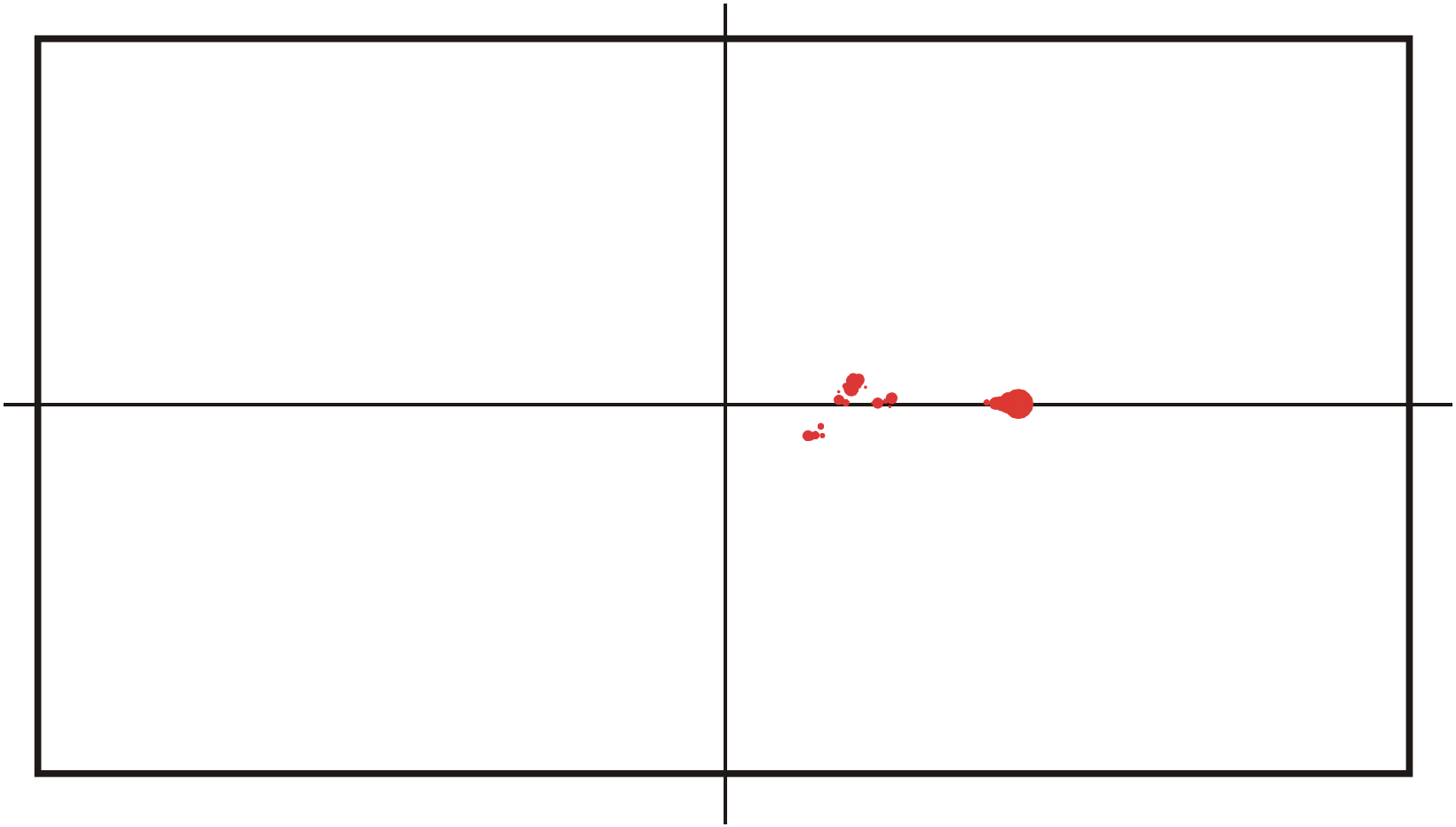
Spreading and evolution of a population on a neutral network : t = 825



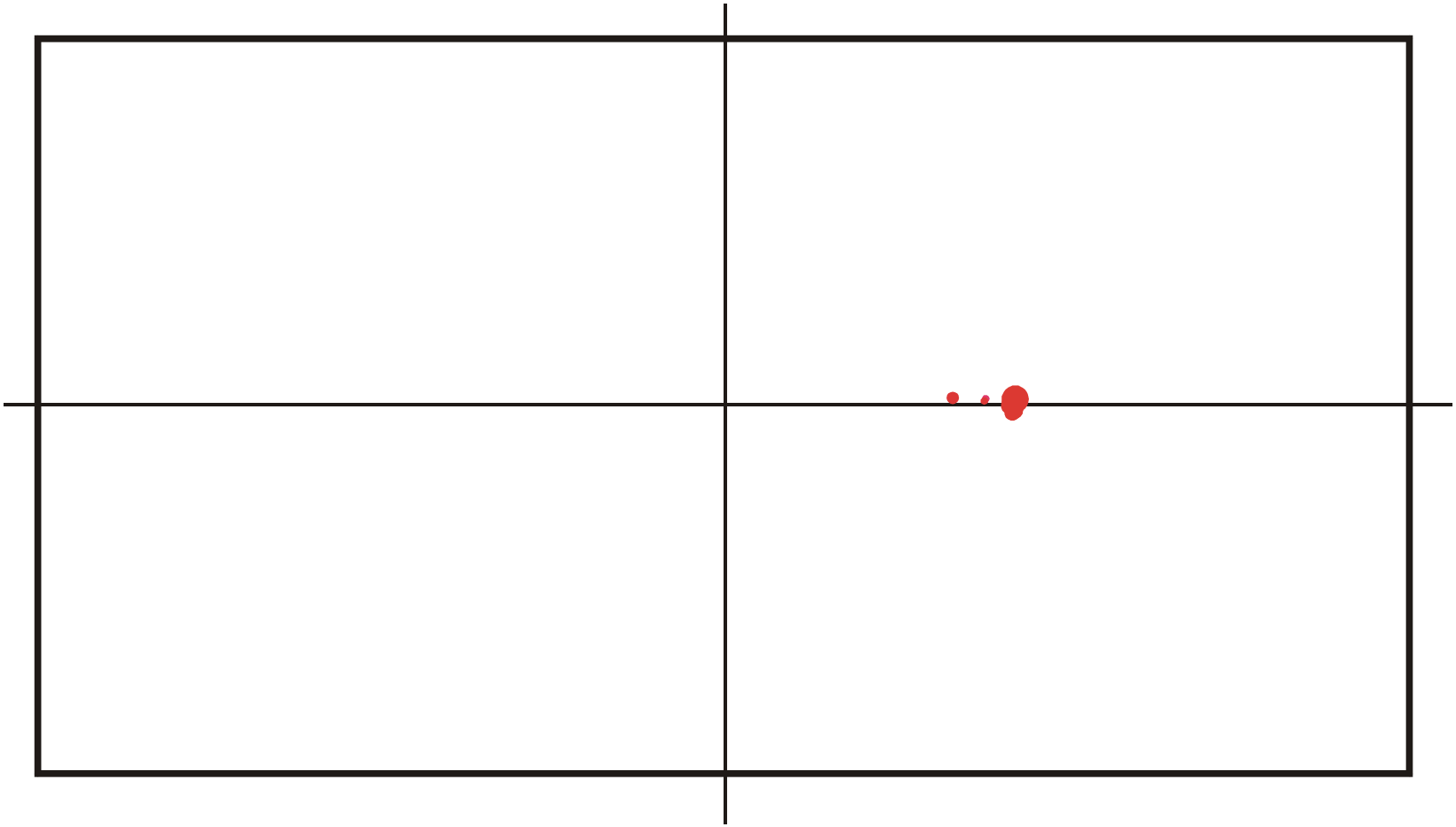
Spreading and evolution of a population on a neutral network : t = 830



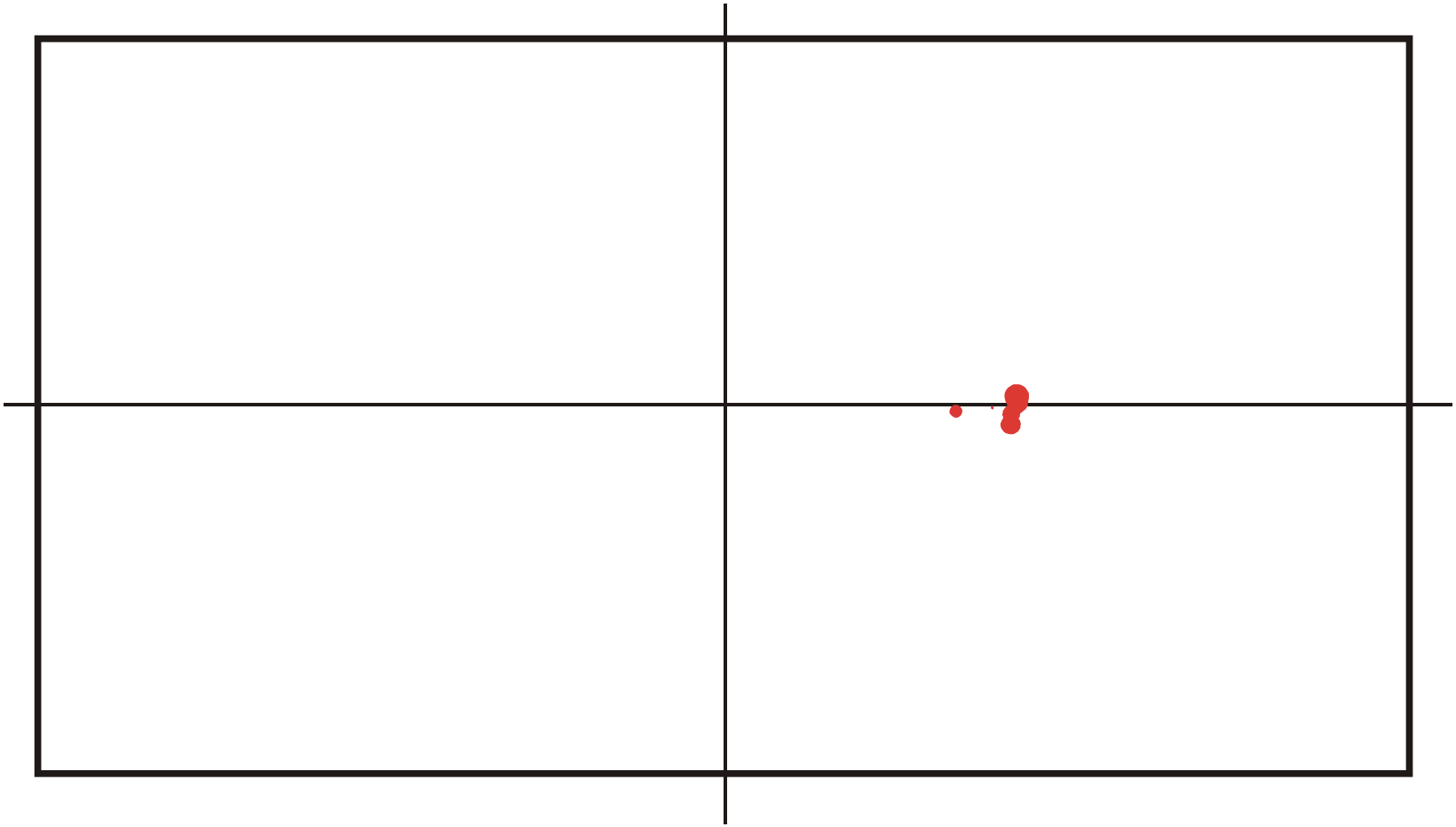
Spreading and evolution of a population on a neutral network : t = 835



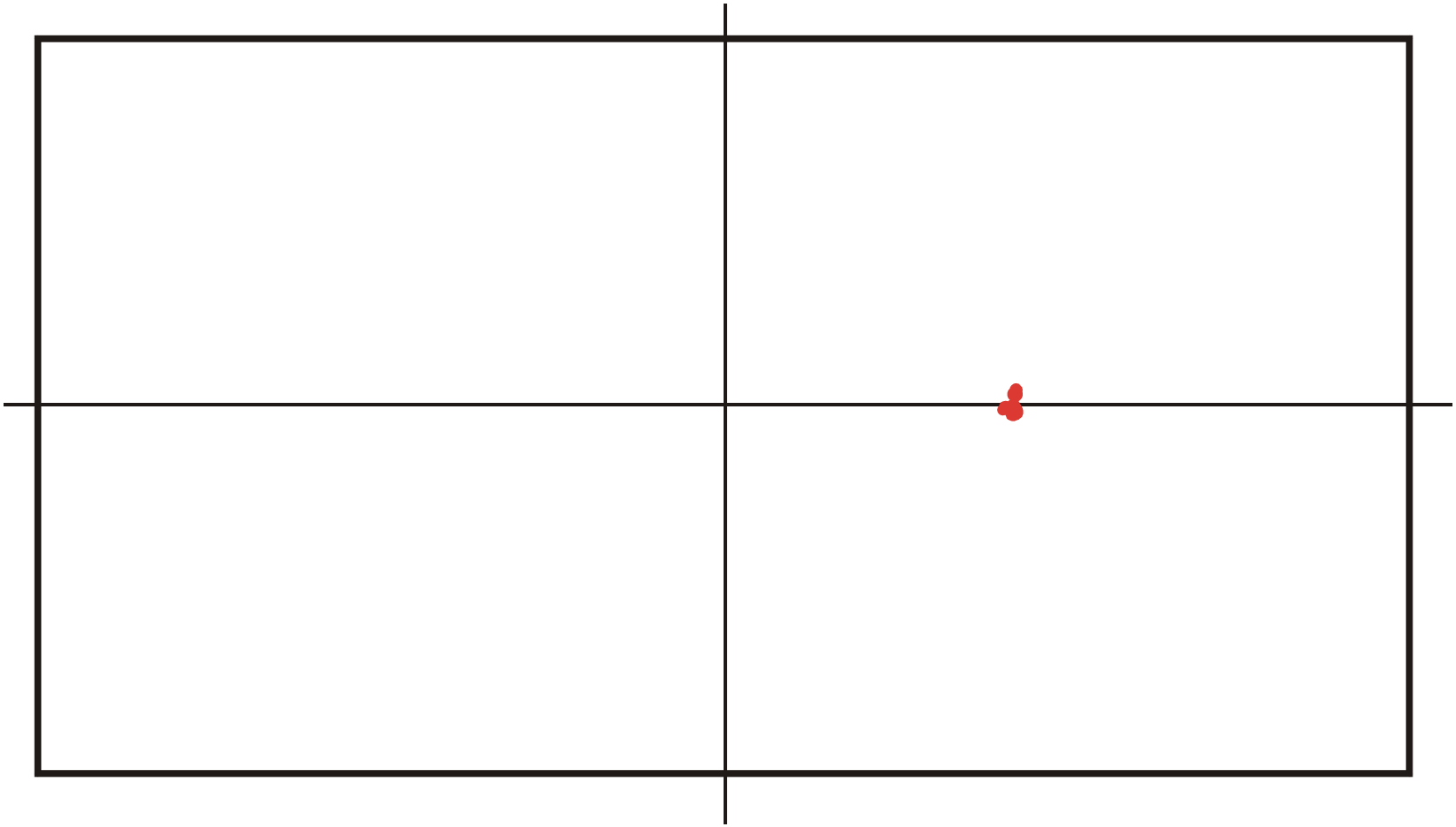
Spreading and evolution of a population on a neutral network : t = 840



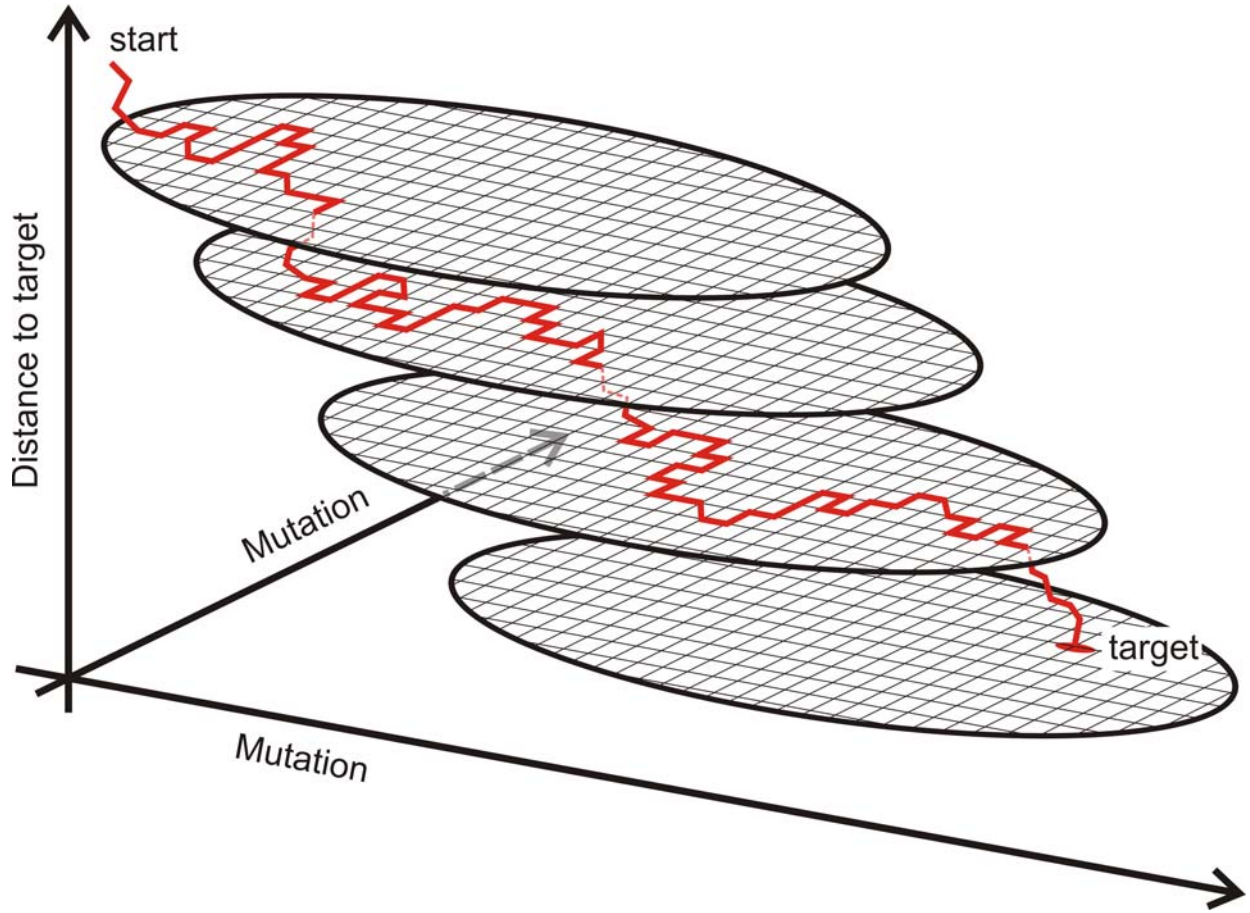
Spreading and evolution of a population on a neutral network : t = 845



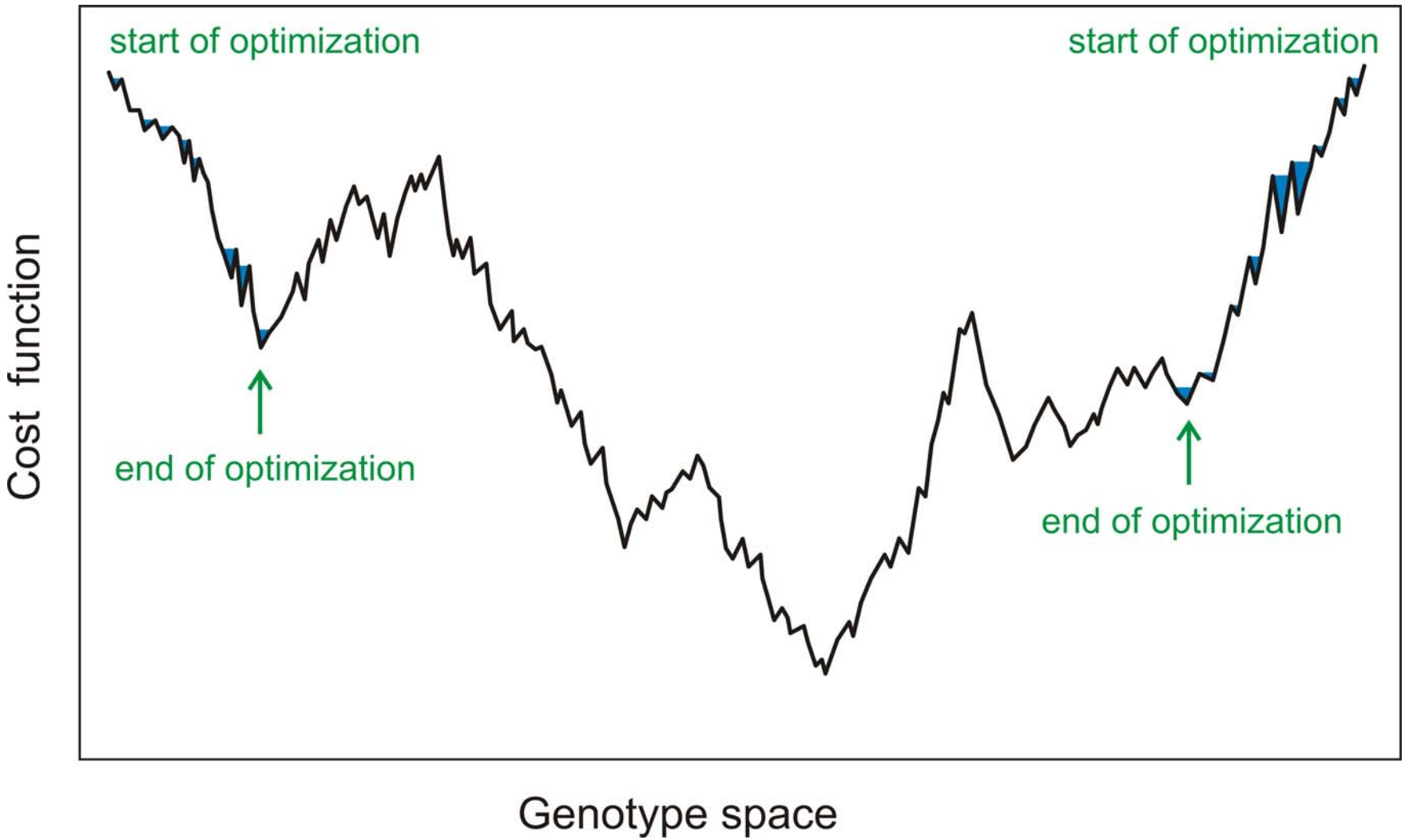
Spreading and evolution of a population on a neutral network : t = 850



Spreading and evolution of a population on a neutral network : t = 855



A sketch of optimization on neutral networks


start of optimization
start of optimization
end of optimization
end of optimization
Cost function
Genotype space

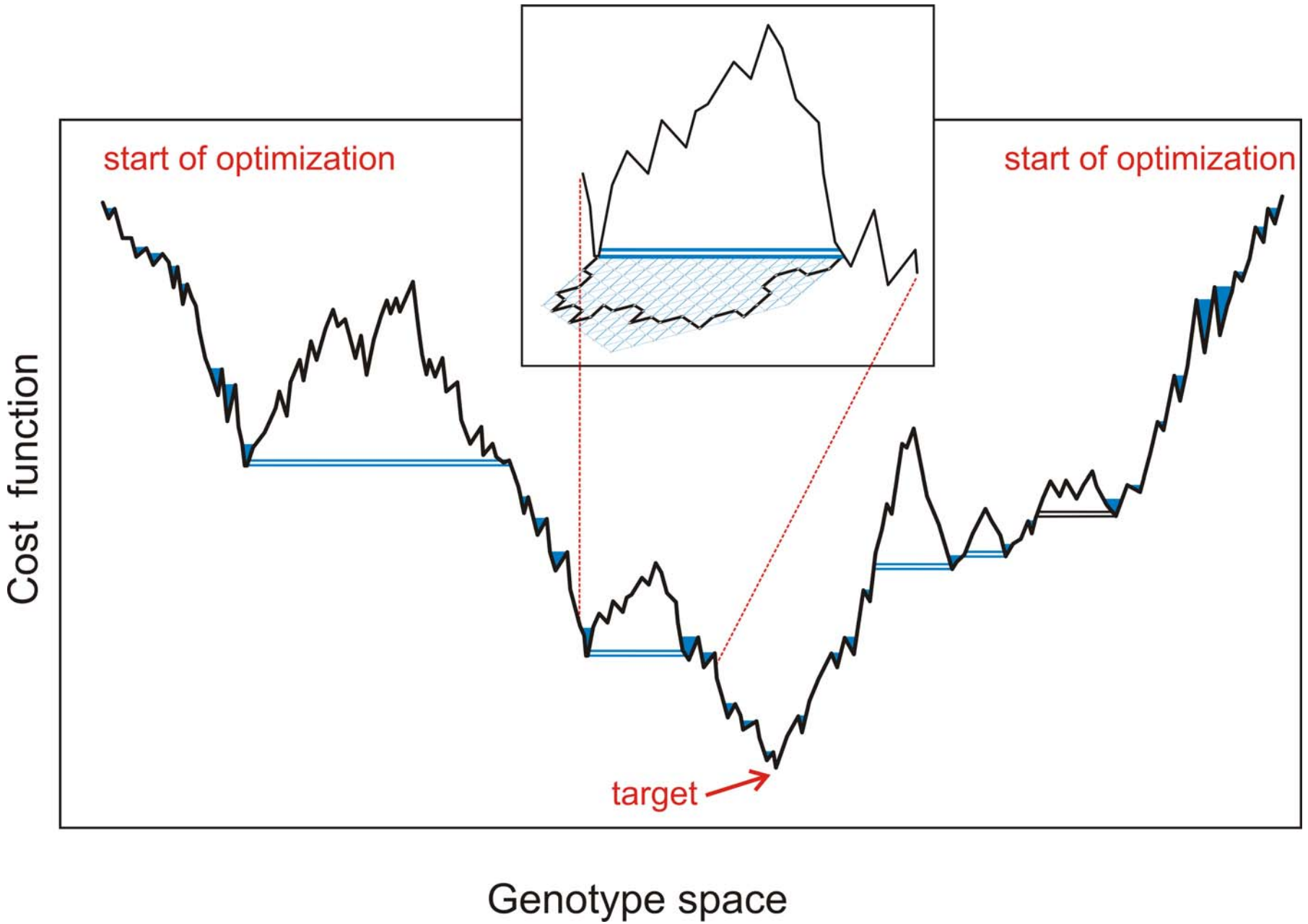

start of optimization
start of optimization
target
Cost function
Genotype space

Table 8. Statistics of the optimization trajectories. The table shows the results of sampled evolutionary trajectories leading from a random initial structure, S_I, to the structure of $tRNA^{phe}$, S_T, as the target[a]. Simulations were performed with an algorithm introduced by Gillespie [55–57]. The time unit is here undefined. A mutation rate of $p = 0.001$ per site and replication were used. The mean and standard deviation were calculated under the assumption of a log-normal distribution that fits well the data of the simulations.

Alphabet	Population size, N	Number of runs, n_R	Real time from start to target		Number of replications [10^7]	
			Mean value	σ	Mean value	σ
AUGC	1 000	120	900	+1380 −542	1.2	+3.1 −0.9
	2 000	120	530	+880 −330	1.4	+3.6 −1.0
	3 000	1199	400	+670 −250	1.6	+4.4 −1.2
	10 000	120	190	+230 −100	2.3	+5.3 −1.6
	30 000	63	110	+97 −52	3.6	+6.7 −2.3
	100 000	18	62	+50 −28	–	–
GC	1 000	46	5160	+15700 −3890	–	–
	3 000	278	1910	+5180 −1460	7.4	+35.8 −6.1
	10 000	40	560	+1620 −420	–	–

[a] The structures S_I and S_T were used in the optimization:

S_I: ((.((((((((((((.....((((....)))).....))))))).)))))))..))...(((.....)))

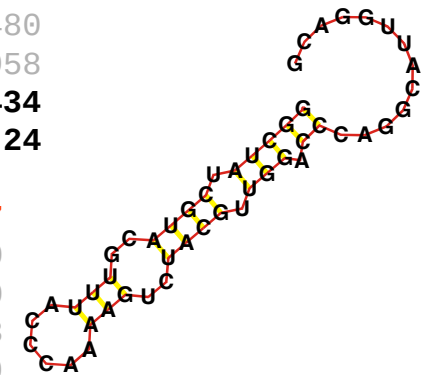
S_T: ((((((...((((........)))).((((.......)))))).....((((.......)))))).))))))....

Is the degree of neutrality in **GC** space much lower than in **AUGC** space ?

Statistics of RNA structure optimization: P. Schuster, Rep.Prog.Phys. 69:1419-1477, 2006

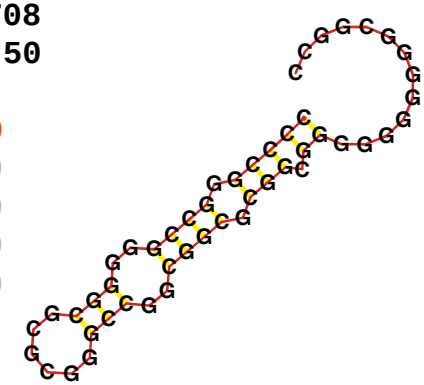
	Number	Mean Value	Variance	Std.Dev.
Total Hamming Distance:	150000	11.647973	23.140715	4.810480
Nonzero Hamming Distance:	99875	16.949991	30.757651	5.545958
Degree of Neutrality:	50125	**0.334167**	0.006961	**0.083434**
Number of Structures:	**1000**	**52.31**	85.30	**9.24**

1	(((((.((((..(((......))))..)))))..)))).....	**50125**	**0.334167**
2	..(((.((((..(((......))))..)))))..))).....	2856	0.019040
3	((((((((((..(((......))))..)))))))).)).....	2799	0.018660
4	(((((.((((..((((....))))..)))))..)))).....	2417	0.016113
5	(((((.((((.((((......))))..)))))..)))).....	2265	0.015100
6	(((((.(((((.(((......)))).)))))..)))).....	2233	0.014887



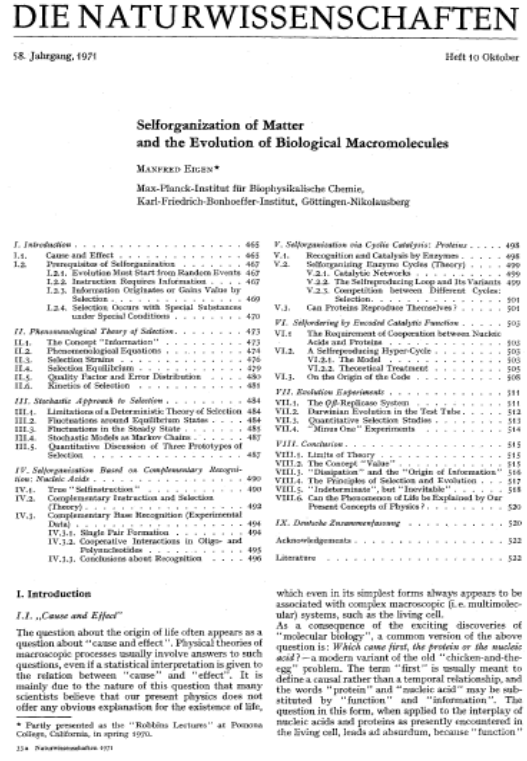
	Number	Mean Value	Variance	Std.Dev.
Total Hamming Distance:	50000	13.673580	10.795762	3.285691
Nonzero Hamming Distance:	45738	14.872054	10.821236	3.289565
Degree of Neutrality:	4262	**0.085240**	0.001824	**0.042708**
Number of Structures:	**1000**	**36.24**	6.27	**2.50**

1	(((((.((((..(((......))))..)))))..)))).....	**4262**	**0.085240**
2	((((((((((..(((......))))..)))))))).)).....	1940	0.038800
3	(((((.(((((.(((......)))).)))))..)))).....	1791	0.035820
4	(((((.((((.((((......))))..)))))..)))).....	1752	0.035040
5	(((((.((((..((((....))))..)))))..)))).....	1423	0.028460



Shadow – Surrounding of an RNA structure in shape space – **AUGC** and **GC** alphabet

1. Reconstruction of evolutionary processes
2. Diffusion in sequence space and shape space
3. Continuous and discontinuous transitions
4. Mechanism of RNA optimization
5. **Major transitions in evolution**



DIE NATURWISSENSCHAFTEN

58. Jahrgang, 1971 — Heft 10 Oktober

Selforganization of Matter and the Evolution of Biological Macromolecules

MANFRED EIGEN*

Max-Planck-Institut für Biophysikalische Chemie,
Karl-Friedrich-Bonhoeffer-Institut, Göttingen-Nikolausberg

* Partly presented as the "Robbins Lectures" at Pomona College, California, in spring 1970.

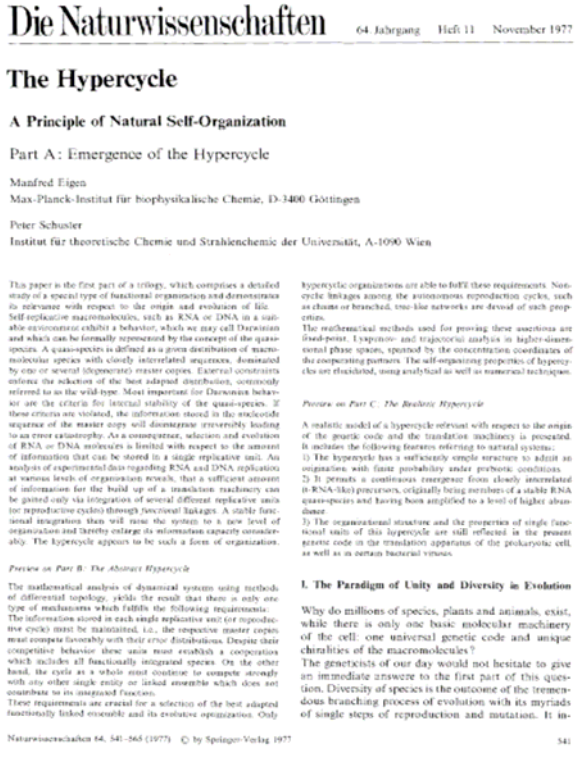
1. Introduction

1.1. „Cause and Effect"

The question about the origin of life often appears as a question about "cause and effect". Physical theories of macroscopic processes usually involve answers to such questions, even if a statistical interpretation is given to the relation between "cause" and "effect". It is mainly due to the nature of this question that many scientists believe that our present physics does not offer any obvious explanation for the existence of life, which even in its simplest forms always appears to be associated with complex macroscopic (i.e. multimolecular) systems, such as the living cell.

As a consequence of the exciting discoveries of "molecular biology", a common version of the above question is: *Which came first, the protein or the nucleic acid?* – a modern variant of the old "chicken-and-the-egg" problem. The term "first" is usually meant to define a causal rather than a temporal relationship, and the words "protein" and "nucleic acid" may be substituted by "function" and "information". The question in this form, when applied to the interplay of nucleic acids and proteins as presently encountered in the living cell, leads ad absurdum, because "function"

1971



Die Naturwissenschaften 64. Jahrgang Heft 11 November 1977

The Hypercycle

A Principle of Natural Self-Organization

Part A: Emergence of the Hypercycle

Manfred Eigen
Max-Planck-Institut für biophysikalische Chemie, D-3400 Göttingen

Peter Schuster
Institut für theoretische Chemie und Strahlenchemie der Universität, A-1090 Wien

This paper is the first part of a trilogy, which comprises a detailed study of a special type of functional organization and demonstrates its relevance with respect to the origin and evolution of life. Self-replicative macromolecules, such as RNA or DNA in a suitable environment exhibit a behavior, which we may call Darwinian and which can be formally represented by the concept of the quasi-species. A quasi-species is defined as a given distribution of macromolecular species with closely interrelated sequences, dominated by one or several (degenerate) master copies. External constraints enforce the selection of the best adapted distribution, commonly referred to as the wild-type. Most important for Darwinian behavior are the criteria for internal stability of the quasi-species. If these criteria are violated, the information stored in the nucleotide sequence of the master copy will deteriorate irreversibly leading to an error catastrophy. As a consequence, selection and evolution of RNA or DNA molecules is limited with respect to the amount of information that can be stored in a single replicative unit. An analysis of experimental data regarding RNA and DNA replication at various levels of organization reveals, that a sufficient amount of information for the build up of a translation machinery can be gained only via integration of several different replicative units (or reproductive cycles) through functional linkages. A stable functional integration then will raise the system to a new level of organization and thereby enlarge its information capacity considerably. The hypercycle appears to be such a form of organization.

Preview on Part B: The Abstract Hypercycle

The mathematical analysis of dynamical systems using methods of differential topology, yields the result that there is only one type of mechanisms which fulfills the following requirements: The information stored in each single replicative unit (or reproductive cycle) must be maintained, i.e., the respective master copies must compete favorably with their error distributions. Despite their competitive behavior these units must establish a cooperation which includes all functionally integrated species. On the other hand, the cycle as a whole must continue to compete strongly with any other single entity or linked ensemble which does not contribute to its integrated function.
These requirements are crucial for a selection of the best adapted functionally linked ensemble and its evolutive optimization. Only hypercyclic organizations are able to fulfil these requirements. Non-cycle linkages among the autonomous reproduction cycles, such as chains or branched, tree-like networks are devoid of such properties.
The mathematical methods used for proving these assertions are fixed-point, Lyapunov- and trajectorial analysis in higher-dimensional phase spaces, spanned by the concentration coordinates of the cooperating partners. The self-organizing properties of hypercycles are elucidated, using analytical as well as numerical techniques.

Preview on Part C: The Realistic Hypercycle

A realistic model of a hypercycle relevant with respect to the origin of the genetic code and the translation machinery is presented. It includes the following features referring to natural systems:
1) The hypercycle has a sufficiently simple structure to admit an origination with finite probability under prebiotic conditions.
2) It permits a continuous emergence from closely interrelated (t-RNA-like) precursors, originally being members of a stable RNA quasi-species and having been amplified to a level of higher abundance.
3) The organizational structure and the properties of single functional units of this hypercycle are still reflected in the present genetic code in the translation apparatus of the prokaryotic cell, as well as in certain bacterial viruses.

I. The Paradigm of Unity and Diversity in Evolution

Why do millions of species, plants and animals, exist, while there is only one basic molecular machinery of the cell: one universal genetic code and unique chiralities of the macromolecules?
The geneticists of our day would not hesitate to give an immediate answere to the first part of this question. Diversity of species is the outcome of the tremendous branching process of evolution with its myriads of single steps of reproduction and mutation. It in-

1977

Molecular Quasi-Species†

Manfred Eigen,* John McCaskill,
Max Planck Institut für biophysikalische Chemie, Am Fassberg, D 3400 Göttingen-Nikolausberg, BRD
and Peter Schuster*
Institut für theoretische Chemie und Strahlenchemie, der Universität Wien, Währinger Strasse 17, A-1090 Wien, Austria (Received: June 9, 1988)

The molecular quasi-species model describes the physicochemical organization of monomers into an ensemble of heteropolymers with combinatorial complexity by ongoing template polymerization. Polynucleotides belong to the simplest class of such molecules. The quasi-species itself represents the stationary distribution of macromolecular sequences maintained by chemical reactions effecting error-prone replication and by transport processes. It is obtained deterministically, by mass-action kinetics, as the dominant eigenvalue of a *value* matrix, **W**, which is derived directly from chemical rate coefficients, but it also exhibits stochastic features, being composed to a significant fraction of unique individual macromolecular sequences. The quasi-species model demonstrates how macromolecular information originates through specific nonequilibrium autocatalytic reactions and thus forms a bridge between reaction kinetics and molecular evolution. Selection and evolutionary optimization appear as new features in physical chemistry. Concentration bias in the production of mutants is a new concept in population genetics, relevant to frequently mutating populations, which is shown to greatly enhance the optimization properties. The present theory relates to asexually replicating ensembles, but this restriction is not essential. A sharp transition is exhibited between a drifting population of essentially random macromolecular sequences and a localized population of close relatives. This transition at a threshold error rate was found to depend on sequence lengths, distributions of selective values, and population sizes. It has been determined generically for complex landscapes and for special cases, and, it was shown to persist generically in the presence of nearly neutral mutants. Replication dynamics has much in common with the equilibrium statistics of complex spin systems: the error threshold is equivalent to a magnetic order-disorder transition. A rational function of the replication accuracy plays the role of temperature. Experimental data obtained from test-tube evolution of polynucleotides and from studies of natural virus populations support the quasi-species model. The error threshold seems to set a limit to the genome lengths of several classes of RNA viruses. In addition, the results are relevant even in eucaryotes where they contribute to the exon-intron debate.

1. Molecular Selection

Our knowledge of physical and chemical systems is, in a final analysis, based on models derived from repeatable experiments. While none of the classic and rather besieged list of properties rounded up to support the intuition of a distinction between the living and nonliving—metabolism, self-reproduction, irritability, and adaptability, for example—intrinsically limit the application of the scientific method, a determining role by unique or *individual* entities comes into conflict with the requirement of repeatability. Combinatorial variety, such as that in heteropolymers based on even very small numbers of different bases, even just two, readily provides numbers of different entities so enormous that neither consecutive nor parallel physical realization is possible. The physical chemistry of finite systems of such macromolecules must deal with both known regularities and the advent of unique copolymeric sequences. Normally this would present no difficulty in a statistical mechanical analysis of typical behavior, where rare events play no significant role, but with autocatalytic polymerization processes even unique single molecules may be amplified to determine the fate of the entire system. Potentially creative, self-organizing around unique events, the dynamics of this simplest living chemical system is invested with regularities that both allow and limit efficient adaptation. The quasi-species model is a study of these regularities.

The fundamental regularity in living organisms that has invited explanation is adaptation. Why are organisms so well fitted to their environments? At a more chemical level, why are enzymes optimal catalysts? Darwin's theory of natural selection has provided biologists with a framework for the answer to this question. The present model is constructed along Darwinian lines but in terms of specific macromolecules, chemical reactions, and physical processes that make the notion of survival of the fittest precise. Not only does the model give an understanding of the physical limitations of adaptation, but also it provides new insight into the role of chance in the process. For an understanding of the structure of this minimal chemical model it is first necessary to recall the conceptual basis of Darwin's theory.

Darwin recognized that new inheritable adaptive properties were not induced by the environment but arose independently in the production of offspring. Lasting adaptive changes in a population could only come about by natural selection of the heritable traits or *genotype* based on the full characteristics or *phenotype* relevant for producing offspring. A process of chance, i.e., uncorrelated with the developed phenotype, controls changes in the genotype from one generation to the next and generates the diversity necessary for selection. Three factors have probably prevented chemists from gaining a clear insight into these phenomena in the past, despite the discovery of the polymeric nature of the genotype (DNA): the complexity of a minimum replication phenotype, the problem of dealing with a huge number of variants, and the nonequilibrium nature of these ongoing processes.

The formulation of a tractable chemical model based on Darwin's principle may be understood in several steps:

1. The major constituents of the system have to be inherently self-reproductive. Only two classes of molecules are presently

† This is an abridged account of the quasi-species theory that has been submitted in comprehensive form to *Advances in Chemical Physics*.[1]

(1) Eigen, M.; McCaskill, J. S.; Schuster, P. *Adv. Chem. Phys.*, in press.

0022-3654/88/2092-6881$01.50/0

1988

Chemical kinetics of molecular evolution

J Mol Evol (1982) 19:47–61

Journal of
Molecular Evolution

Stages of Emerging Life – Five Principles of Early Organization*

M. Eigen[1] and P. Schuster[2]

1 Max Planck Institut für Biophysikalische Chemie D–34 Göttingen, Federal Republic of Germany
2 Institut für Theoretische Chemie und Strahlenchemie der Universität Wien, A 1090 Wien, Austria

Summary. Five principles underly the evolution of the genetic language: formation of stereoregular heteropolymers, selection through self-replication, evolution of quasispecies towards optimal structures, regulated cooperation between competitors through catalytic hypercycles and evaluation of translation products through compartmentalization. These principles are formulated and illustrated by means of experimental results

abundant molecules on the very early Earth became clear. Most of the natural amino acids were among them. The building blocks of polynucleotides – the four bases, ribose and phosphate – were available too under prebiotic conditions. An enormous amount of low molecular weight organic material and energy-rich compounds was synthesized in those early days. Material was provided from steadily refilling pools for the formation of polymers, among them polypeptides and polynucleotides.

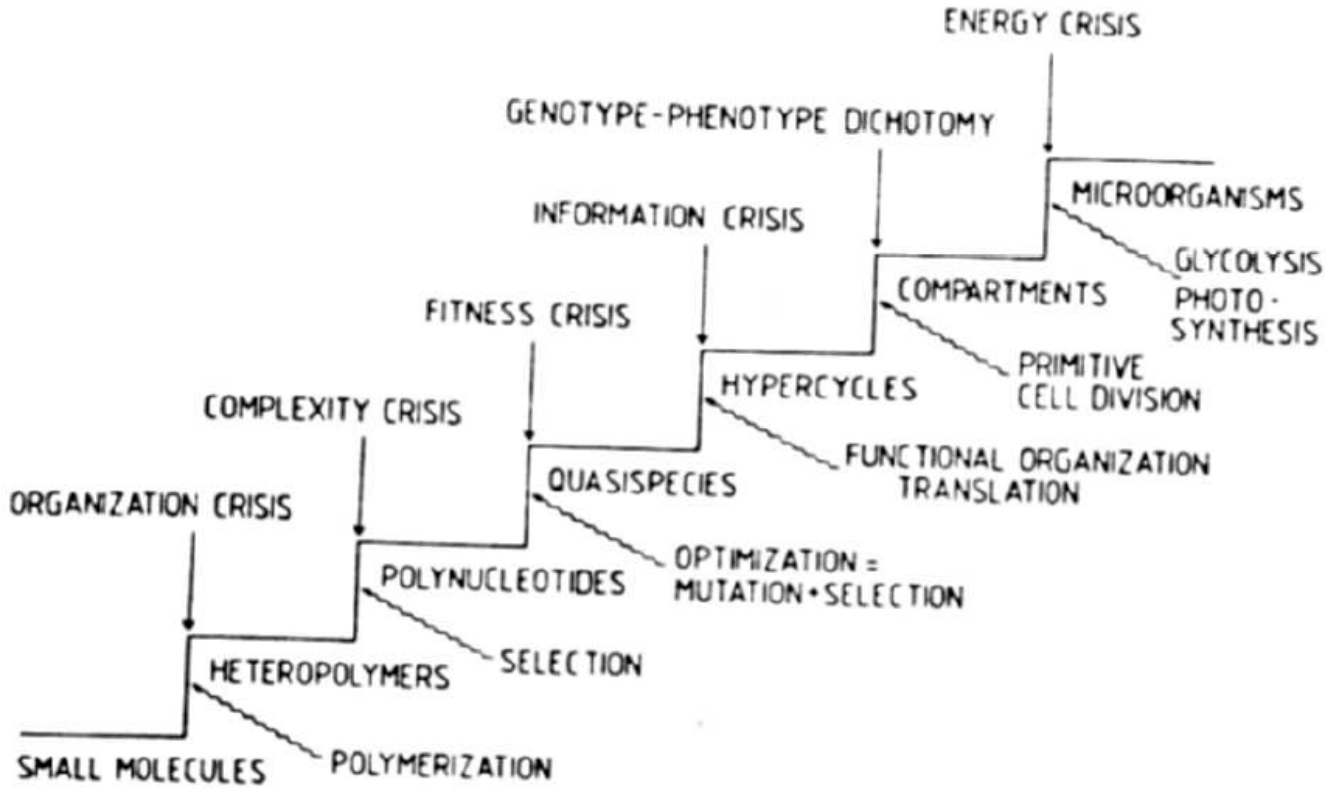
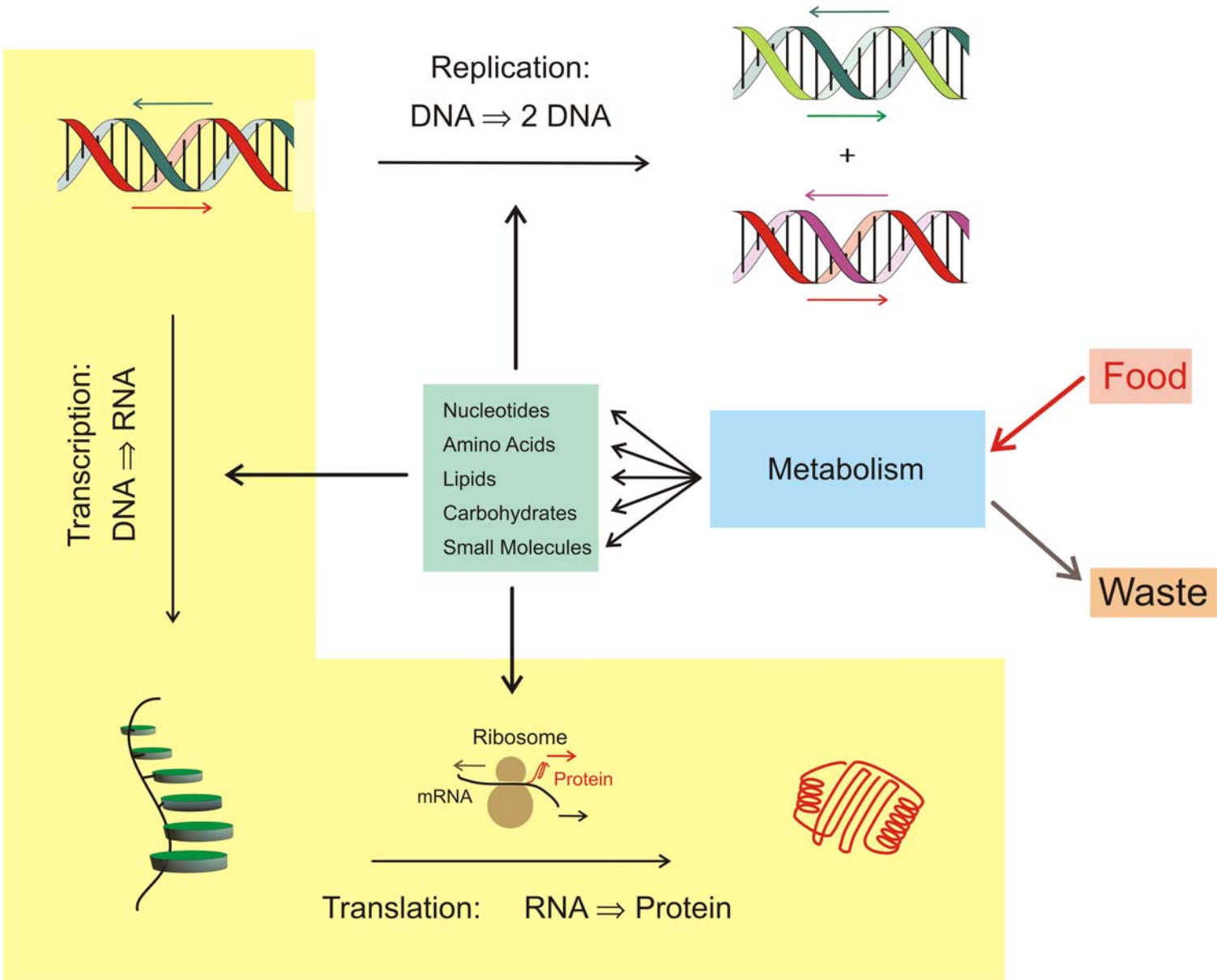


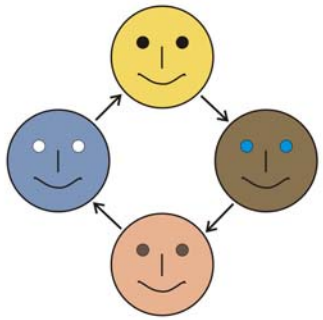
Fig. 11. Six critical steps during early evolution. On its way from small molecules to microorganisms the evolving system had to pass at least the six "crises" indicated by arrows. A "technical" innovation or new principle helped to overcome the obstacle the system had run into. The new concept is indicated by wavy arrows



A sketch of cellular information processing

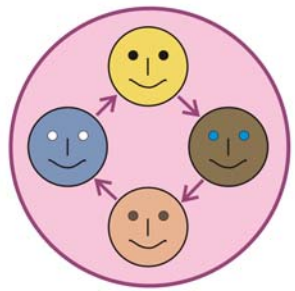
State I:
Independent Competing
Replicators

State II:
Functionally Coupled
Replicators



Parasite

State III:
New Unit of Selection



Parasite

State IV = State I:
Independent Competing
Units

A mechanism for major transitions
towards a hierarchically higher level
(nach Manfred Eigen und Peter Schuster)

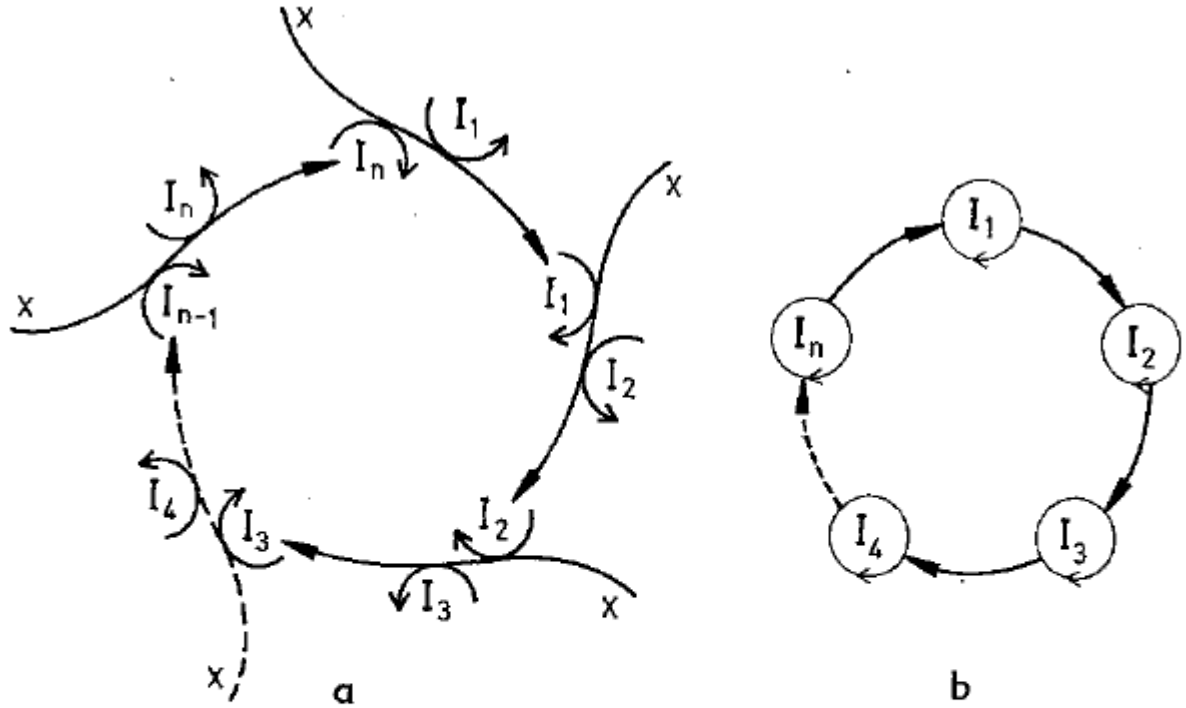
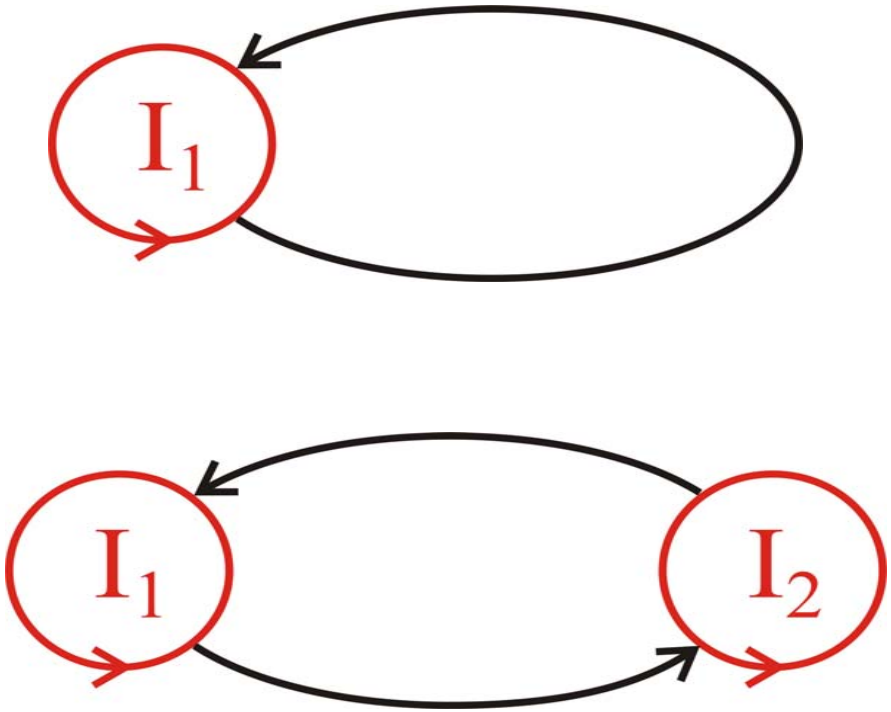
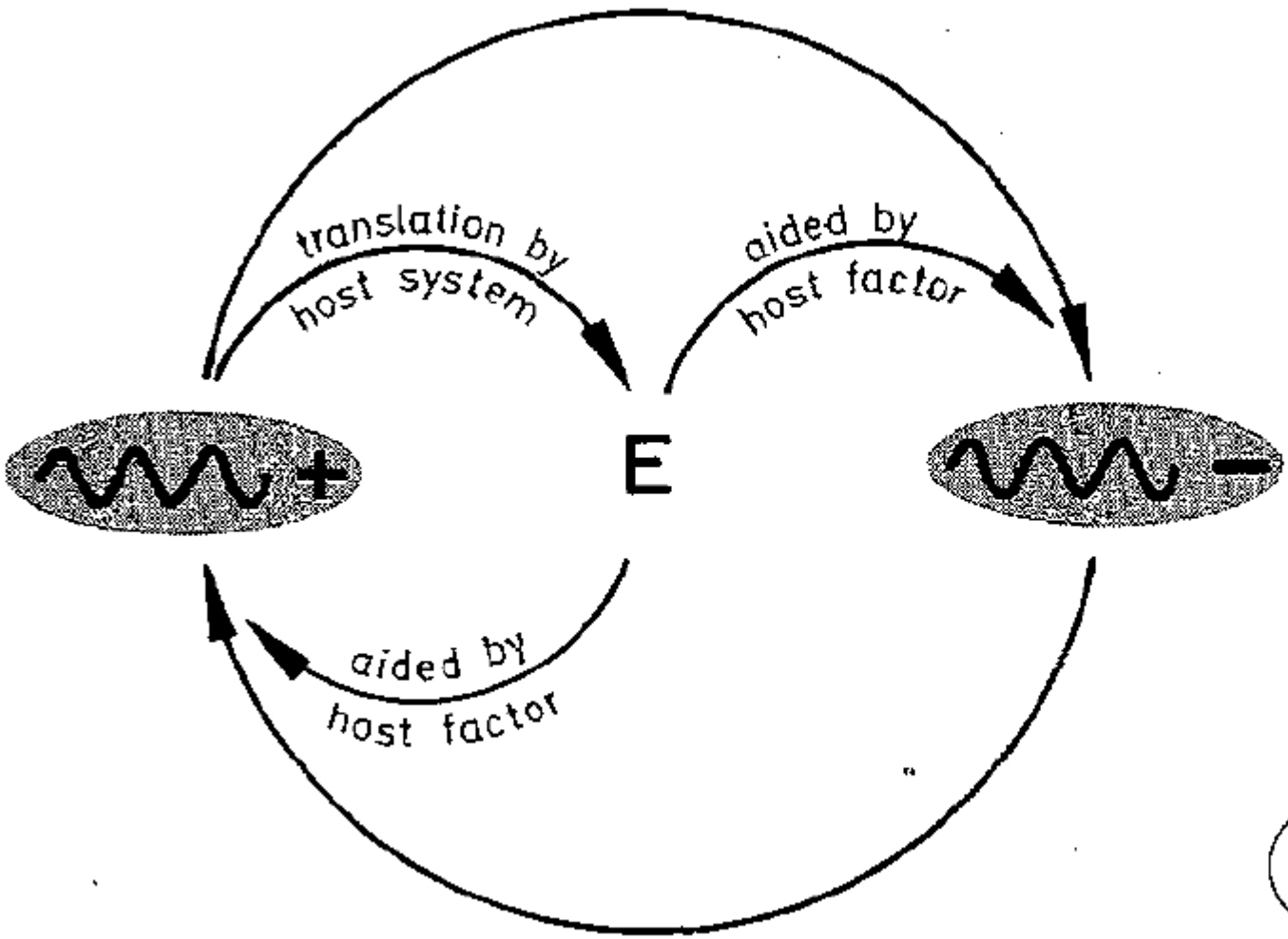


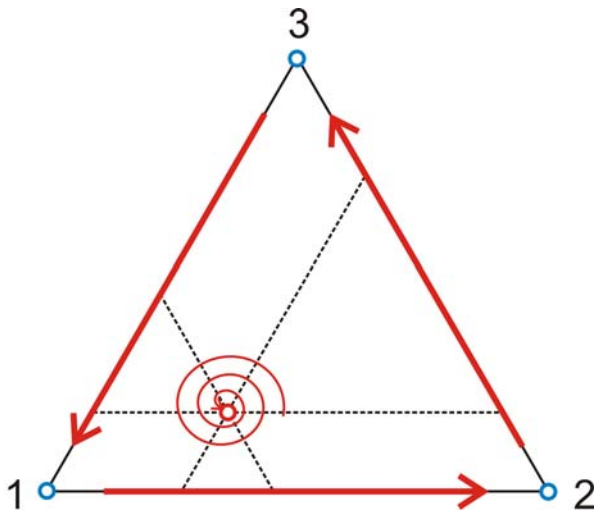
Fig. 7. *A catalytic hypercycle* consists of self-instructive units I_i with two-fold catalytic functions. As autocatalysts or—more generally—as catalytic cycles the intermediates I_i are able to instruct their own reproduction and, in addition, provide catalytic support for the reproduction of the subsequent intermediate (using the energy-rich building material X). The simplified graph (b) indicates the cyclic hierarchy

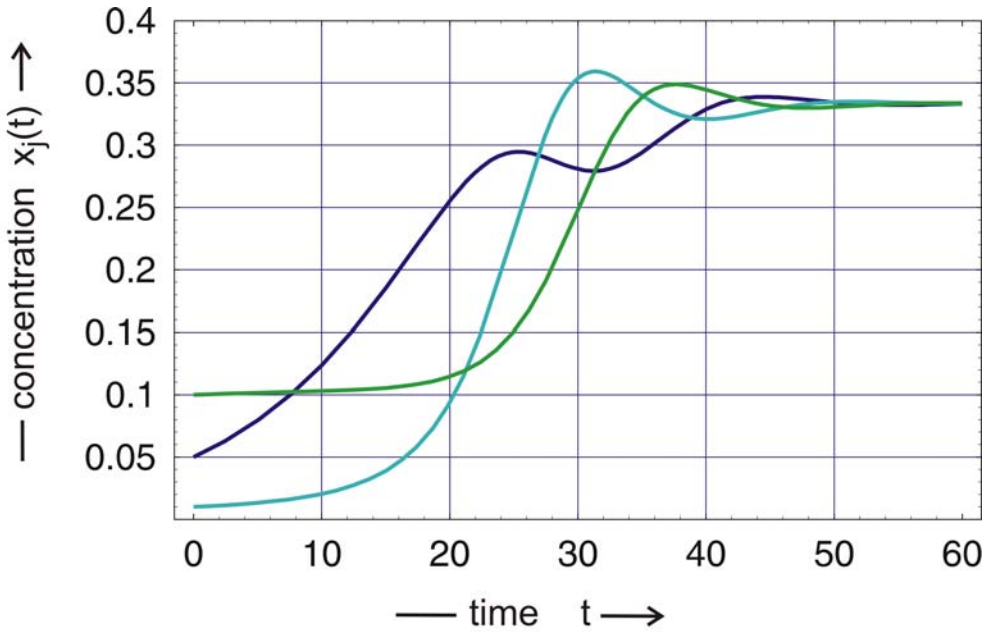


Hypercycles with one and two members are common in nature.

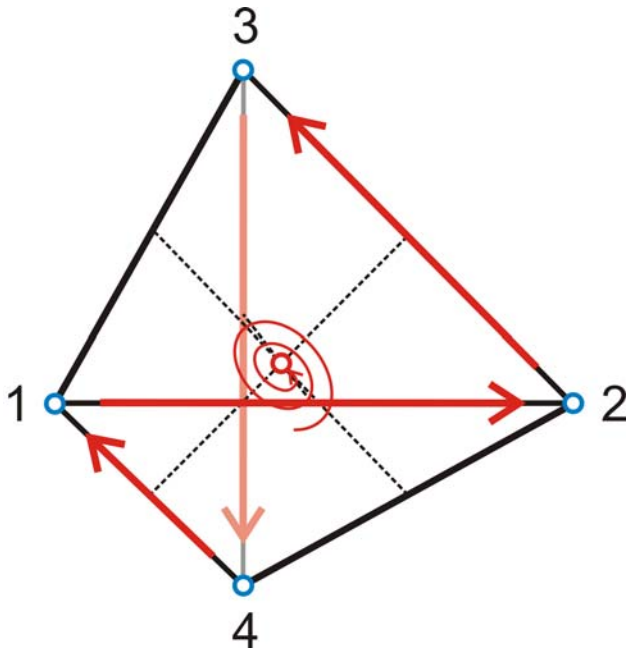

translation by
host system
aided by
host factor
+
E
−
aided by
host factor

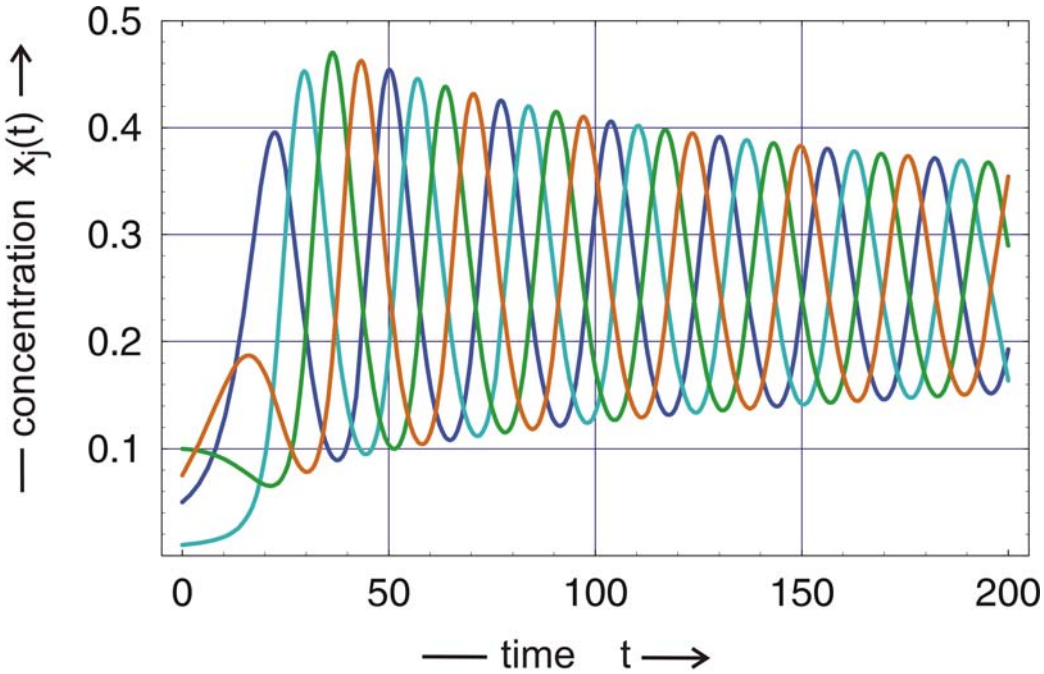

I
E



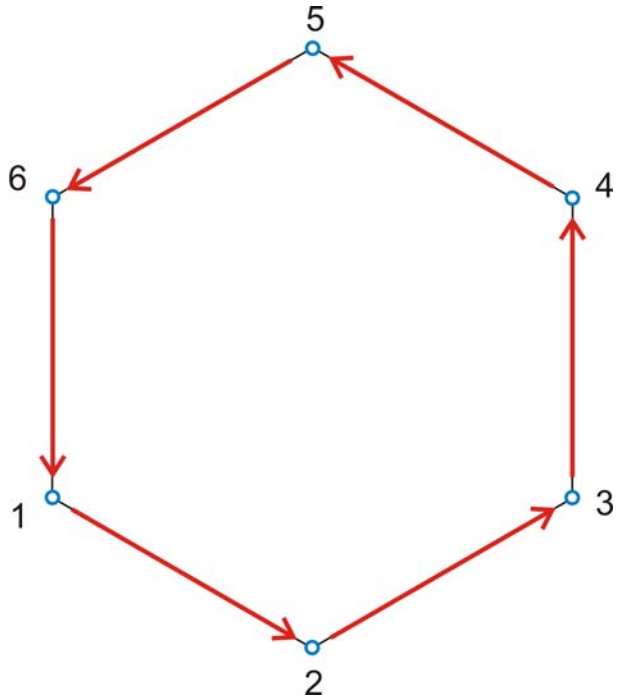


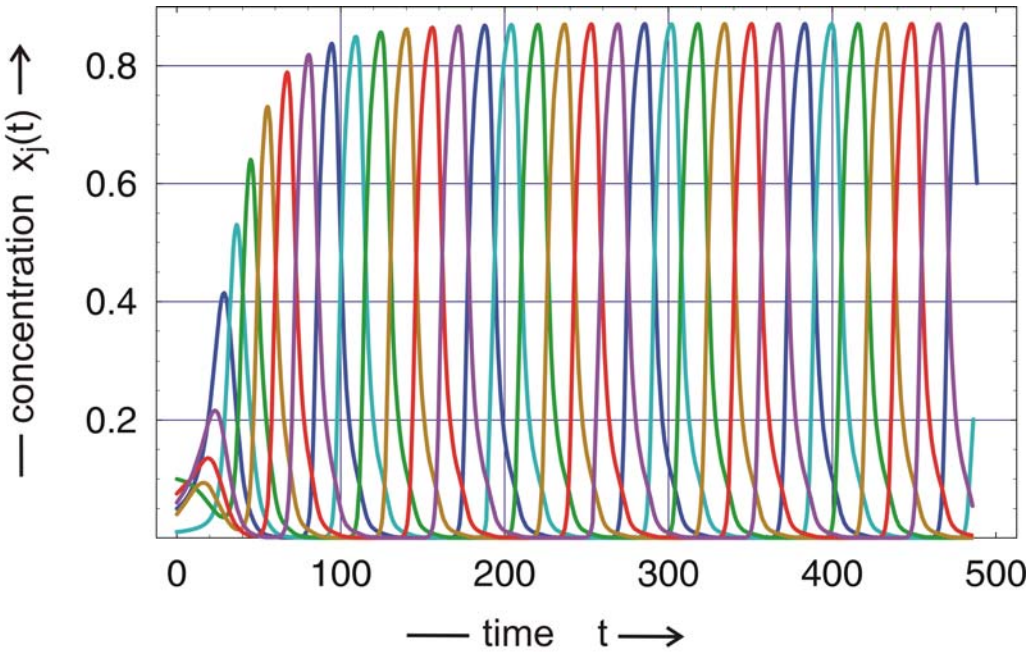
Hypercycle dynamics for n=3





Hypercycle dynamics for n=4





Hypercycle dynamics for n=6

Major transitions in evolution (John Maynard Smith and Eörs Szathmáry)

		membranes, controlled division	
Replicating molecules	⇒		molecules in compartments
		ligation of molecules, joint replication	
Independent replicators	⇒		chromosomes
		genetic code, ribosome	
RNA as gen und enzyme	⇒		DNA and protein
		endosymbiosis	
Prokaryotes	⇒		eukaryotes
		origin of sexual replication	
Asexually replicating clones	⇒		sexually replicating populations
		cell differentiation and development	
Protists	⇒		plants, fungi and animals
		formation of non-reproductive casts	
Solitary individuals	⇒		animal colonies
		language, writing, culture, …	
Primate societies	⇒		human societies

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks