1. Autocatalytic chemical reactions in the flow reactor

2. Replication, mutation, selection and Shannon information

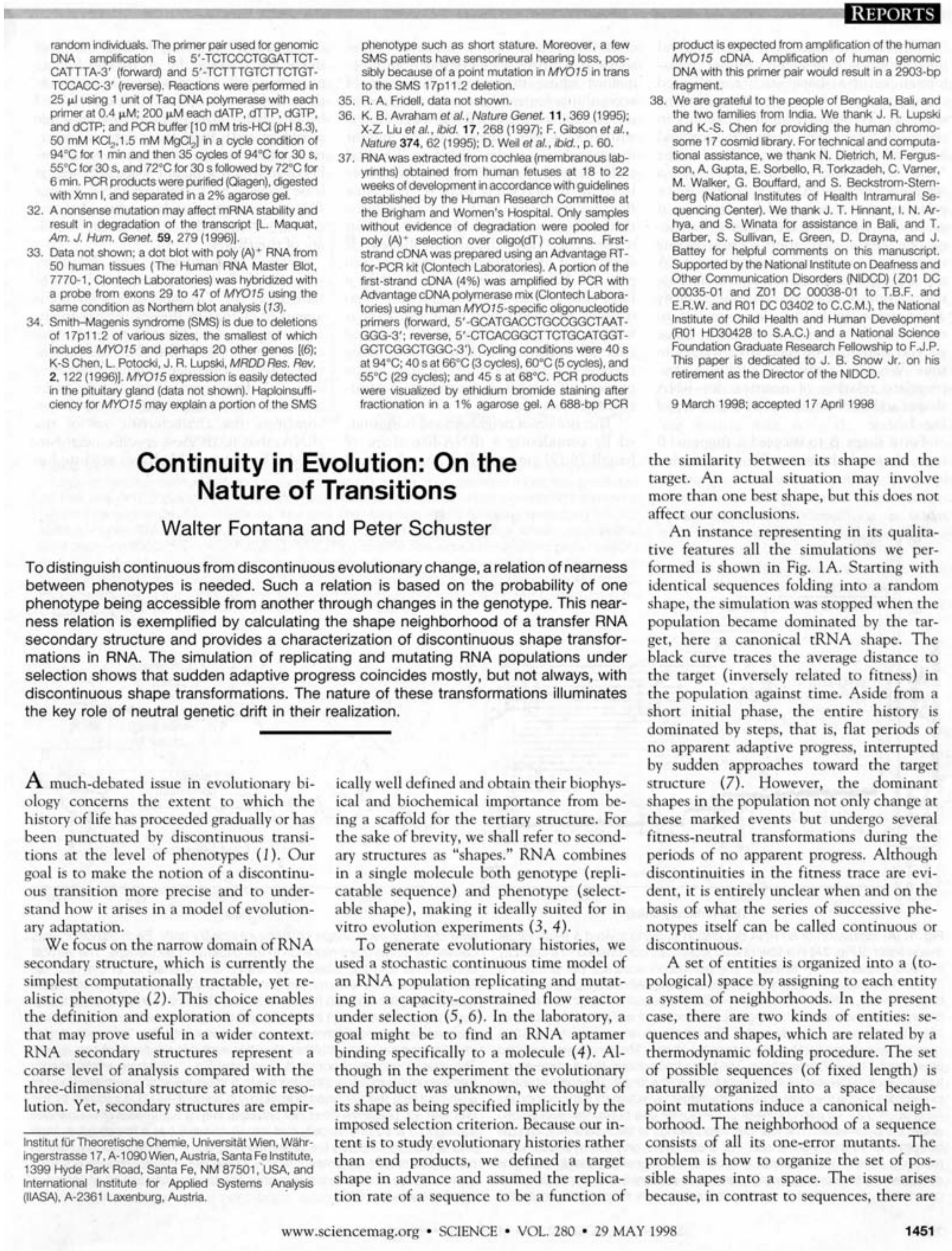
3. **Evolution *in silico* and optimization of RNA structures**

4. Random walks and ‚ensemble learning‘

5. Sequence-structure maps, neutral networks, and intersections

Evolution *in silico*

W. Fontana, P. Schuster,
Science **280** (1998), 1451-1455



random individuals. The primer pair used for genomic DNA amplification is 5′-TCTCCCTGGATTCTCATTTA-3′ (forward) and 5′-TCTTTGTCTTCTGTTCCACC-3′ (reverse). Reactions were performed in 25 μl using 1 unit of Taq DNA polymerase with each primer at 0.4 μM; 200 μM each dATP, dTTP, dGTP, and dCTP; and PCR buffer [10 mM tris-HCl (pH 8.3), 50 mM KCl_2,1.5 mM $MgCl_2$] in a cycle condition of 94°C for 1 min and then 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s followed by 72°C for 6 min. PCR products were purified (Qiagen), digested with Xmn I, and separated in a 2% agarose gel.

32. A nonsense mutation may affect mRNA stability and result in degradation of the transcript [L. Maquat, *Am. J. Hum. Genet.* **59**, 279 (1996)].
33. Data not shown; a dot blot with poly $(A)^+$ RNA from 50 human tissues (The Human RNA Master Blot, 7770-1, Clontech Laboratories) was hybridized with a probe from exons 29 to 47 of *MYO15* using the same condition as Northern blot analysis (*13*).
34. Smith–Magenis syndrome (SMS) is due to deletions of 17p11.2 of various sizes, the smallest of which includes *MYO15* and perhaps 20 other genes [(*6*); K-S Chen, L. Potocki, J. R. Lupski, *MRDD Res. Rev.* **2**, 122 (1996)]. *MYO15* expression is easily detected in the pituitary gland (data not shown). Haploinsufficiency for *MYO15* may explain a portion of the SMS phenotype such as short stature. Moreover, a few SMS patients have sensorineural hearing loss, possibly because of a point mutation in *MYO15* in trans to the SMS 17p11.2 deletion.
35. R. A. Fridell, data not shown.
36. K. B. Avraham *et al.*, *Nature Genet.* **11**, 369 (1995); X-Z. Liu *et al.*, *ibid.* **17**, 268 (1997); F. Gibson *et al.*, *Nature* **374**, 62 (1995); D. Weil *et al.*, *ibid.*, p. 60.
37. RNA was extracted from cochlea (membranous labyrinths) obtained from human fetuses at 18 to 22 weeks of development in accordance with guidelines established by the Human Research Committee at the Brigham and Women's Hospital. Only samples without evidence of degradation were pooled for poly $(A)^+$ selection over oligo(dT) columns. First-strand cDNA was prepared using an Advantage RT-for-PCR kit (Clontech Laboratories). A portion of the first-strand cDNA (4%) was amplified by PCR with Advantage cDNA polymerase mix (Clontech Laboratories) using human *MYO15*-specific oligonucleotide primers (forward, 5′-GCATGACCTGCCGGCTAATGGG-3′; reverse, 5′-CTCACGGCTTCTGCATGGTGCTCGGCTGGC-3′). Cycling conditions were 40 s at 94°C; 40 s at 66°C (3 cycles), 60°C (5 cycles), and 55°C (29 cycles); and 45 s at 68°C. PCR products were visualized by ethidium bromide staining after fractionation in a 1% agarose gel. A 688-bp PCR product is expected from amplification of the human *MYO15* cDNA. Amplification of human genomic DNA with this primer pair would result in a 2903-bp fragment.
38. We are grateful to the people of Bengkala, Bali, and the two families from India. We thank J. R. Lupski and K.-S. Chen for providing the human chromosome 17 cosmid library. For technical and computational assistance, we thank N. Dietrich, M. Fergusson, A. Gupta, E. Sorbello, R. Torkzadeh, C. Varner, M. Walker, G. Bouffard, and S. Beckstrom-Sternberg (National Institutes of Health Intramural Sequencing Center). We thank J. T. Hinnant, I. N. Arhya, and S. Winata for assistance in Bali, and T. Barber, S. Sullivan, E. Green, D. Drayna, and J. Battey for helpful comments on this manuscript. Supported by the National Institute on Deafness and Other Communication Disorders (NIDCD) (Z01 DC 00035-01 and Z01 DC 00038-01 to T.B.F. and E.R.W. and R01 DC 03402 to C.C.M.), the National Institute of Child Health and Human Development (R01 HD30428 to S.A.C.) and a National Science Foundation Graduate Research Fellowship to F.J.P. This paper is dedicated to J. B. Snow Jr. on his retirement as the Director of the NIDCD.

9 March 1998; accepted 17 April 1998

Continuity in Evolution: On the Nature of Transitions

Walter Fontana and Peter Schuster

To distinguish continuous from discontinuous evolutionary change, a relation of nearness between phenotypes is needed. Such a relation is based on the probability of one phenotype being accessible from another through changes in the genotype. This nearness relation is exemplified by calculating the shape neighborhood of a transfer RNA secondary structure and provides a characterization of discontinuous shape transformations in RNA. The simulation of replicating and mutating RNA populations under selection shows that sudden adaptive progress coincides mostly, but not always, with discontinuous shape transformations. The nature of these transformations illuminates the key role of neutral genetic drift in their realization.

A much-debated issue in evolutionary biology concerns the extent to which the history of life has proceeded gradually or has been punctuated by discontinuous transitions at the level of phenotypes (*1*). Our goal is to make the notion of a discontinuous transition more precise and to understand how it arises in a model of evolutionary adaptation.

We focus on the narrow domain of RNA secondary structure, which is currently the simplest computationally tractable, yet realistic phenotype (*2*). This choice enables the definition and exploration of concepts that may prove useful in a wider context. RNA secondary structures represent a coarse level of analysis compared with the three-dimensional structure at atomic resolution. Yet, secondary structures are empirically well defined and obtain their biophysical and biochemical importance from being a scaffold for the tertiary structure. For the sake of brevity, we shall refer to secondary structures as "shapes." RNA combines in a single molecule both genotype (replicatable sequence) and phenotype (selectable shape), making it ideally suited for in vitro evolution experiments (*3*, *4*).

To generate evolutionary histories, we used a stochastic continuous time model of an RNA population replicating and mutating in a capacity-constrained flow reactor under selection (*5*, *6*). In the laboratory, a goal might be to find an RNA aptamer binding specifically to a molecule (*4*). Although in the experiment the evolutionary end product was unknown, we thought of its shape as being specified implicitly by the imposed selection criterion. Because our intent is to study evolutionary histories rather than end products, we defined a target shape in advance and assumed the replication rate of a sequence to be a function of the similarity between its shape and the target. An actual situation may involve more than one best shape, but this does not affect our conclusions.

An instance representing in its qualitative features all the simulations we performed is shown in Fig. 1A. Starting with identical sequences folding into a random shape, the simulation was stopped when the population became dominated by the target, here a canonical tRNA shape. The black curve traces the average distance to the target (inversely related to fitness) in the population against time. Aside from a short initial phase, the entire history is dominated by steps, that is, flat periods of no apparent adaptive progress, interrupted by sudden approaches toward the target structure (*7*). However, the dominant shapes in the population not only change at these marked events but undergo several fitness-neutral transformations during the periods of no apparent progress. Although discontinuities in the fitness trace are evident, it is entirely unclear when and on the basis of what the series of successive phenotypes itself can be called continuous or discontinuous.

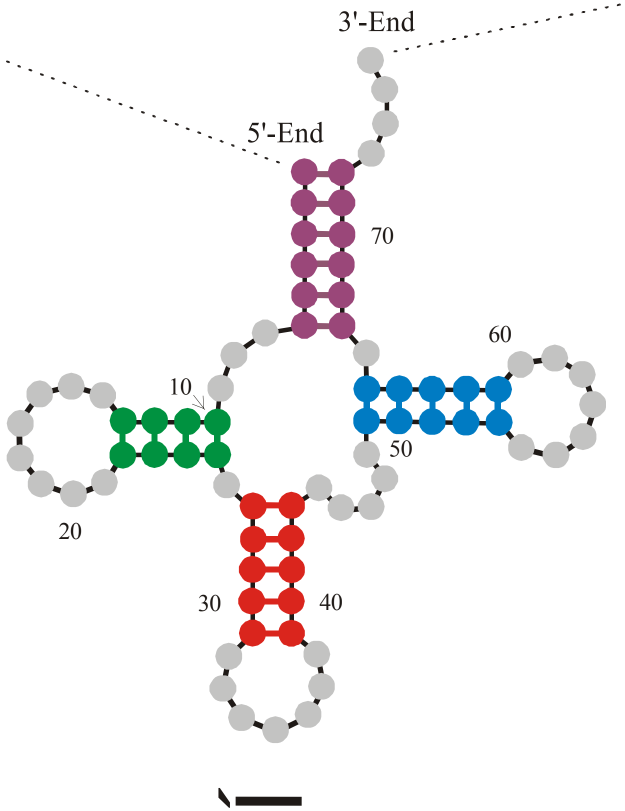
A set of entities is organized into a (topological) space by assigning to each entity a system of neighborhoods. In the present case, there are two kinds of entities: sequences and shapes, which are related by a thermodynamic folding procedure. The set of possible sequences (of fixed length) is naturally organized into a space because point mutations induce a canonical neighborhood. The neighborhood of a sequence consists of all its one-error mutants. The problem is how to organize the set of possible shapes into a space. The issue arises because, in contrast to sequences, there are

Institut für Theoretische Chemie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA, and International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria.

Sequence

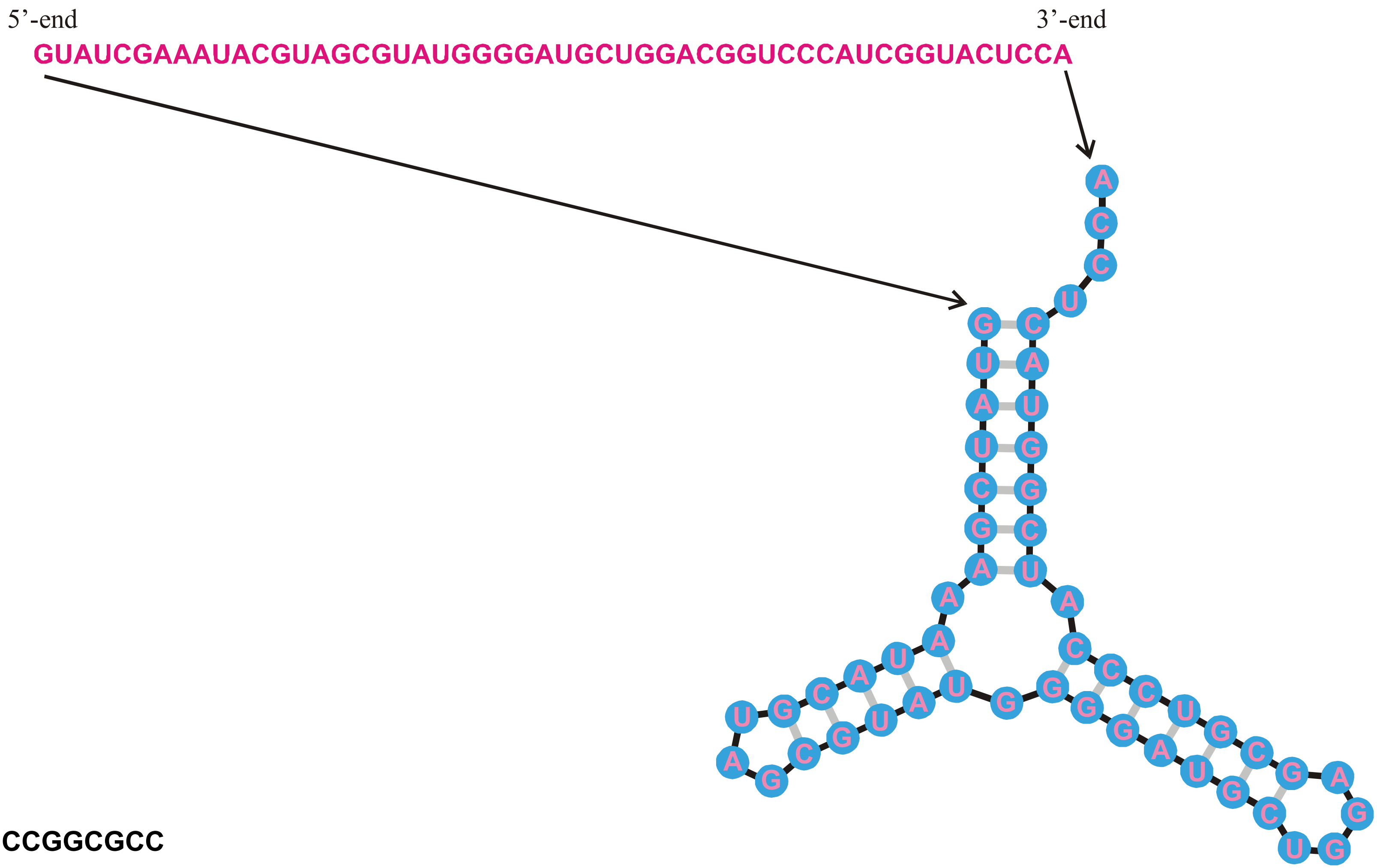
5'-End GCGGAUUUAGCUCAGDDGGGAGAGCMCCAGACUGAAYAUCUGGAGMUCCUGUGTPCGAUCCACAGAAUUCGCACCA 3'-End

Secondary structure



Symbolic notation 5'-End ((((((···((((········))))·(((((·······)))))·····(((((·······)))))·))))))···· 3'-End

A symbolic notation of RNA secondary structure that is equivalent to the conventional graphs



5’-end

3’-end

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

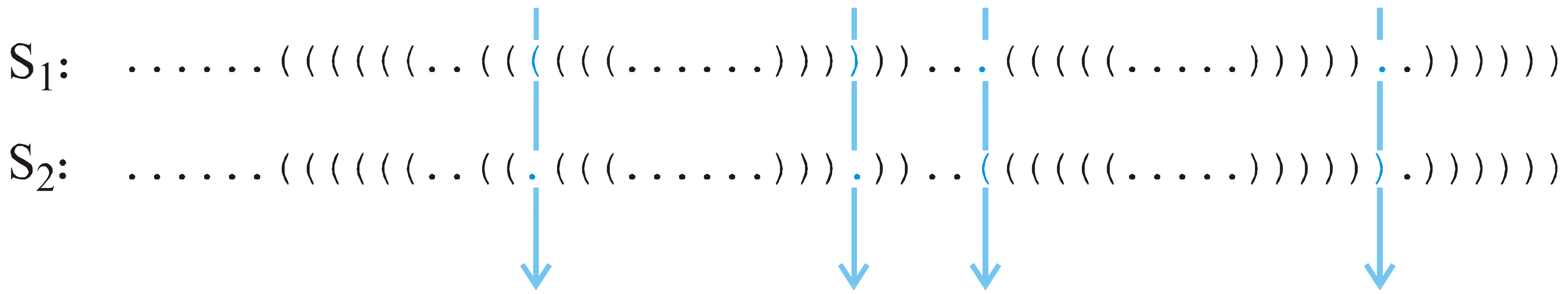
RNAStudio.lnk

GGCGCGCCCGGCGCC

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

UGGUUACGCGUUGGGGUAACGAAGAUUCCGAGAGGAGUUUAGUGACUAGAGG

Folding of RNA sequences into secondary structures of minimal free energy, DG_0^{300}



```
S1:   ......((((((..((((((......))))))...(((((.....)))))..))))))
S2:   ......((((((..((.(((......))).))..((((((.....)))))).))))))
```

Hamming distance $d_H(S_1,S_2) = 4$

(i) $d_H(S_1,S_1) = 0$

(ii) $d_H(S_1,S_2) = d_H(S_2,S_1)$

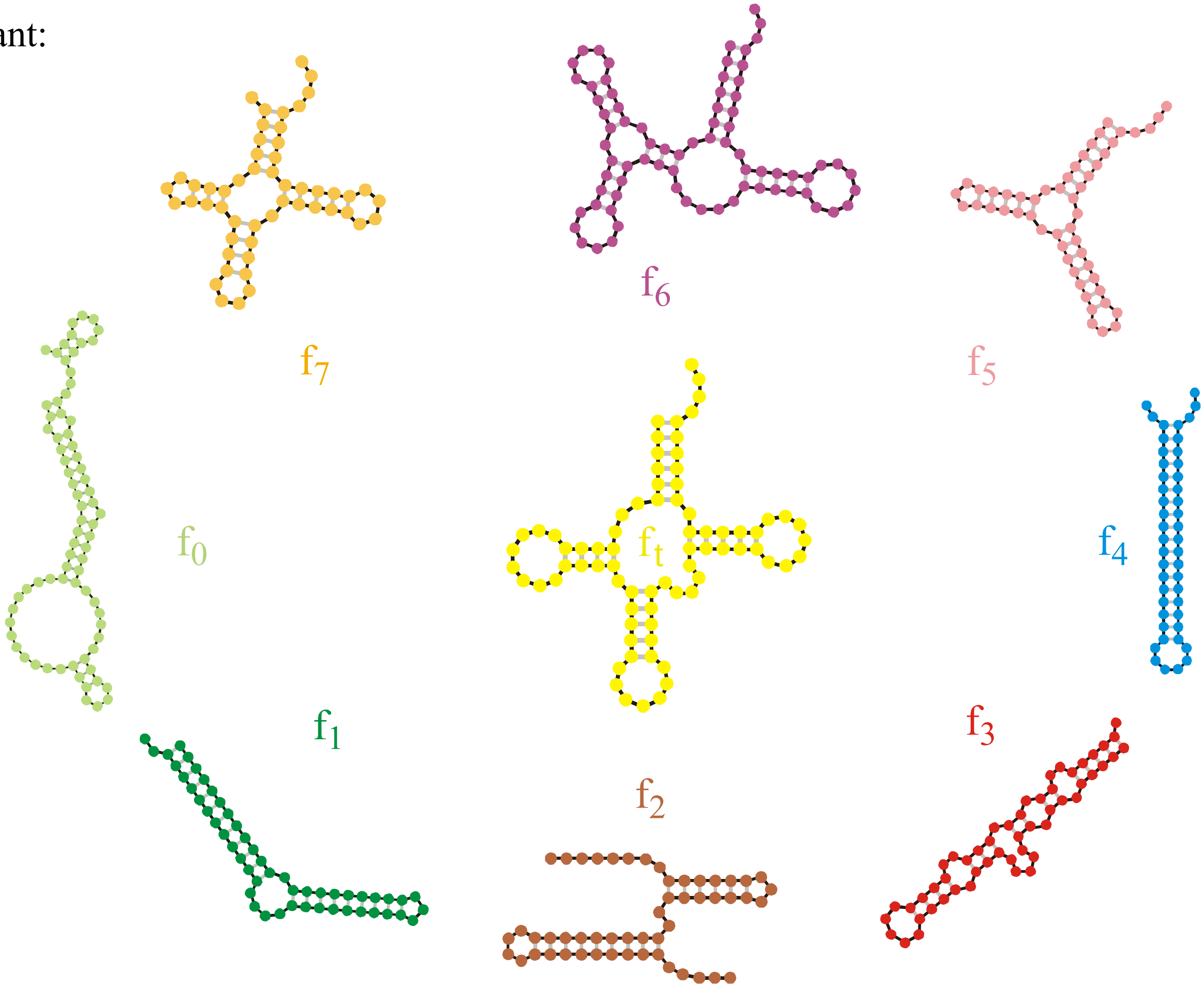
(iii) $d_H(S_1,S_3) \P d_H(S_1,S_2) + d_H(S_2,S_3)$

The Hamming distance between structures in parentheses notation forms a metric in structure space

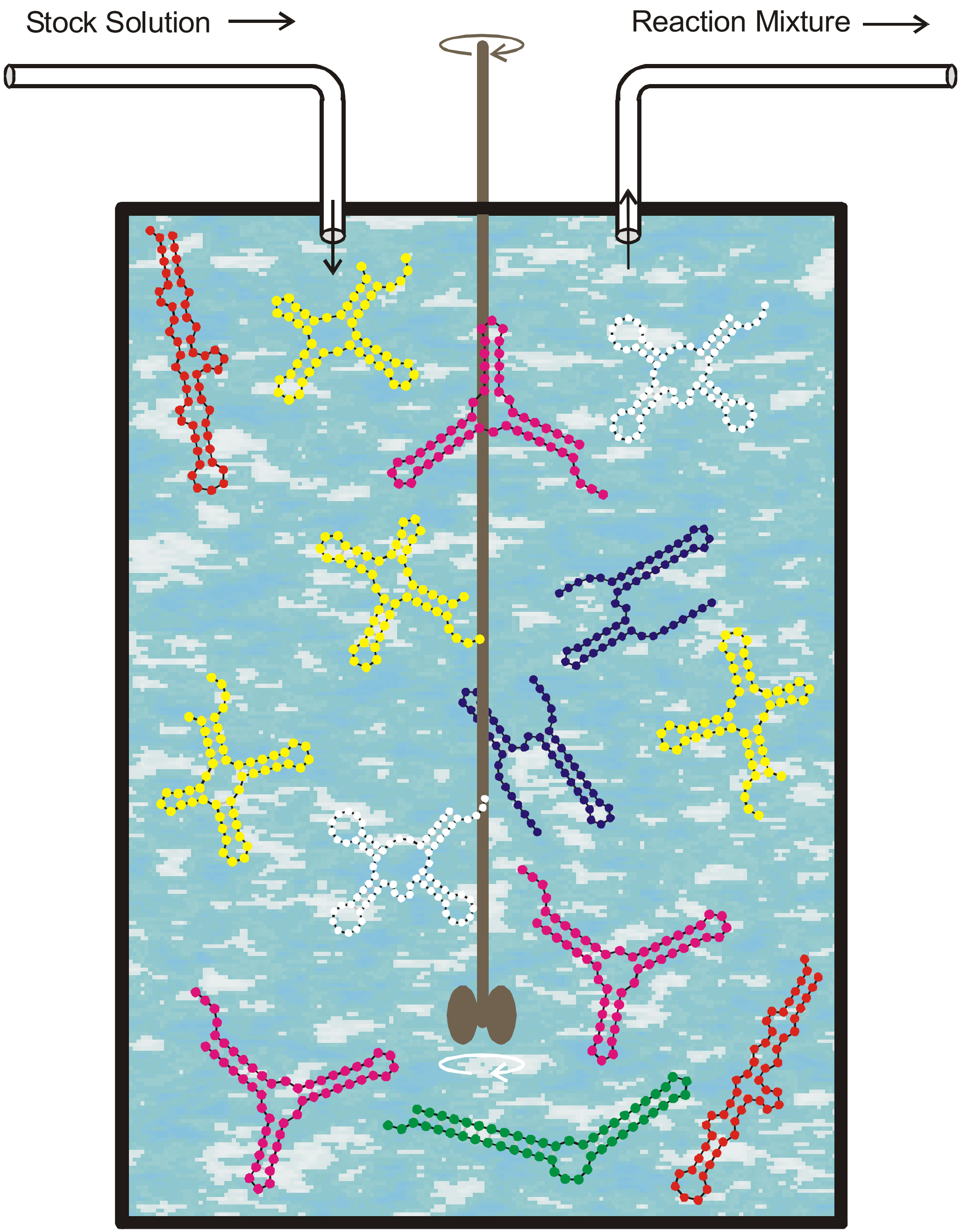
Replication rate constant:

$$f_k = g \,/\, [a + Dd_S^{\,(k)}]$$

$$Dd_S^{\,(k)} = d_H(S_k,S_t)$$



Evaluation of RNA secondary structures yields replication rate constants



Replication rate constant:

$$f_k = g \,/\, [a + Dd_S^{\,(k)}]$$

$$Dd_S^{\,(k)} = d_H(S_k,S_t)$$

Selection constraint:

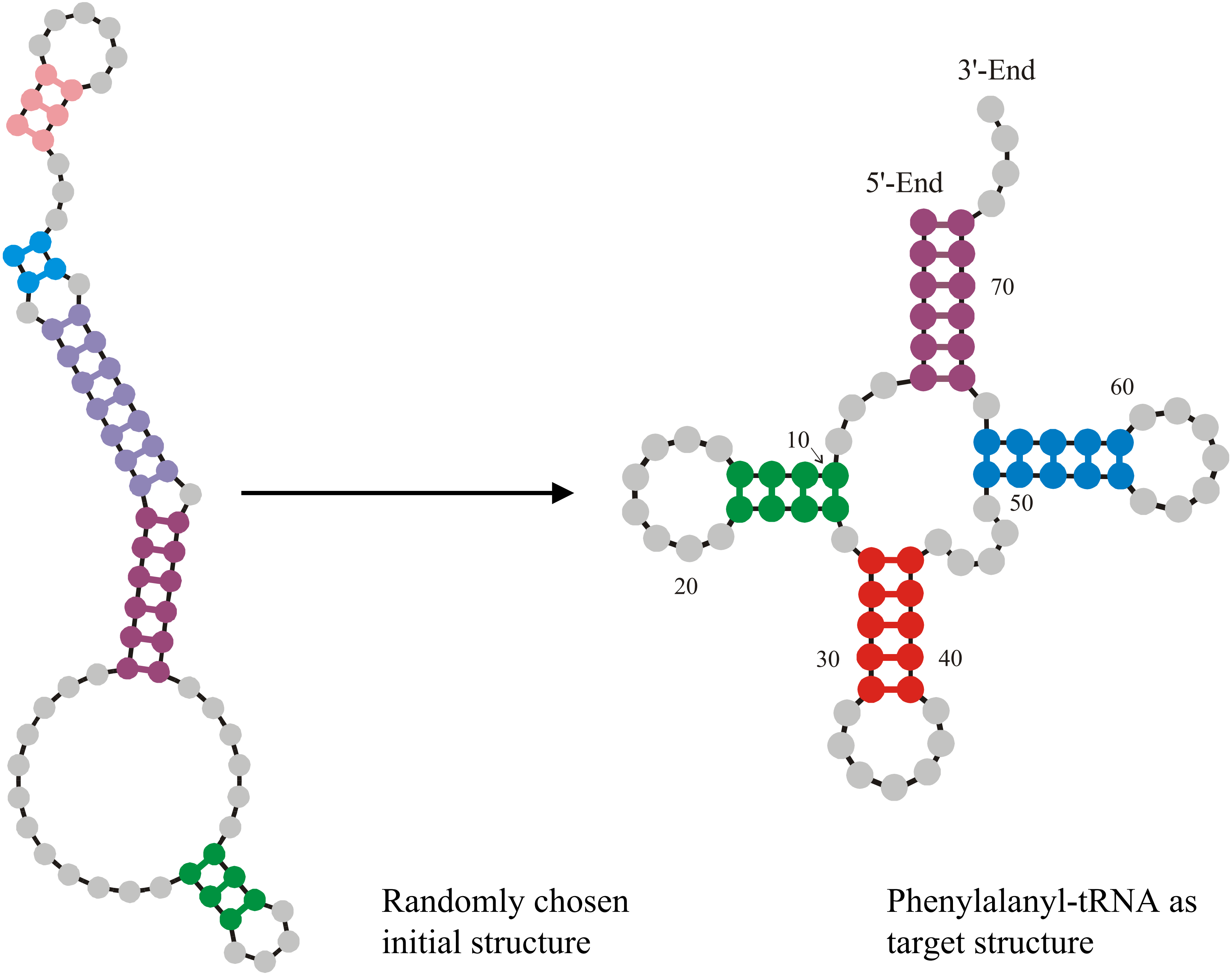
# RNA molecules is controlled by the flow

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

Constant mutation rate:

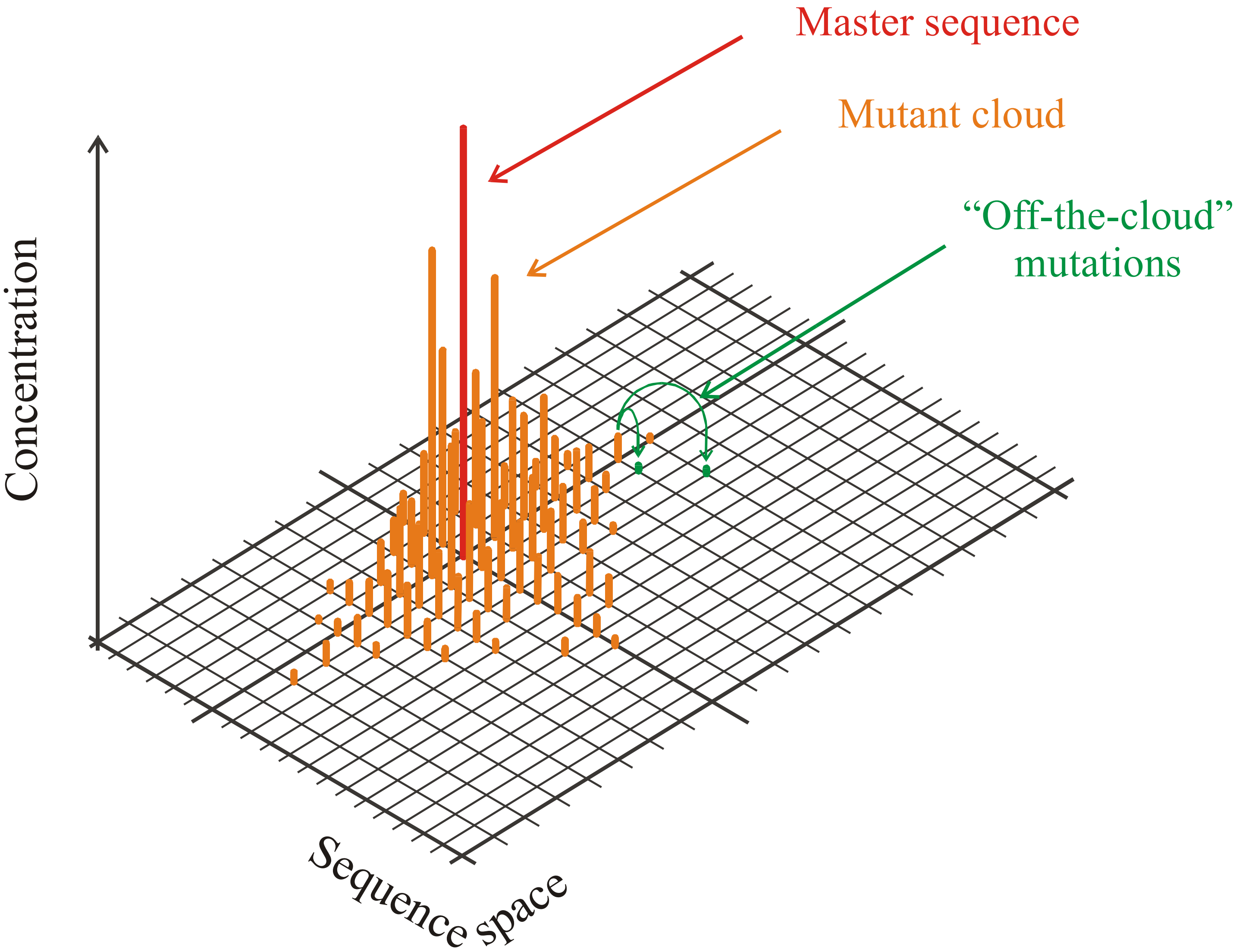
p = 0.001 per site and replication

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

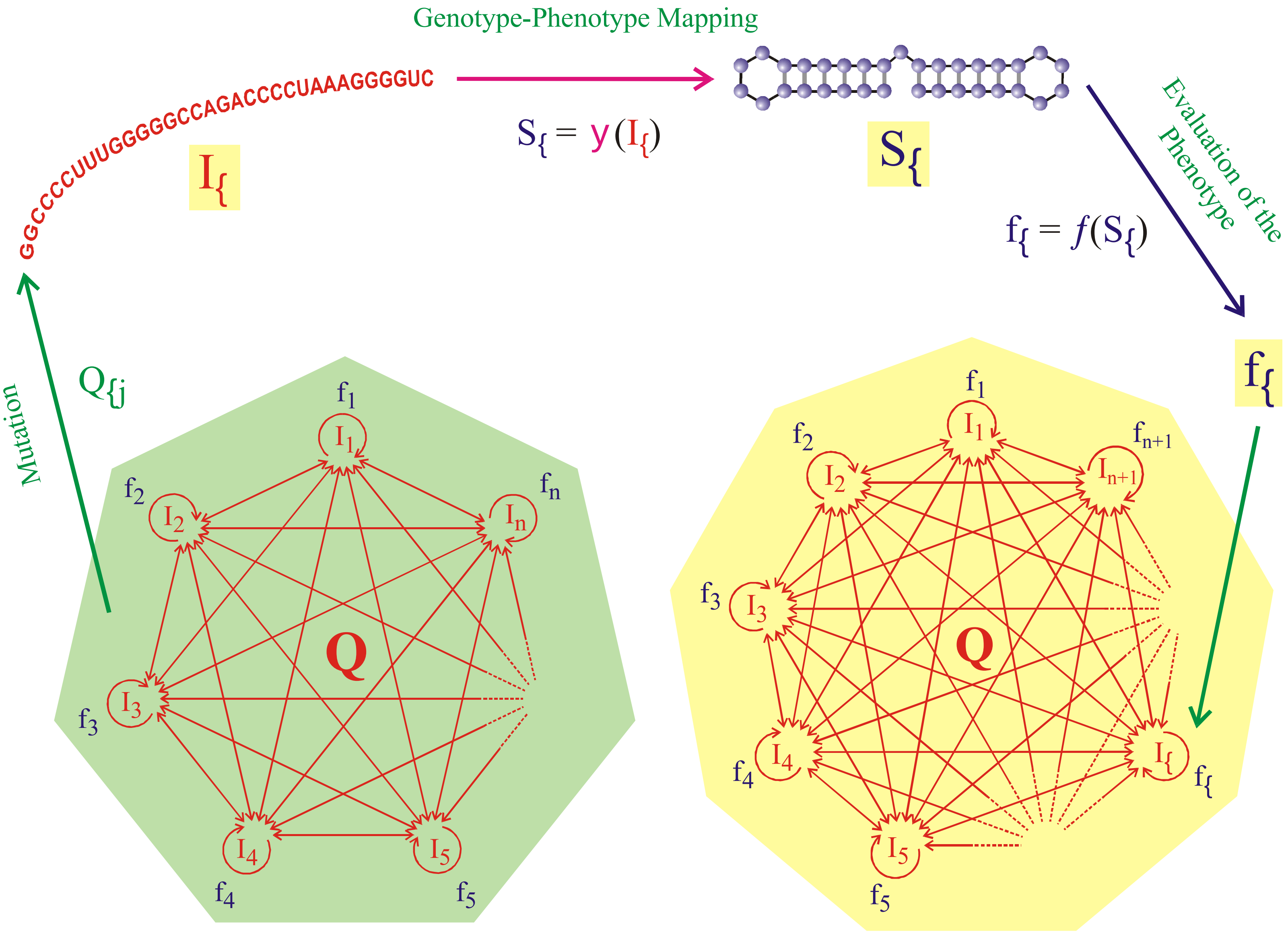


Randomly chosen initial structure

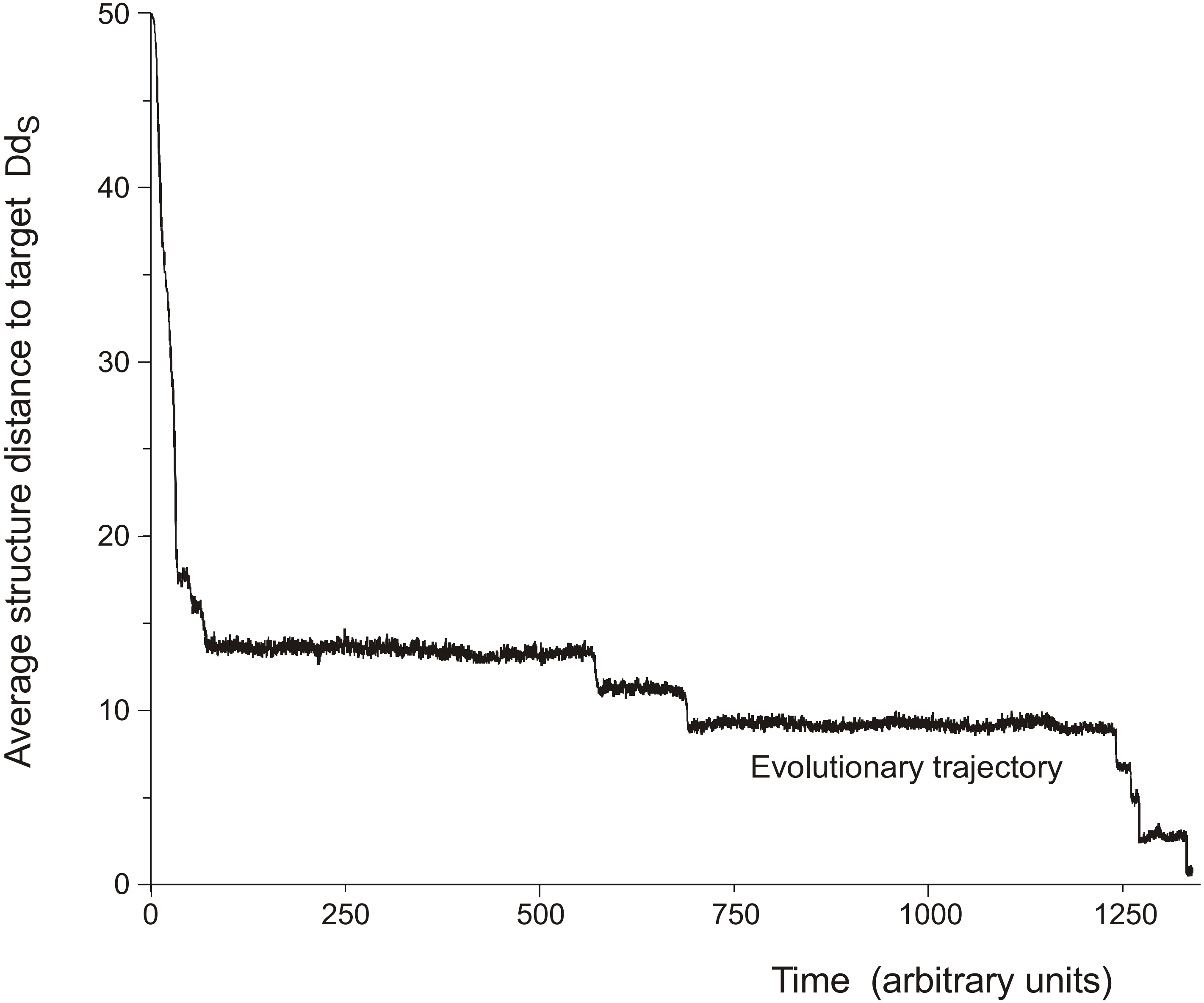
Phenylalanyl-tRNA as target structure



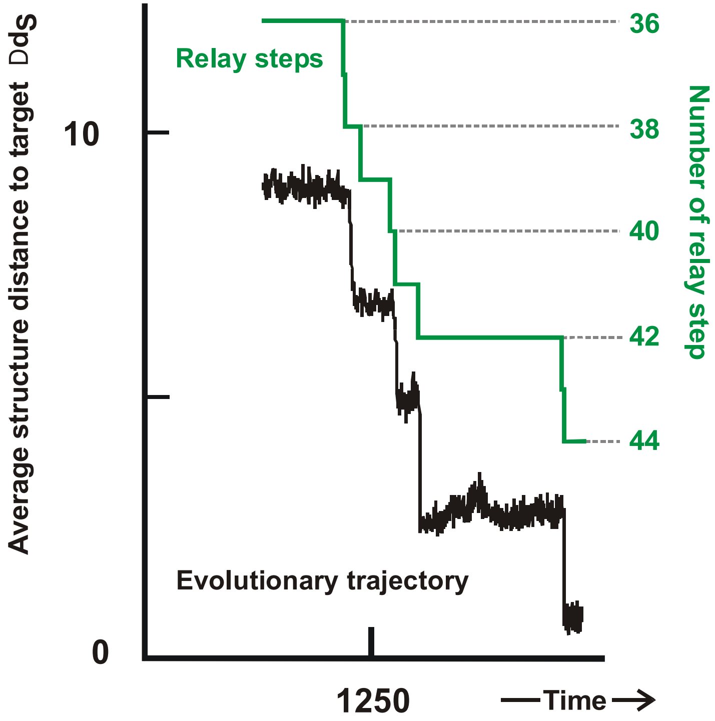
The molecular quasispecies
in sequence space

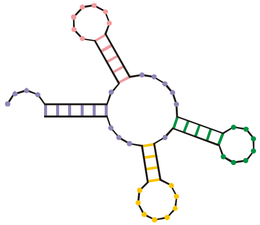


Evolutionary dynamics
including molecular phenotypes



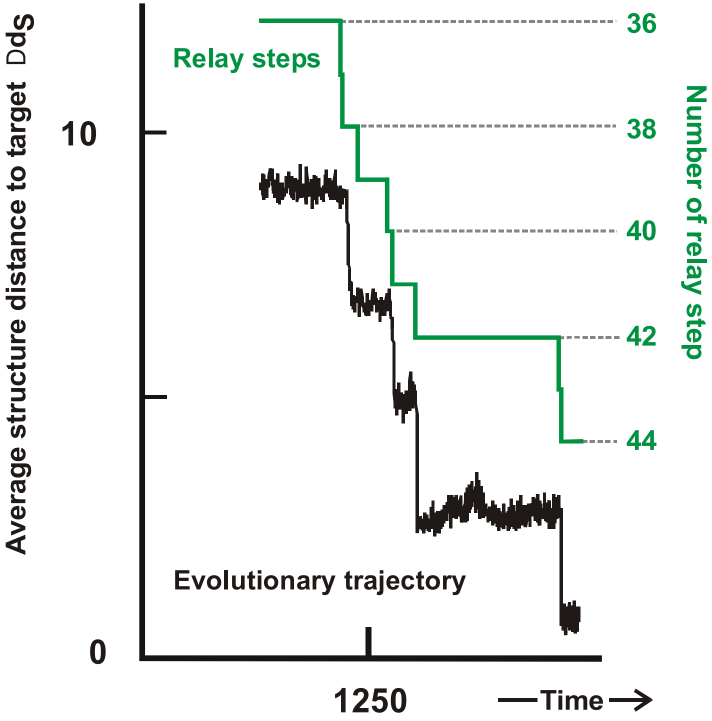
In silico optimization in the flow reactor: trajectory

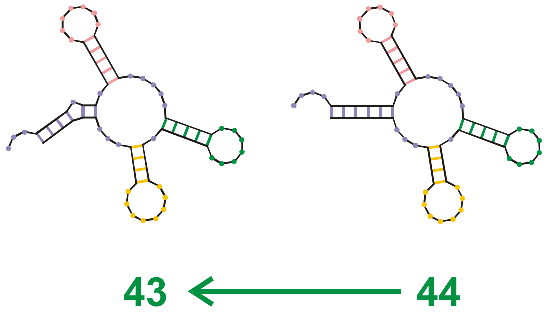




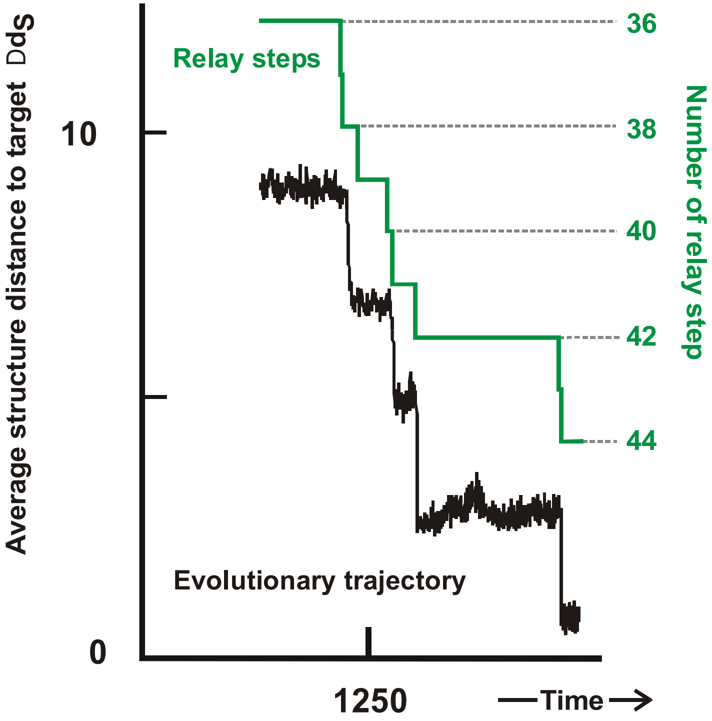
44

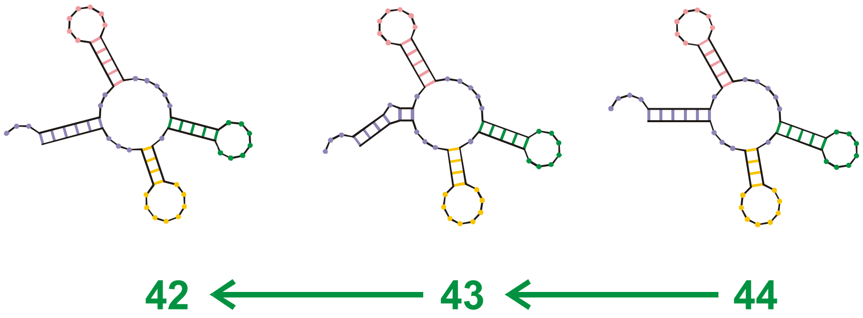
Final structure of the optimization process



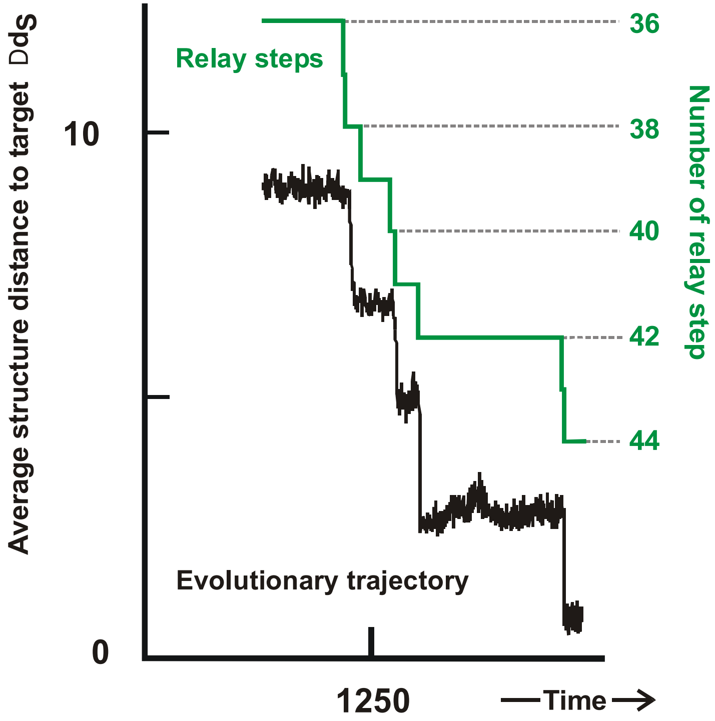


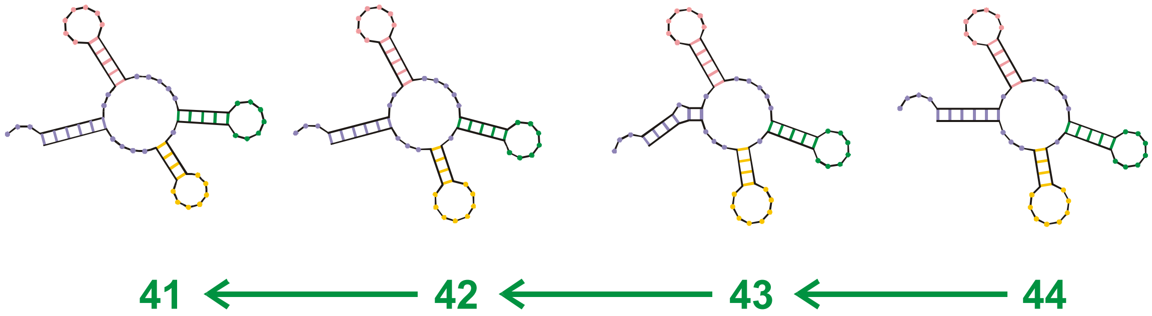
Reconstruction of the last step 43 Á 44



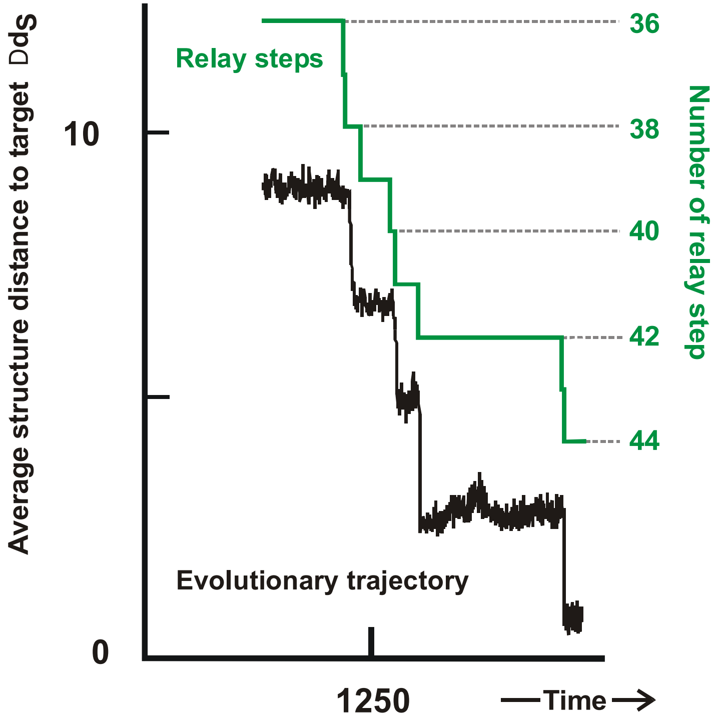


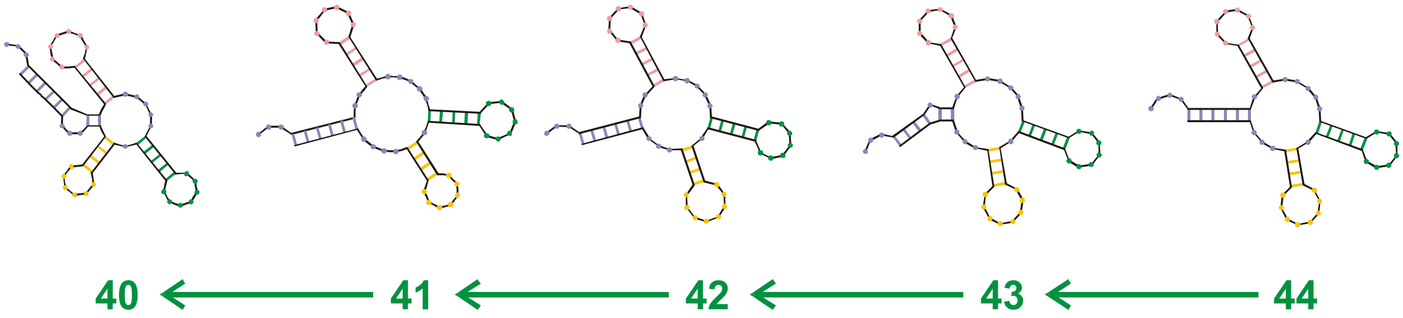
Reconstruction of last-but-one step 42 Á 43 (Á 44)



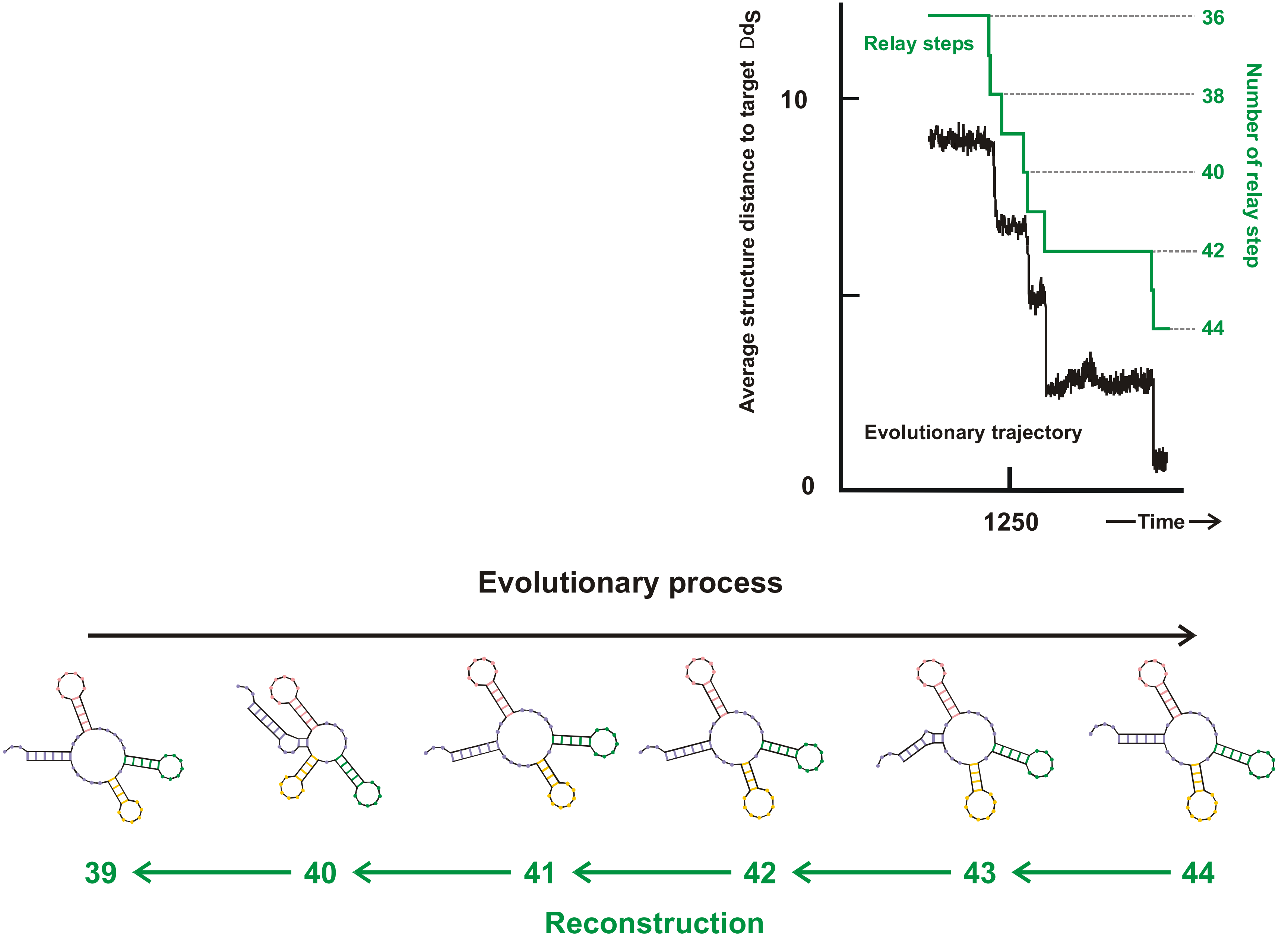


Reconstruction of step 41 Á 42 (Á 43 Á 44)





Reconstruction of step 40 Á 41 (Á 42 Á 43 Á 44)



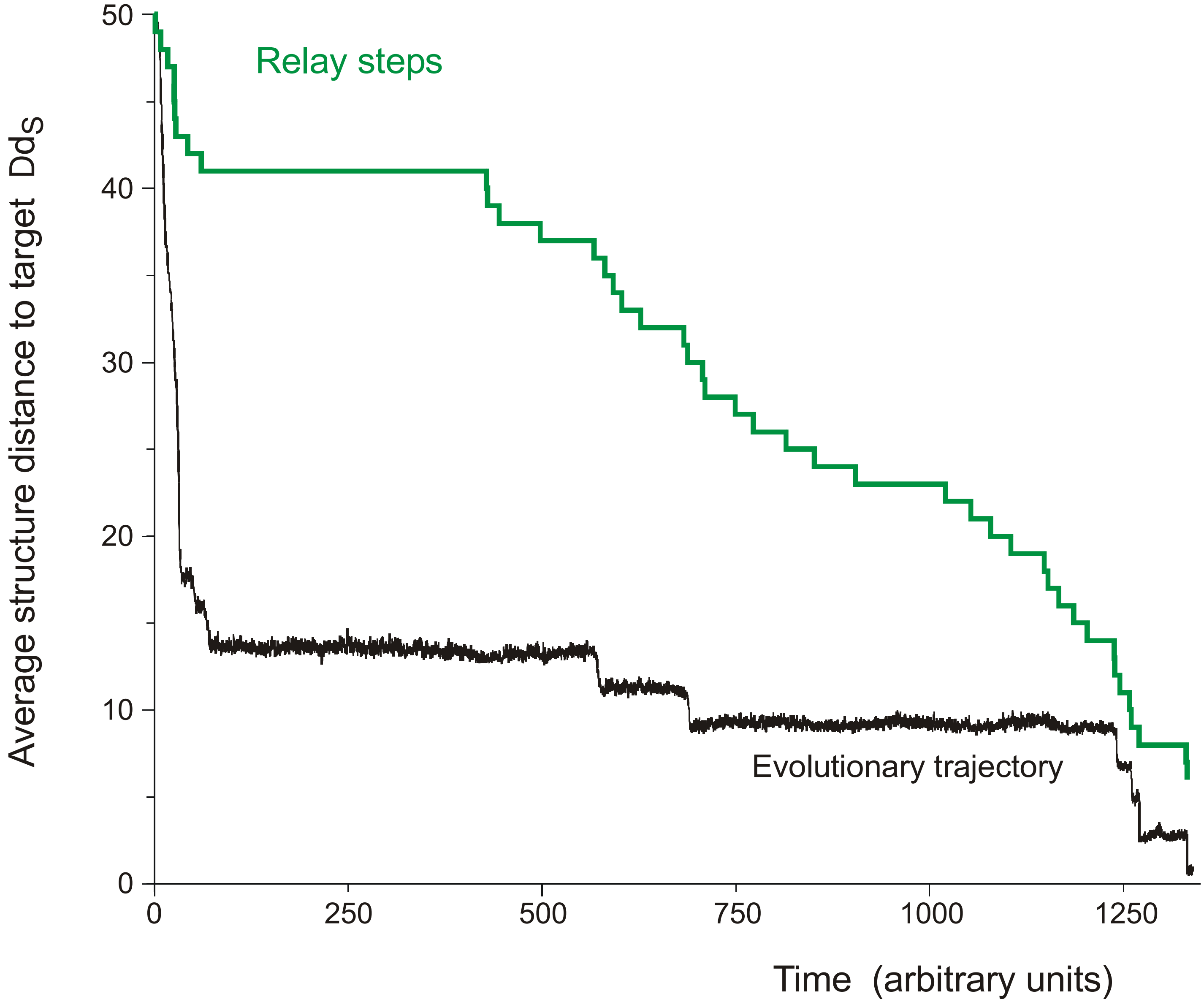
Reconstruction of the relay series

```
entry   GGGAUACAUGUGGCCCCUCAAGGCCCUAGCGAAACUGCUGCUGAAACCGUGUGAAUAAUCCGCACCCUGUCCCCGA
 39     ((((((.....((((......)))).(((((.......))))).....(((((.......)))))..))))))...
exit    GGGAUAUACGAGGCCCGUCAAGGCCGUAGCGAACCGACUGUUGAAACUGUGCGAAUAAUCCGCACCCUGUCCCGGG
entry   GGGAUAUACGGGGCCCGUCAAGGCCGUAGCGAACCGACUGUUGAAACUGUGCGAAUAAUCCGCACCCUGUCCCGGG
 40     ((((((...((((((......)))).(((((.......))))).....(((((.......)))))))))))))...
exit    GGGAUAUACGGGGCCCGUCAAGGCCGUAGCGAACCGACUGUUGAGACUGUGCGAAUAAUCCGCACCCUGUCCCGGG
entry   GGGAUAUACGGGCCCCGUCAAGGCCGUAGCGAACCGACUGUUGAGACUGUGCGAAUAAUCCGCACCCUGUCCCGGG
 41     ((((((....((((.......)))).(((((.......))))).....(((((.......)))))..))))))...
exit    GGGAUAUACGGGCCCCUUCAAGGCCAUAGCGAACCGACUGUUGAAACUGUGCGAAUAAUCCGCACCCUGUCCCGGA
entry   GGGAUAUACGGGCCCCUUCAAGCCCAUAGCGAACCGACUGUUGAAACUGUGCGAAUAAUCCGCACCCUGUCCCGGA
 42     ((((((...((((........)))).(((((.......))))).....(((((.......)))))..))))))...
exit    GGGAUGAUAGGGCGUGUGAUAGCCCAUAGCGAACCCCCCGCUGAGCUUGUGCGACGUUUGUGCACCCUGUCCCGCU
entry   GGGAAGAUAGGGCGUGUGAUAGCCCAUAGCGAACCCCCCGCUGAGCUUGUGCGACGUUUGUGCACCCUGUCCCGCU
 43     ((((((...((((........)))).(((((.......))))).....(((((.......))))).)).))))...
exit    GGGAAGAUAGGGCGUGUGAUAGCCCAUAGCGAACCCCCCGCUGAGCUUGUGCGACGUUUGUGCACCCUGUCCCGCU
entry   GGGCAGAUAGGGCGUGUGAUAGCCCAUAGCGAACCCCCCGCUGAGCUUGUGCGACGUUUGUGCACCCUGUCCCGCU
 44    ((((((...((((........)))).(((((.......))))).....(((((.......))))).))))))....
```

Transition inducing point mutations

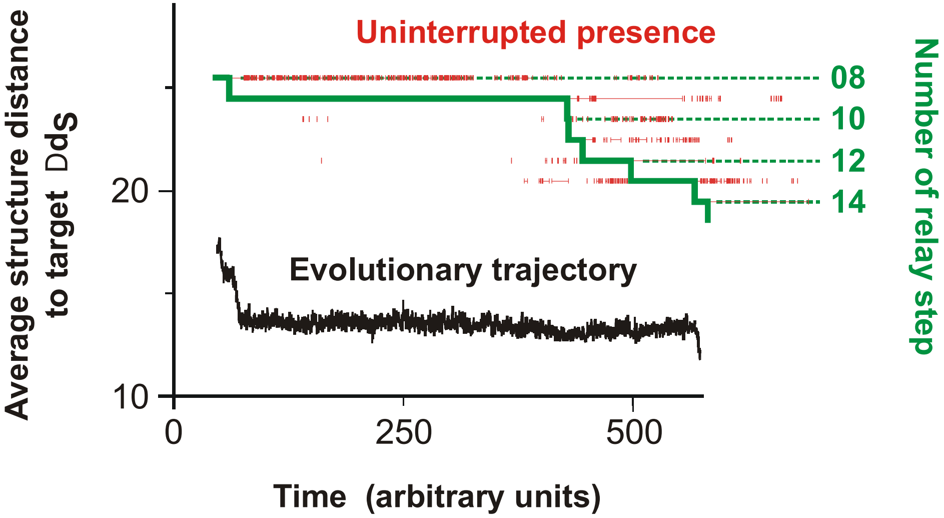
Neutral point mutations

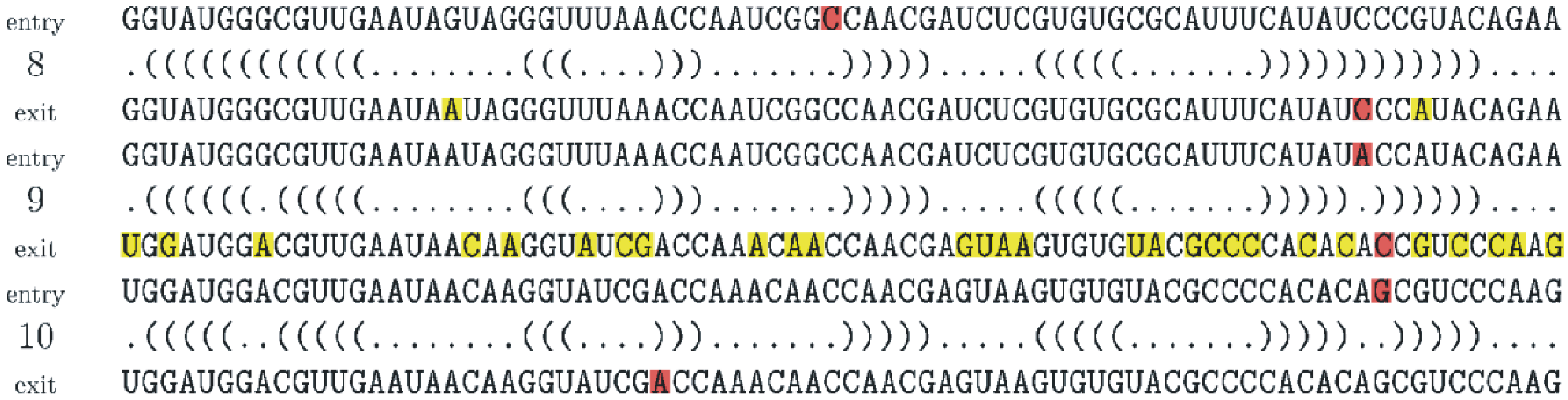
Change in RNA sequences during the final five relay steps 39 Á 44



In silico optimization in the flow reactor: Trajectory and relay steps

28 neutral point mutations during a long quasi-stationary epoch



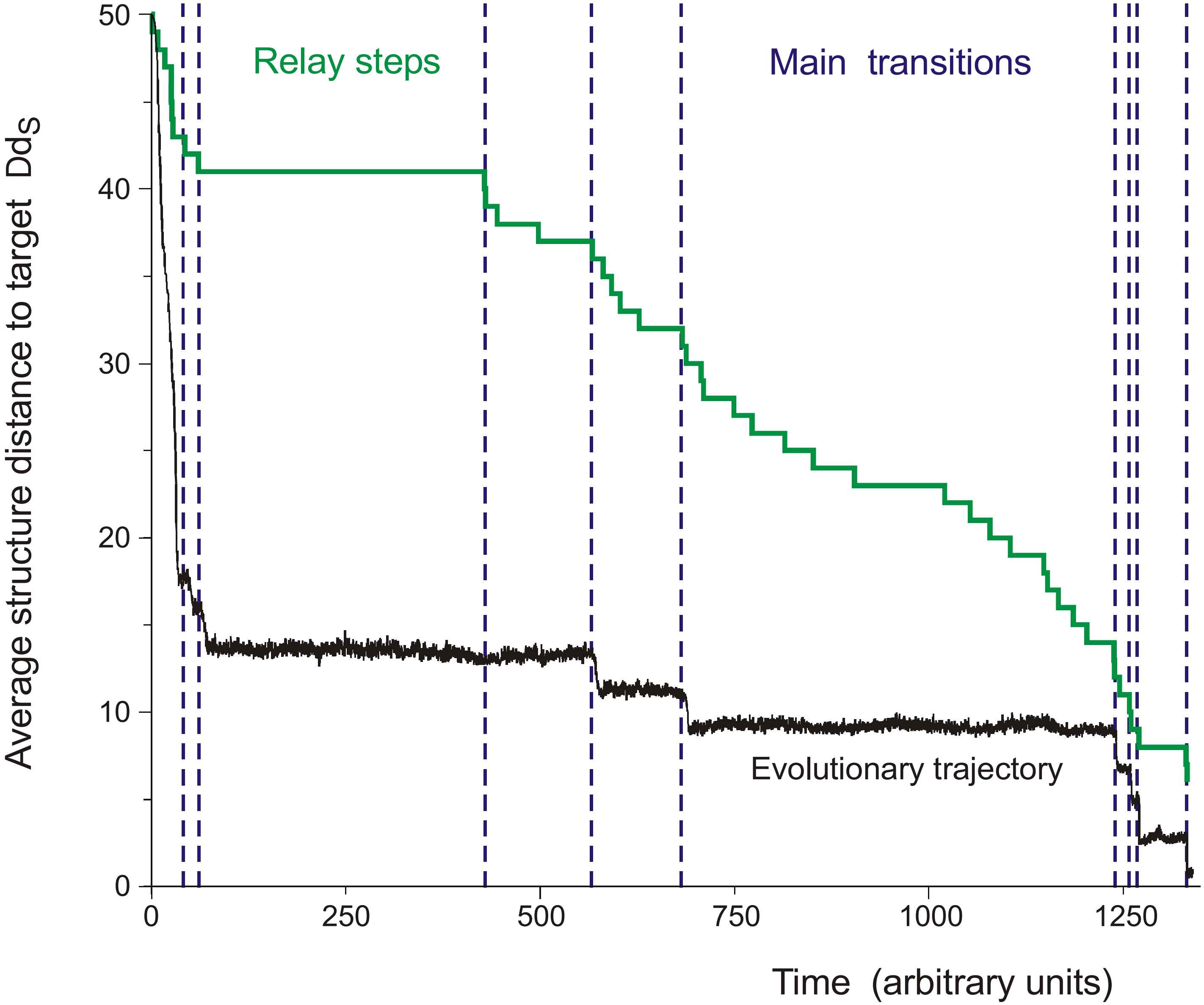


```
entry  GGUAUGGGCGUUGAAUAGUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCGUACAGAA
8      .((((((((((((........(((....))).......))))).....(((((.......))))))))))))....
exit   GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCAUACAGAA
entry  GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUACCAUACAGAA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

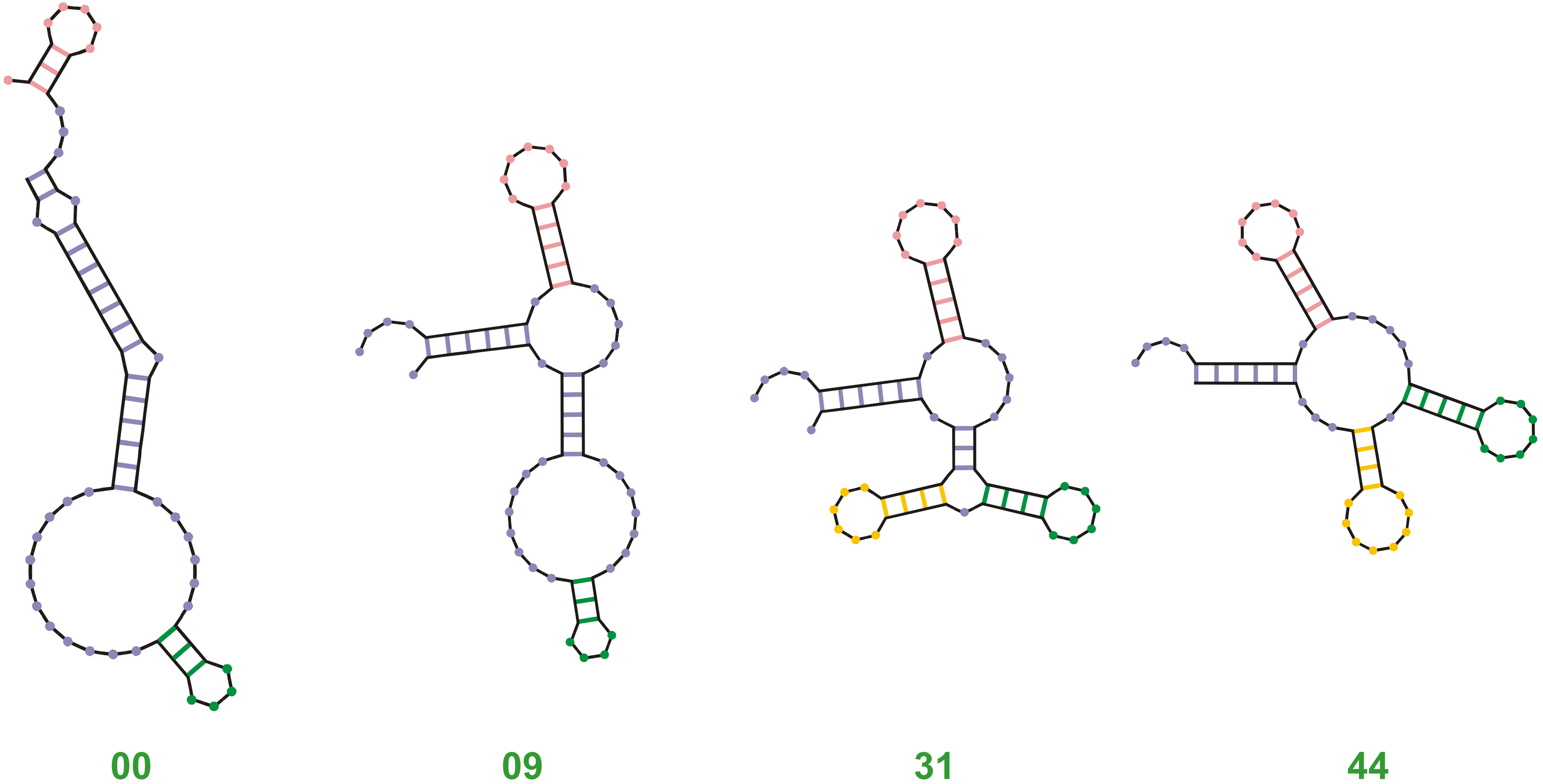
Transition inducing point mutations

Neutral point mutations

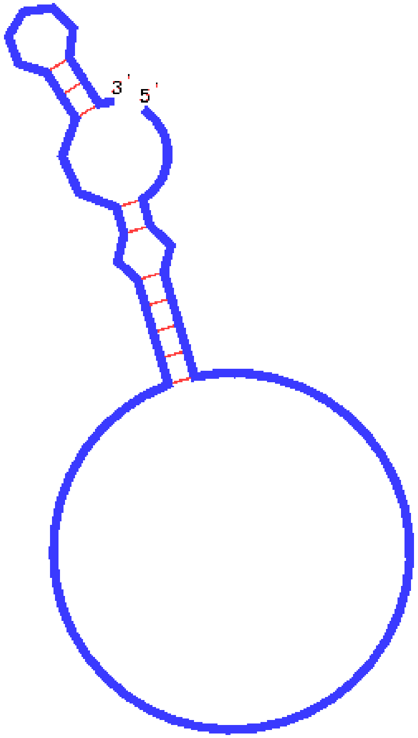
Neutral genotype evolution during phenotypic stasis



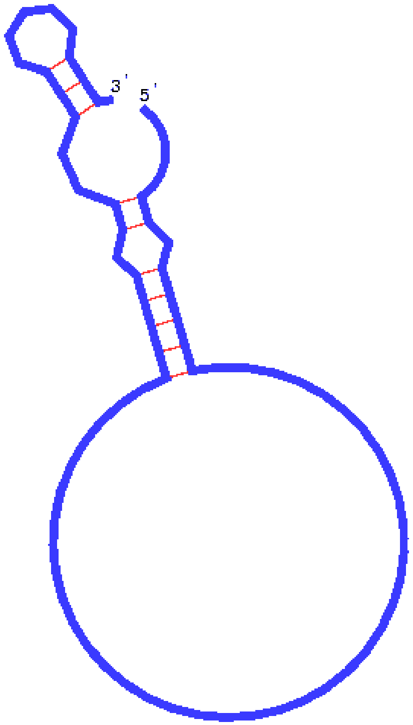
In silico optimization in the flow reactor: Main transitions



Three important steps in the formation of the tRNA clover leaf from a randomly chosen initial structure corresponding to three **main transitions**.

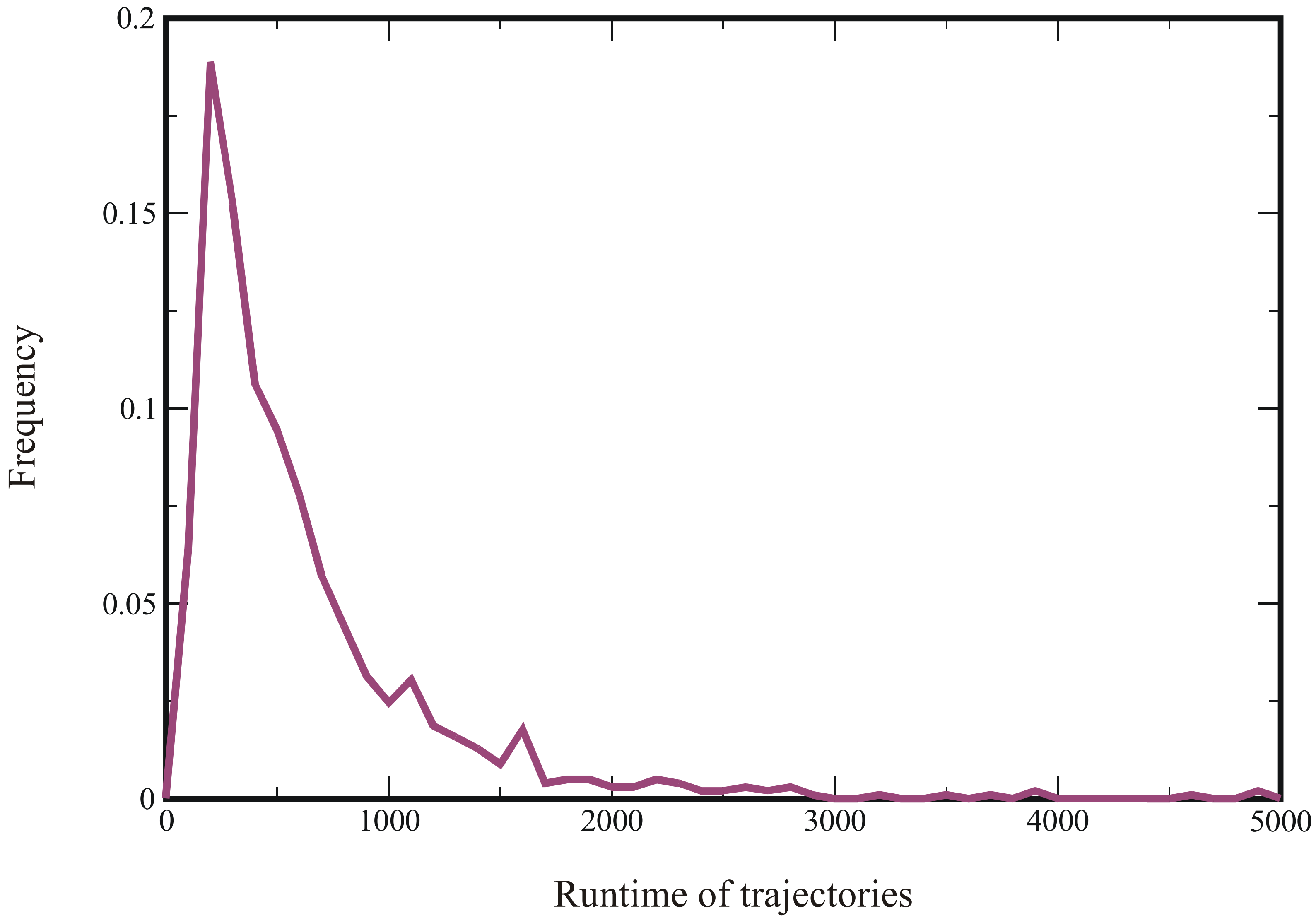


AUGC



GC

Movies of optimization trajectories over the **AUGC** and the **GC** alphabet



Statistics of the lengths of trajectories from initial structure to target (**AUGC**-sequences)

Alphabet	Runtime	Transitions	Main transitions	No. of runs
AUGC	385.6	22.5	12.6	1017
GUC	448.9	30.5	16.5	611
GC	2188.3	40.0	20.6	107

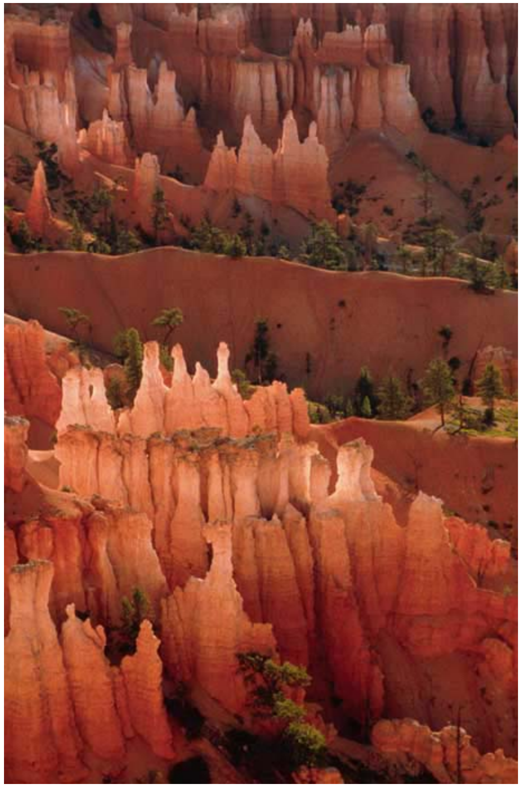
Statistics of trajectories and relay series (mean values of log-normal distributions)

Massif Central

Mount Fuji

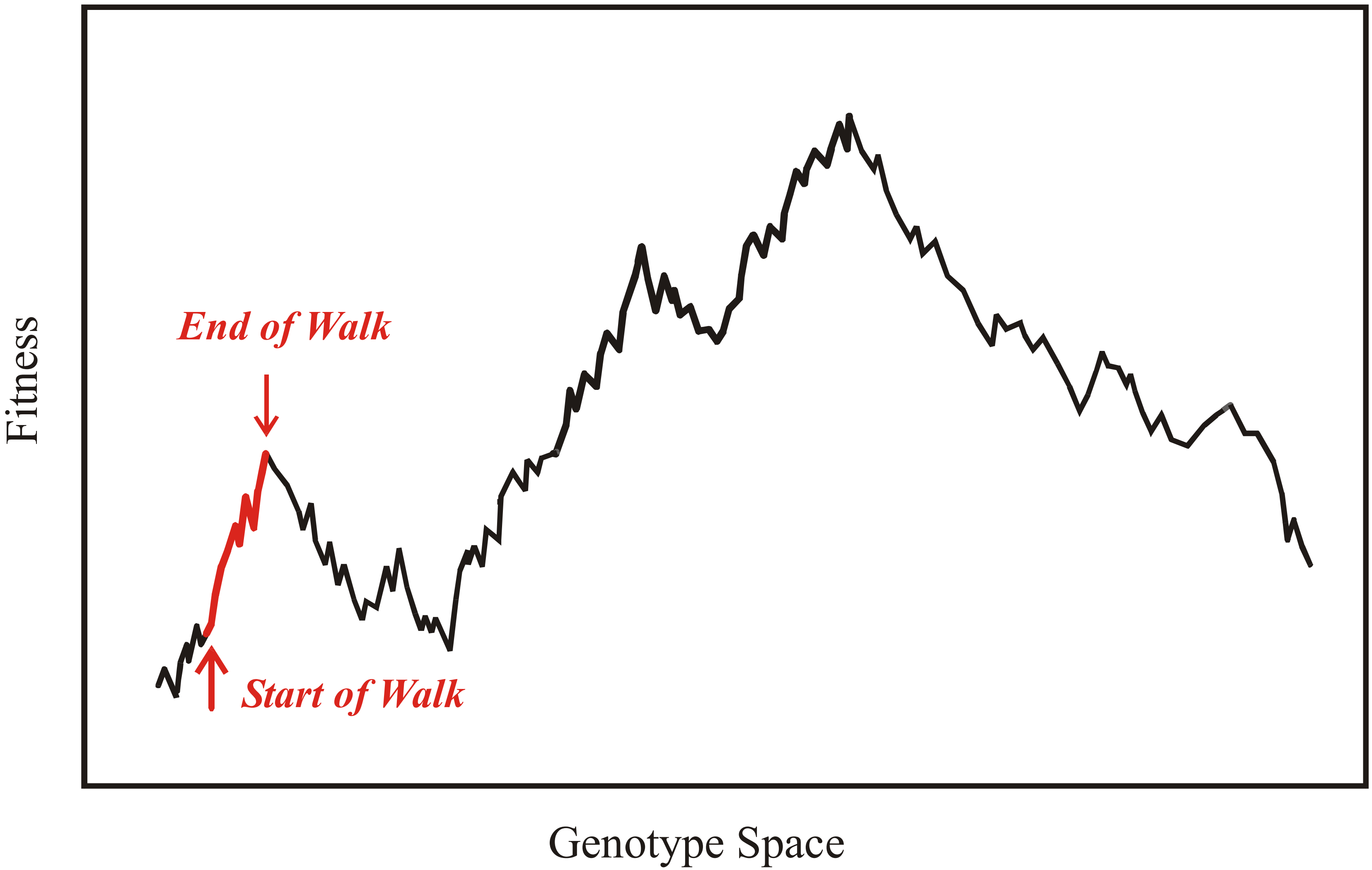
Examples of smooth landscapes on Earth

Dolomites

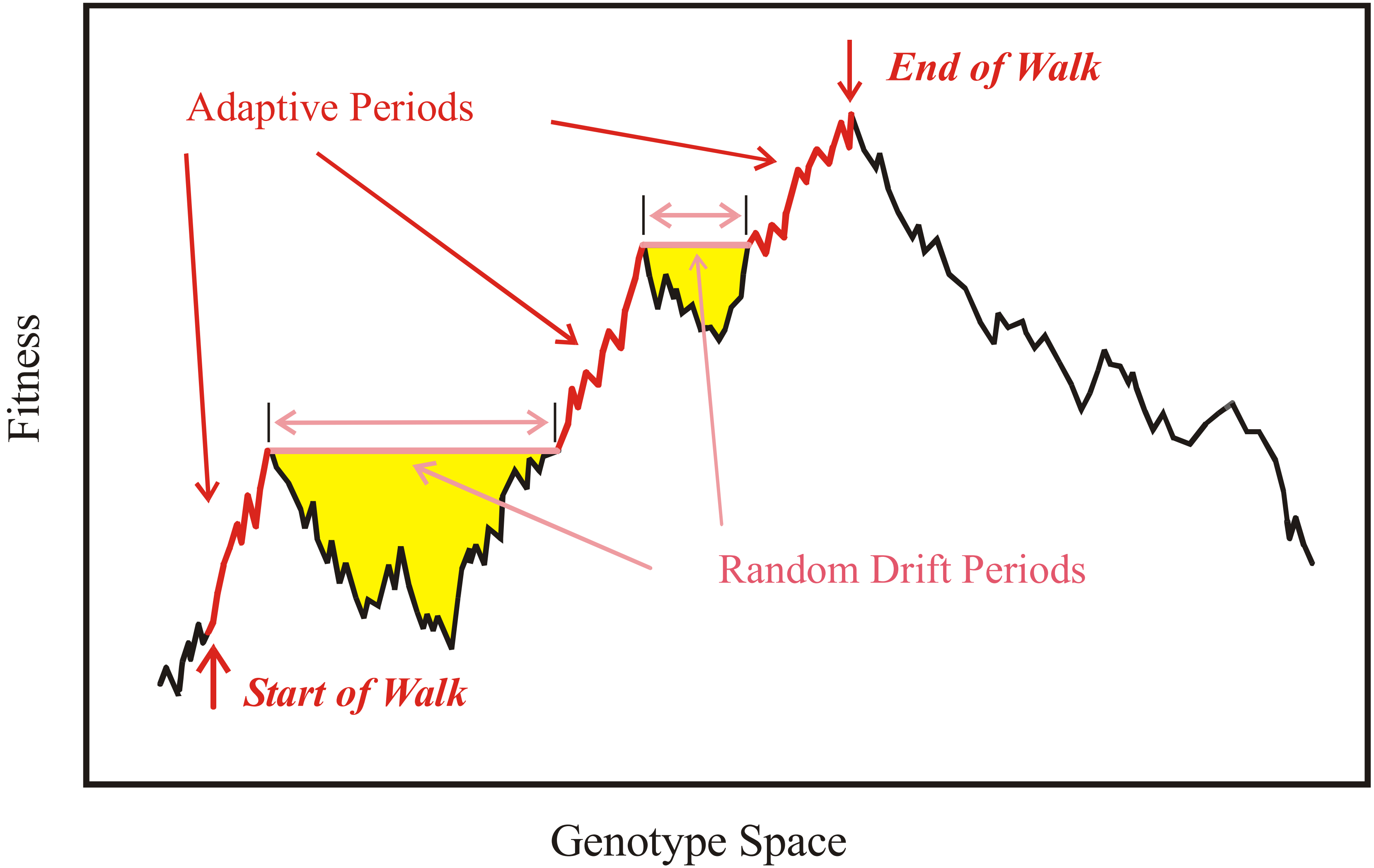


Bryce Canyon

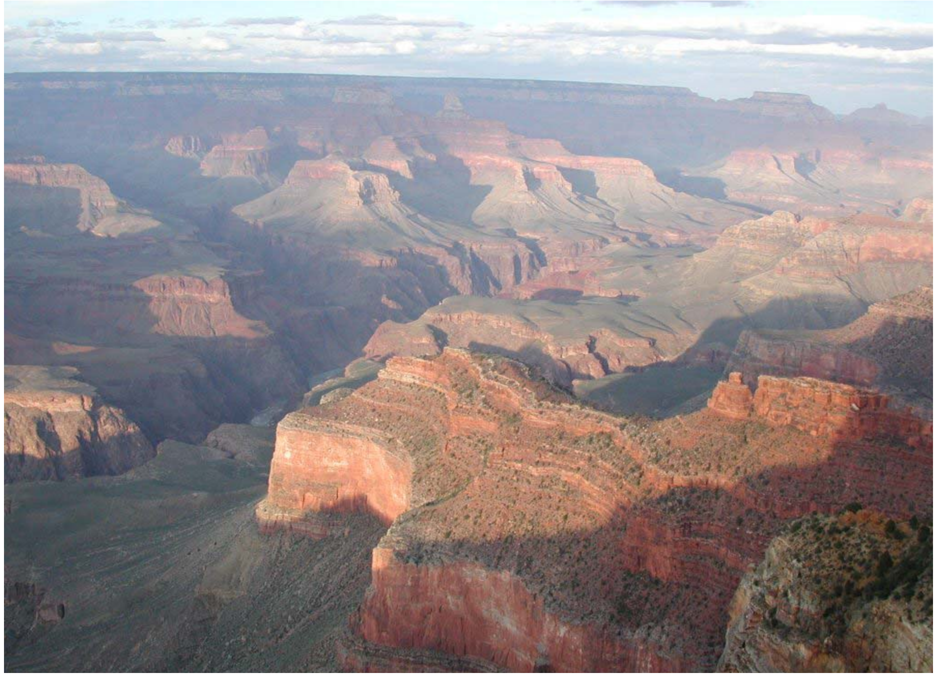
Examples of rugged landscapes on Earth



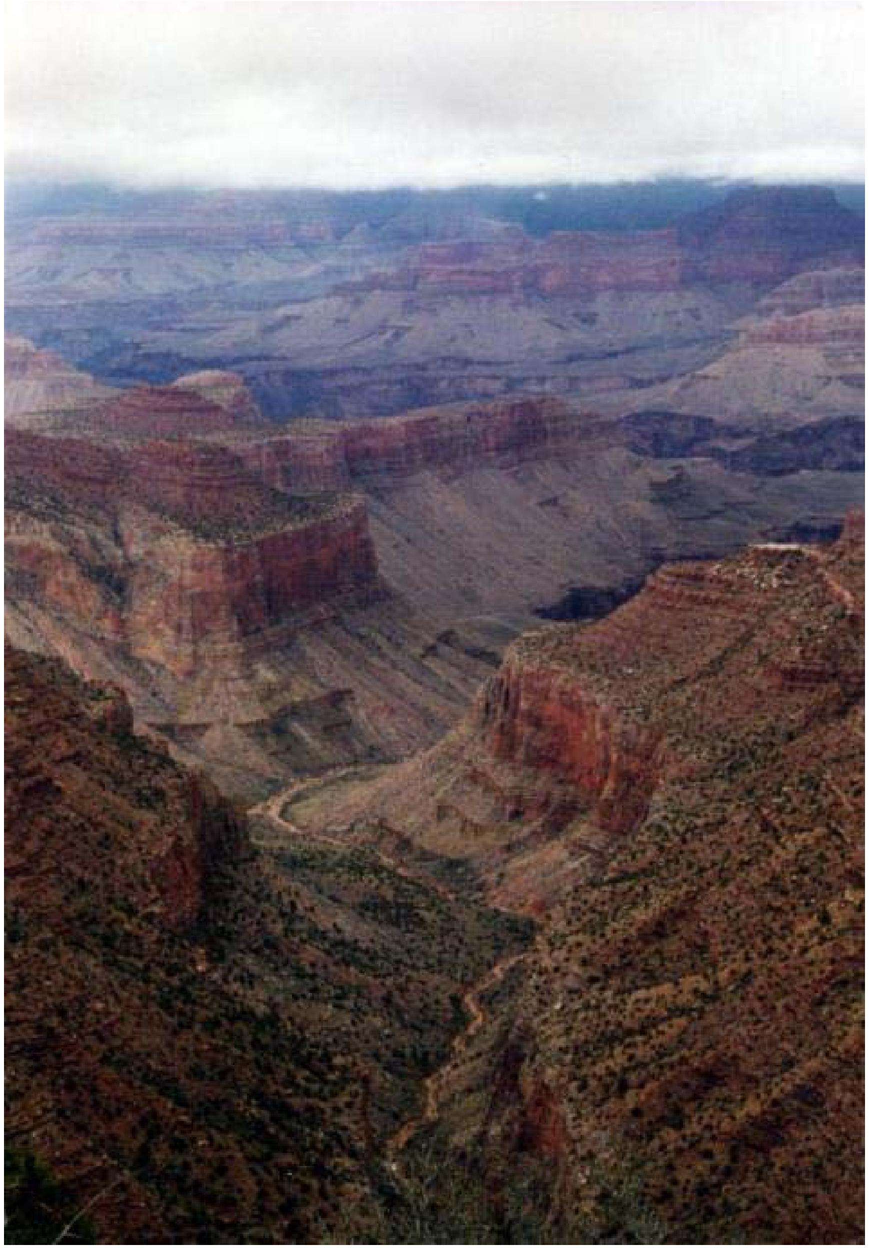
Evolutionary optimization in absence of neutral paths in sequence space



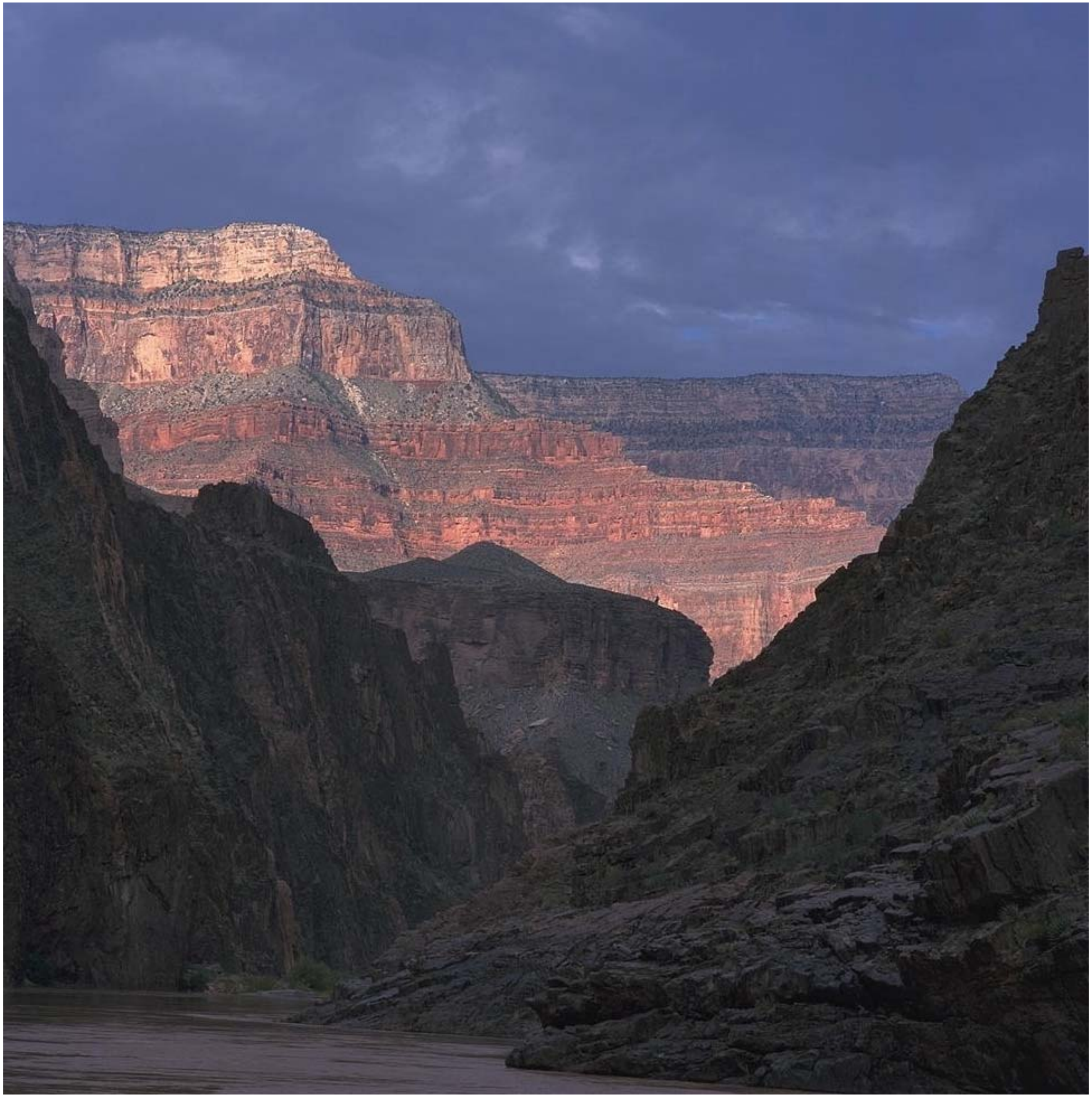
Evolutionary optimization including neutral paths in sequence space



Grand Canyon



Example of a landscape on Earth with ‘neutral’ ridges and plateaus



Neutral ridges and plateus

1. Autocatalytic chemical reactions in the flow reactor

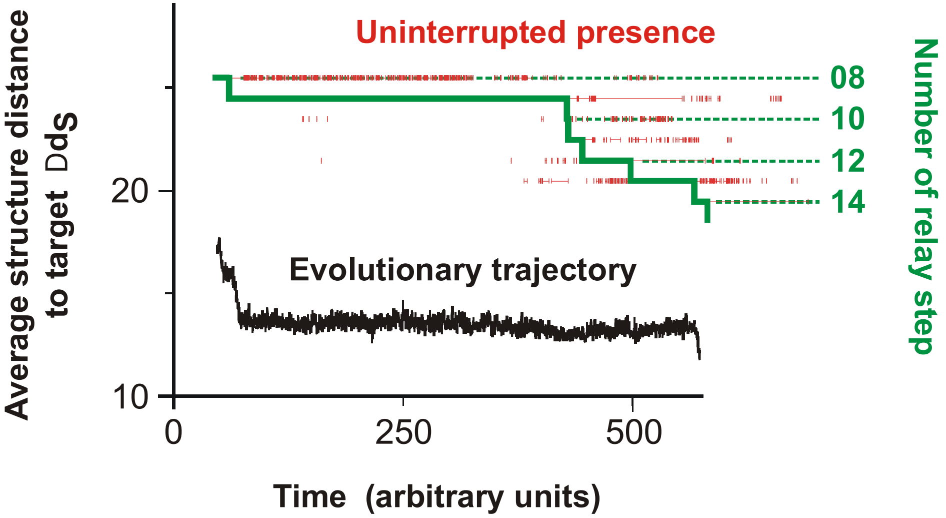
2. Replication, mutation, selection and Shannon information

3. Evolution *in silico* and optimization of RNA structures

4. **Random walks and ‚ensemble learning‘**

5. Sequence-structure maps, neutral networks, and intersections

28 neutral point mutations during a long quasi-stationary epoch

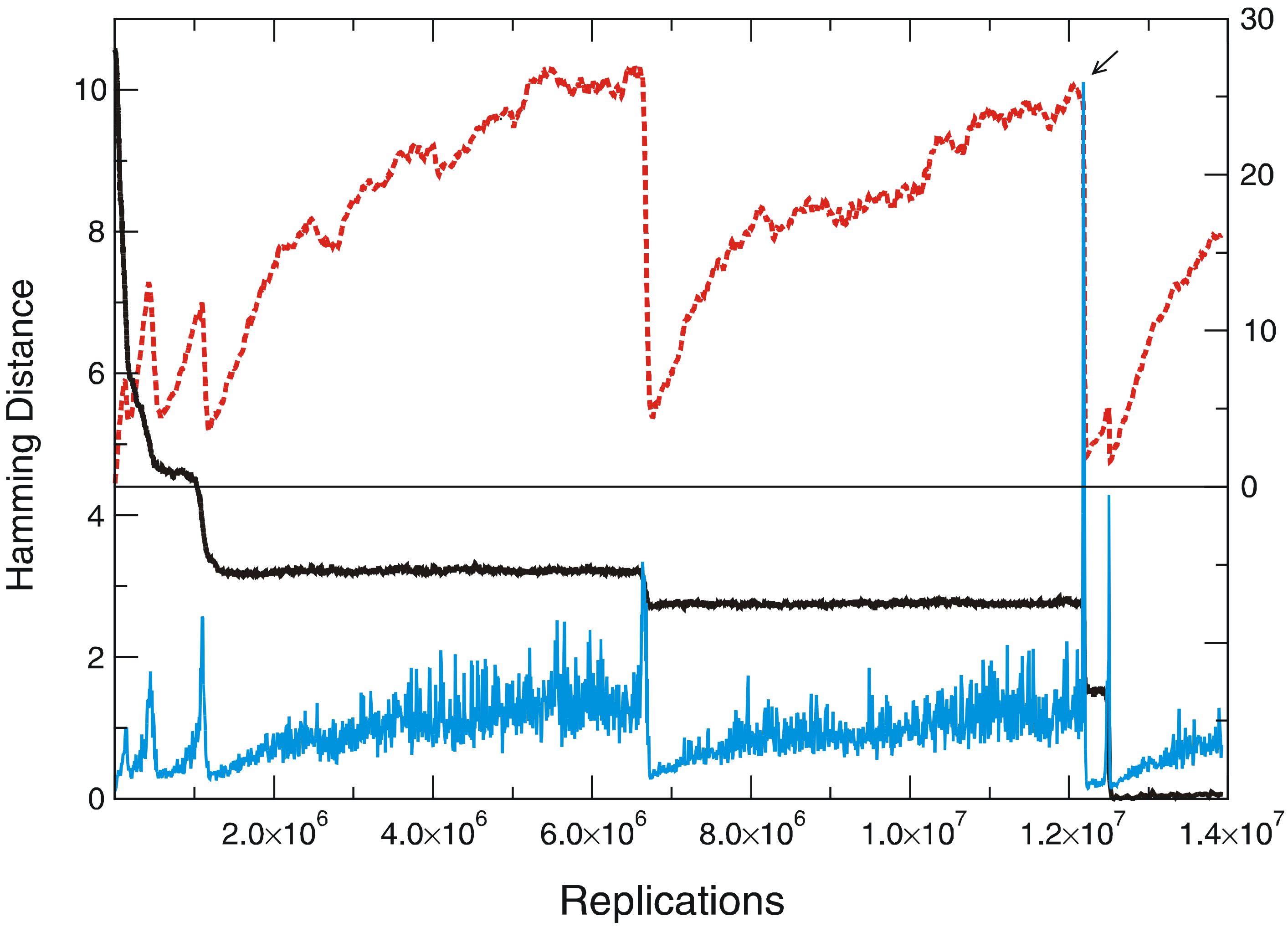


```
entry  GGUAUGGGCGUUGAAUAGUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCGUACAGAA
  8    .((((((((((((........(((....))).......))))).....(((((.......)))))))))))).....
exit   GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCAUACAGAA
entry  GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUACCAUACAGAA
  9    .((((((.(((((........(((....))).......))))).....(((((.......))))).))))))....
exit   UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACACCGUCCAAG
entry  UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCAAG
 10    .(((((..(((((........(((....))).......))))).....(((((.......)))))..)))))....
exit   UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCAAG
```

Transition inducing point mutations

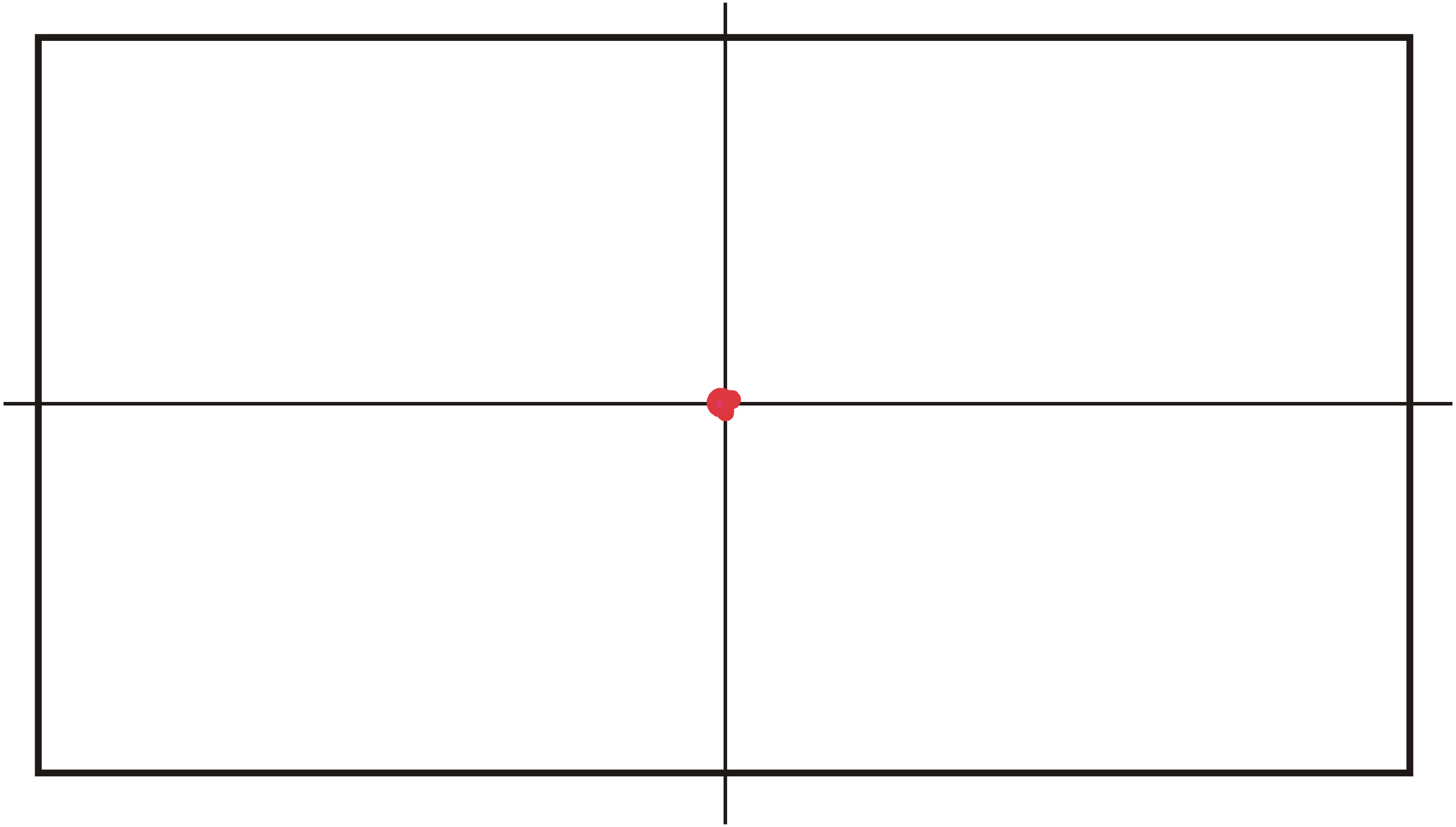
Neutral point mutations

Neutral genotype evolution during phenotypic stasis

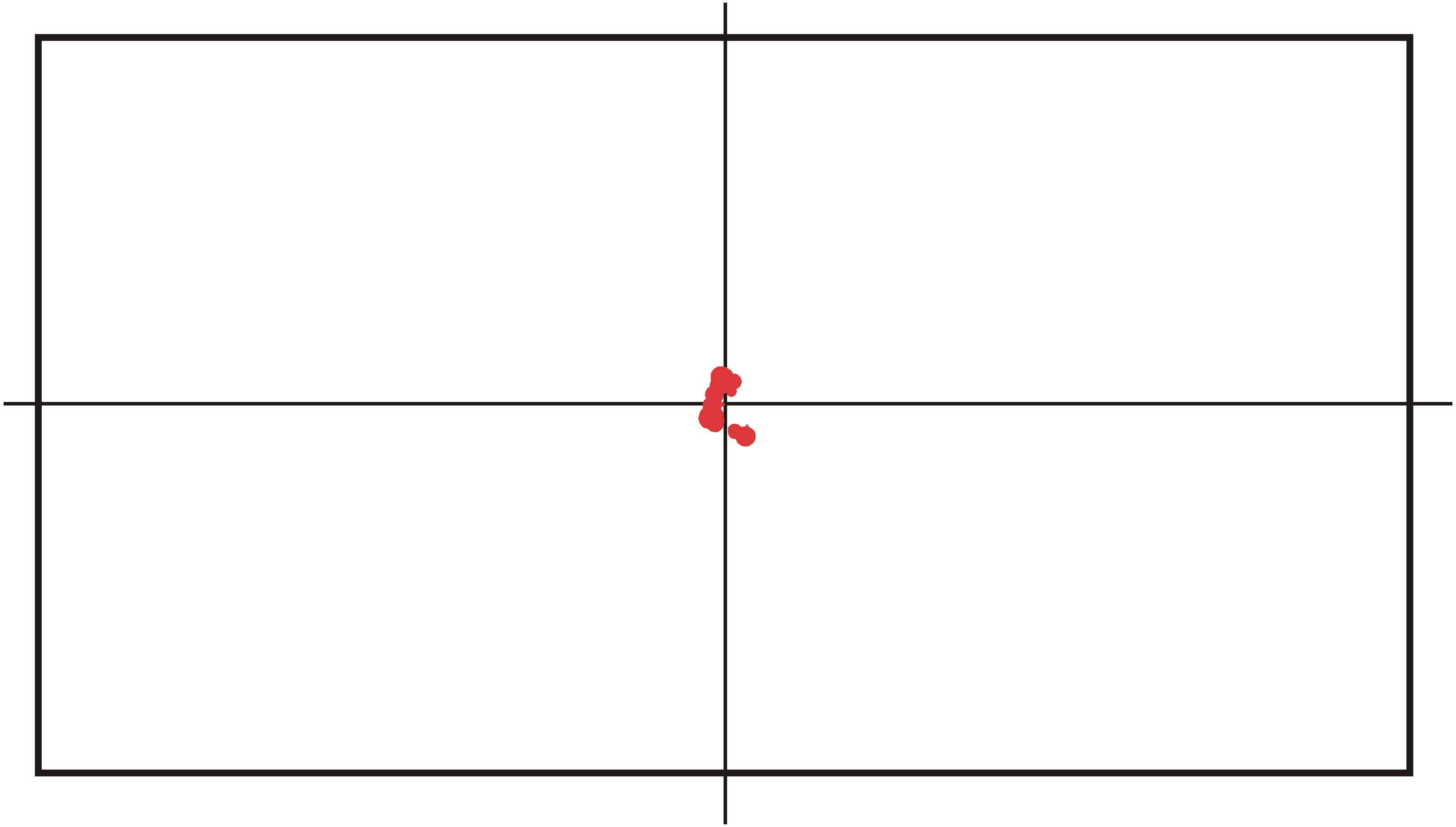


Variation in genotype space during optimization of phenotypes

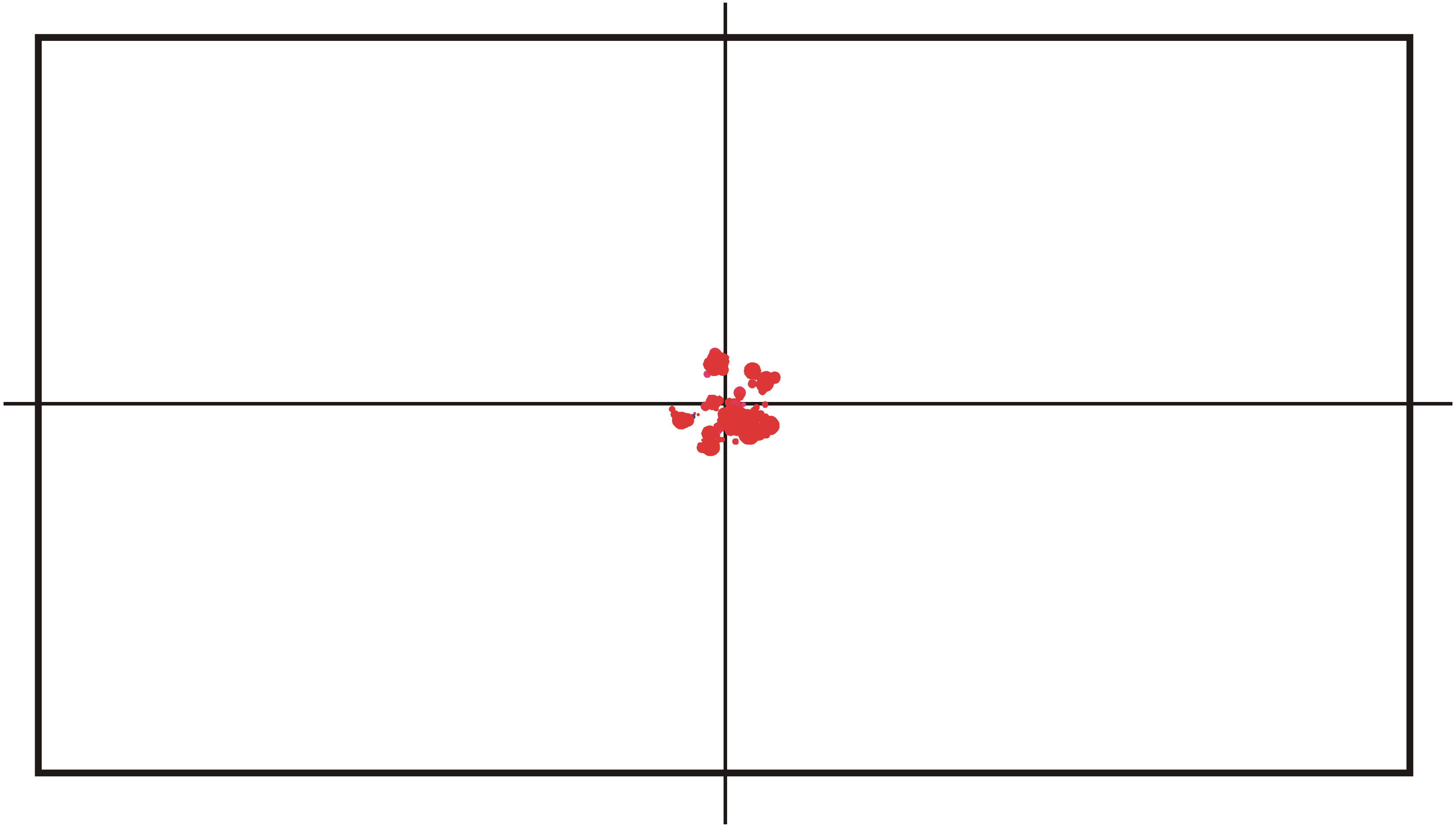
Mean Hamming distance within the population and **drift velocity of the population center** in sequence space.



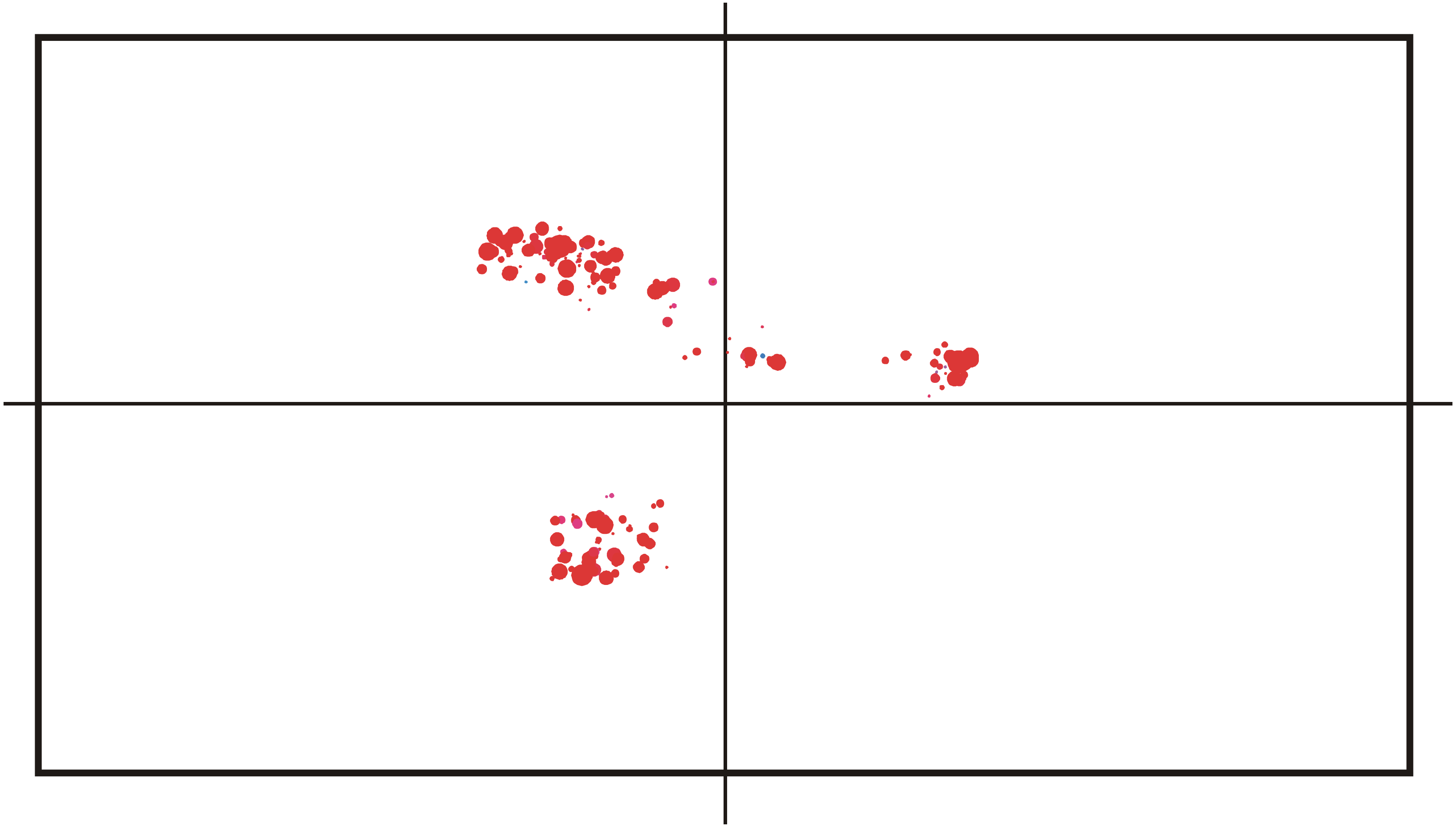
Spread of population in sequence space during a quasistationary epoch: t = 150



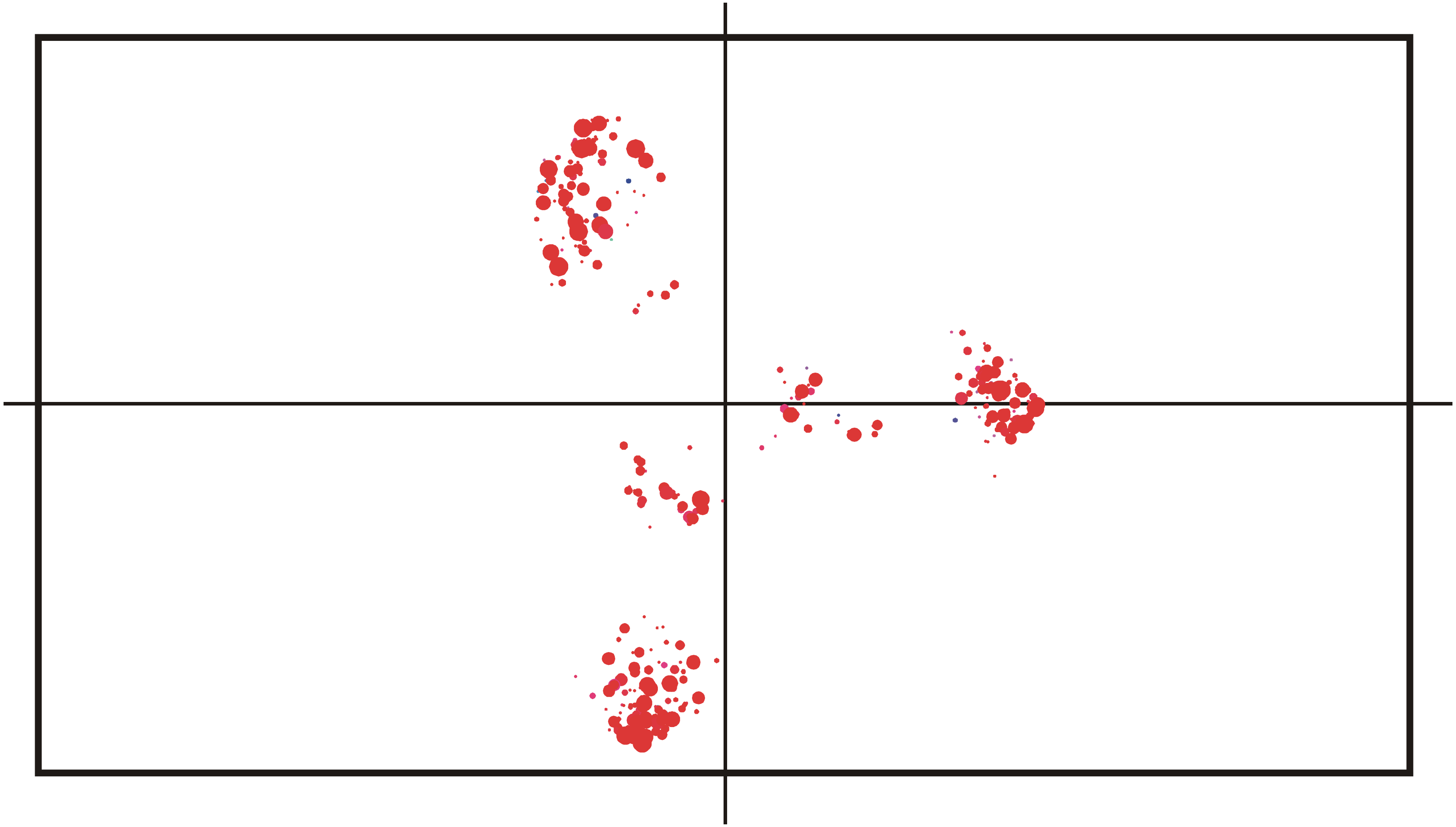
Spread of population in sequence space during a quasistationary epoch: t = 170



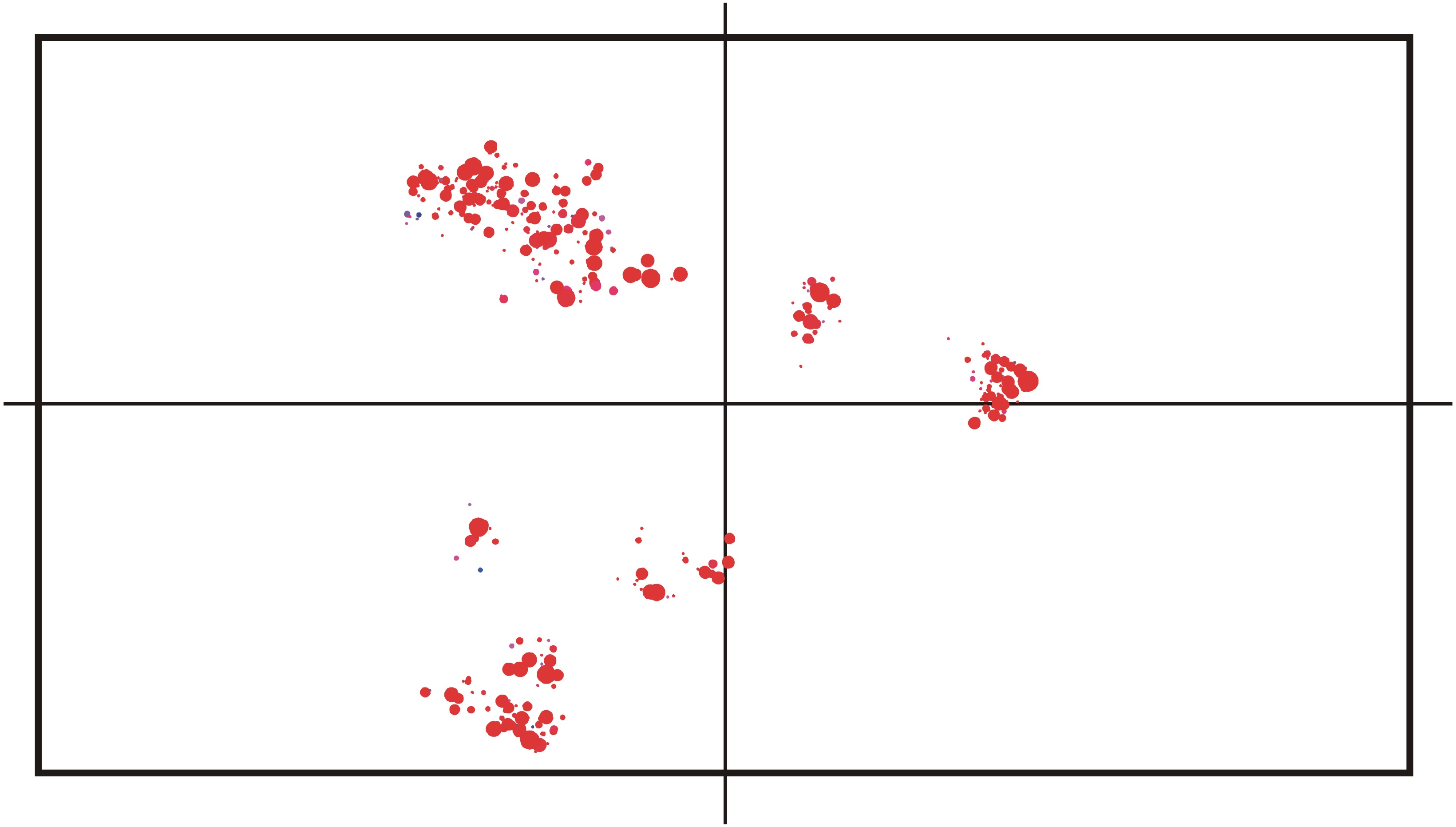
Spread of population in sequence space during a quasistationary epoch: t = 200



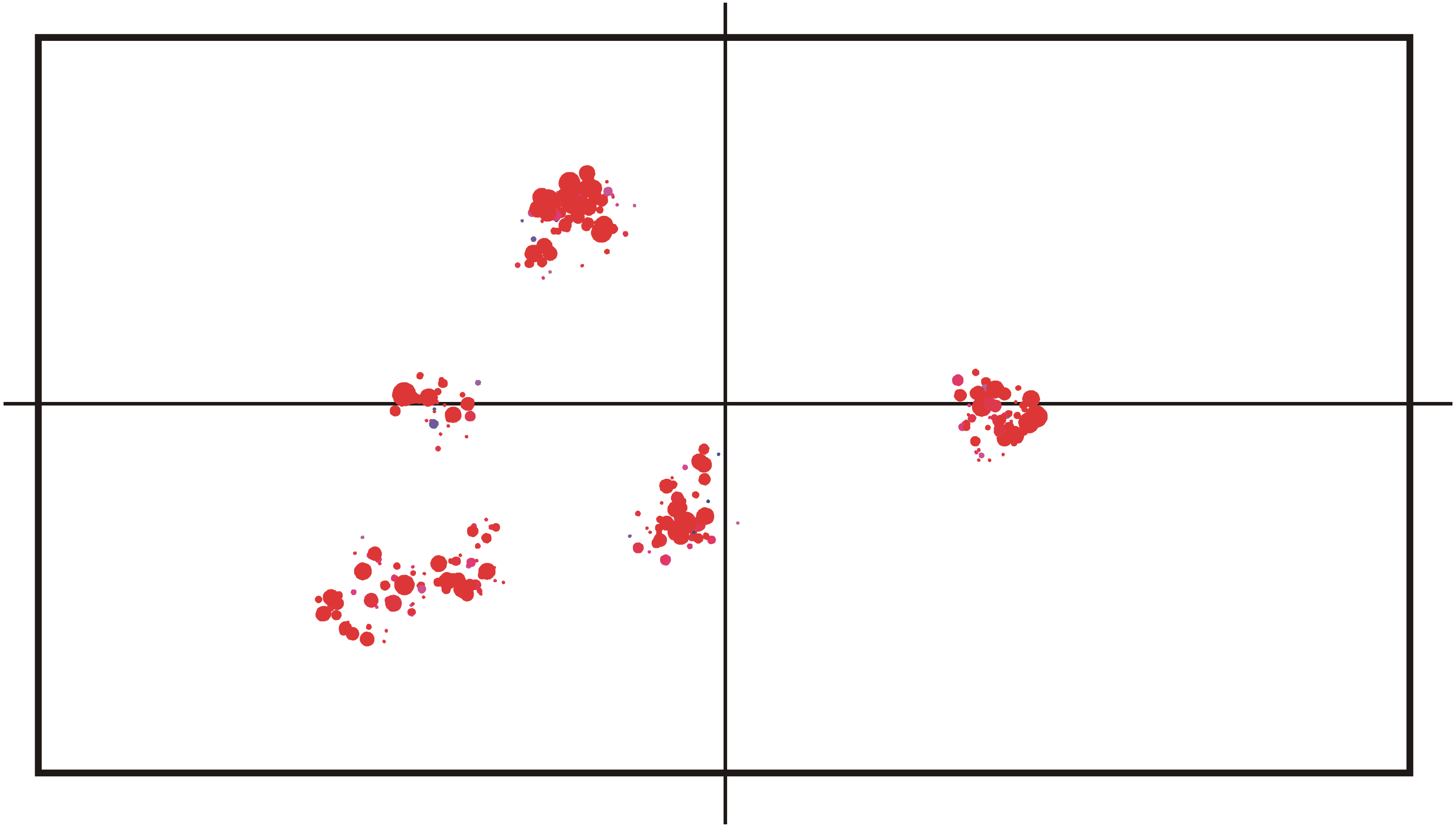
Spread of population in sequence space during a quasistationary epoch: t = 350



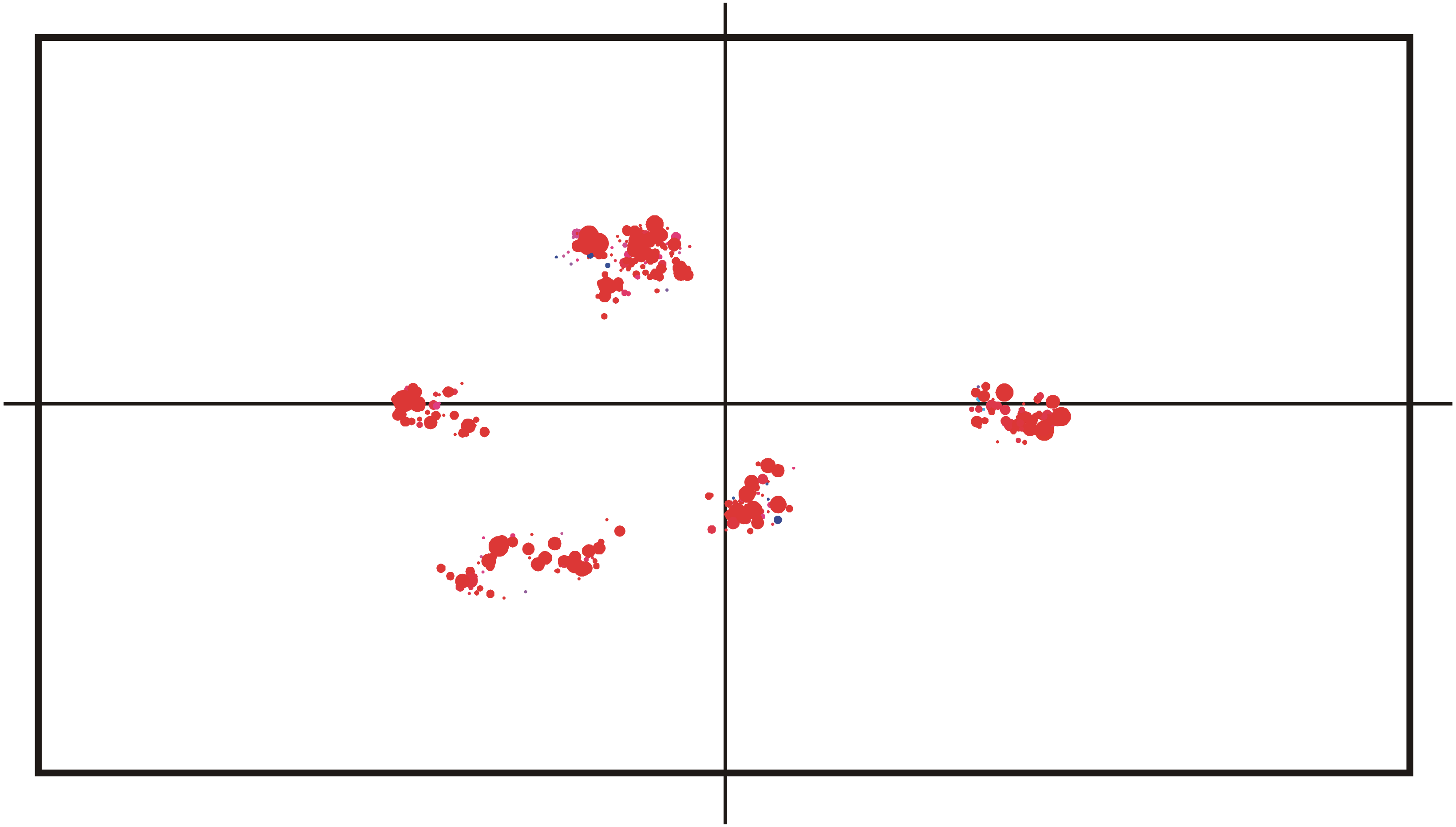
Spread of population in sequence space during a quasistationary epoch: t = 500



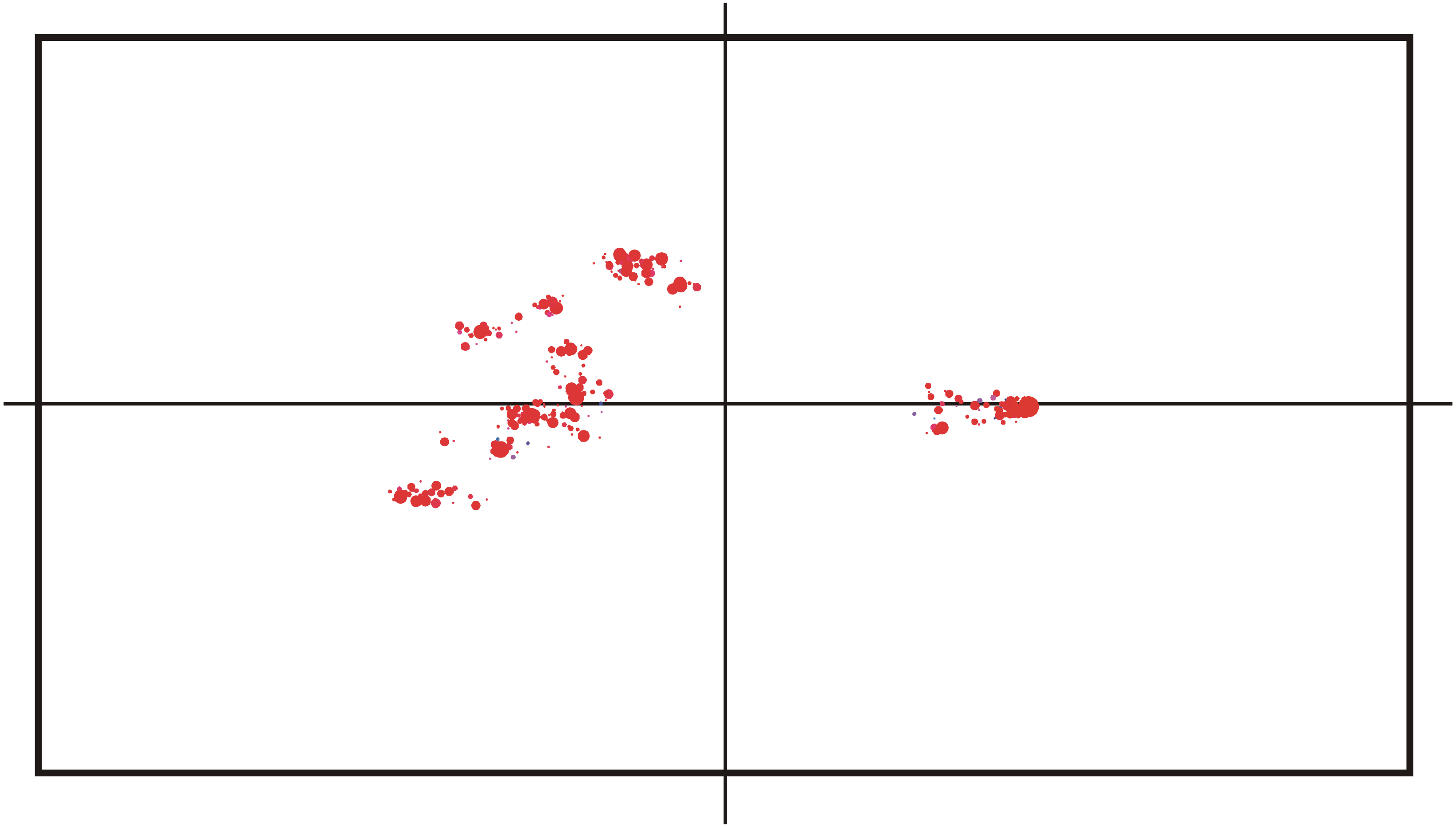
Spread of population in sequence space during a quasistationary epoch: t = 650



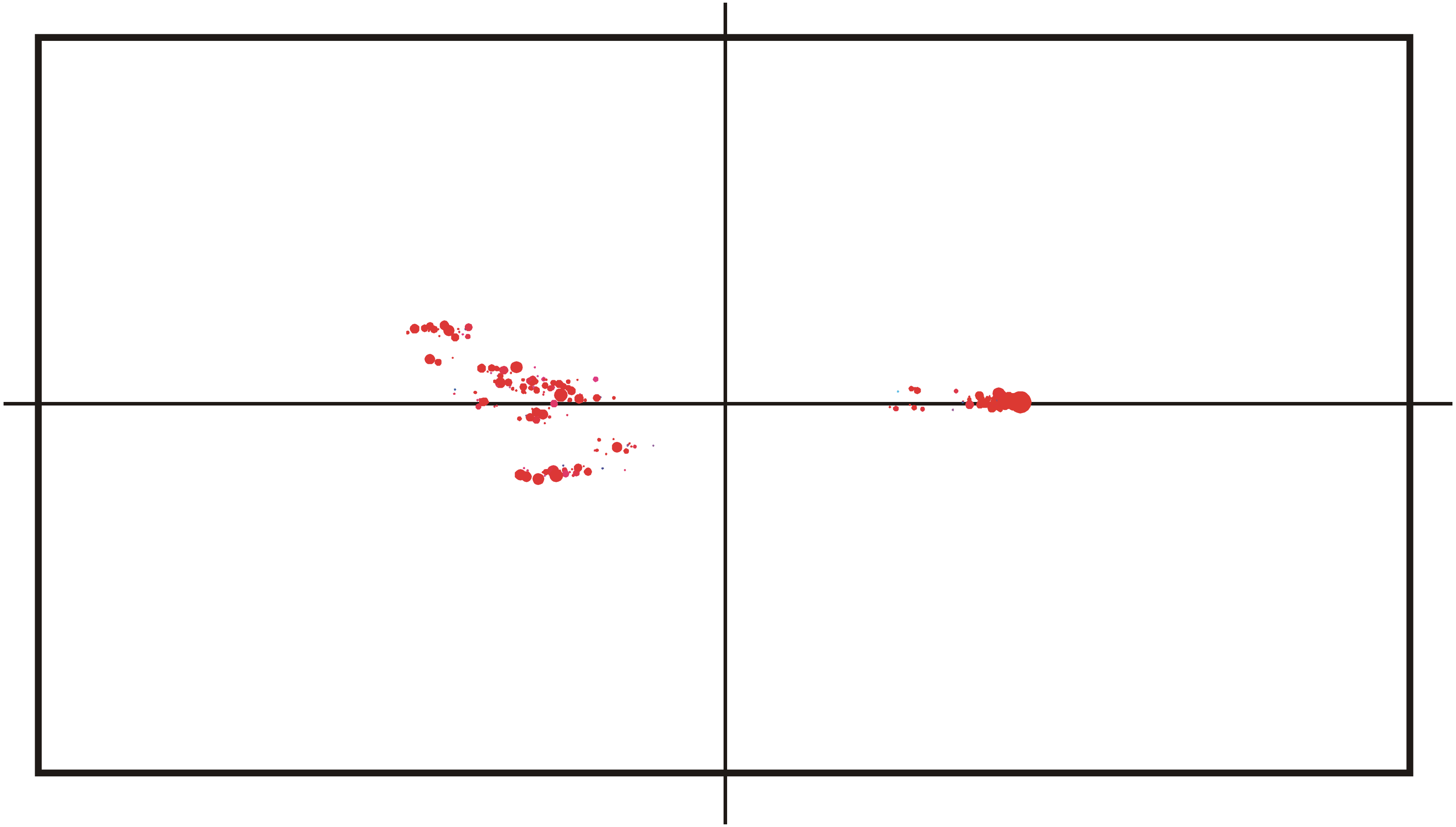
Spread of population in sequence space during a quasistationary epoch: t = 820



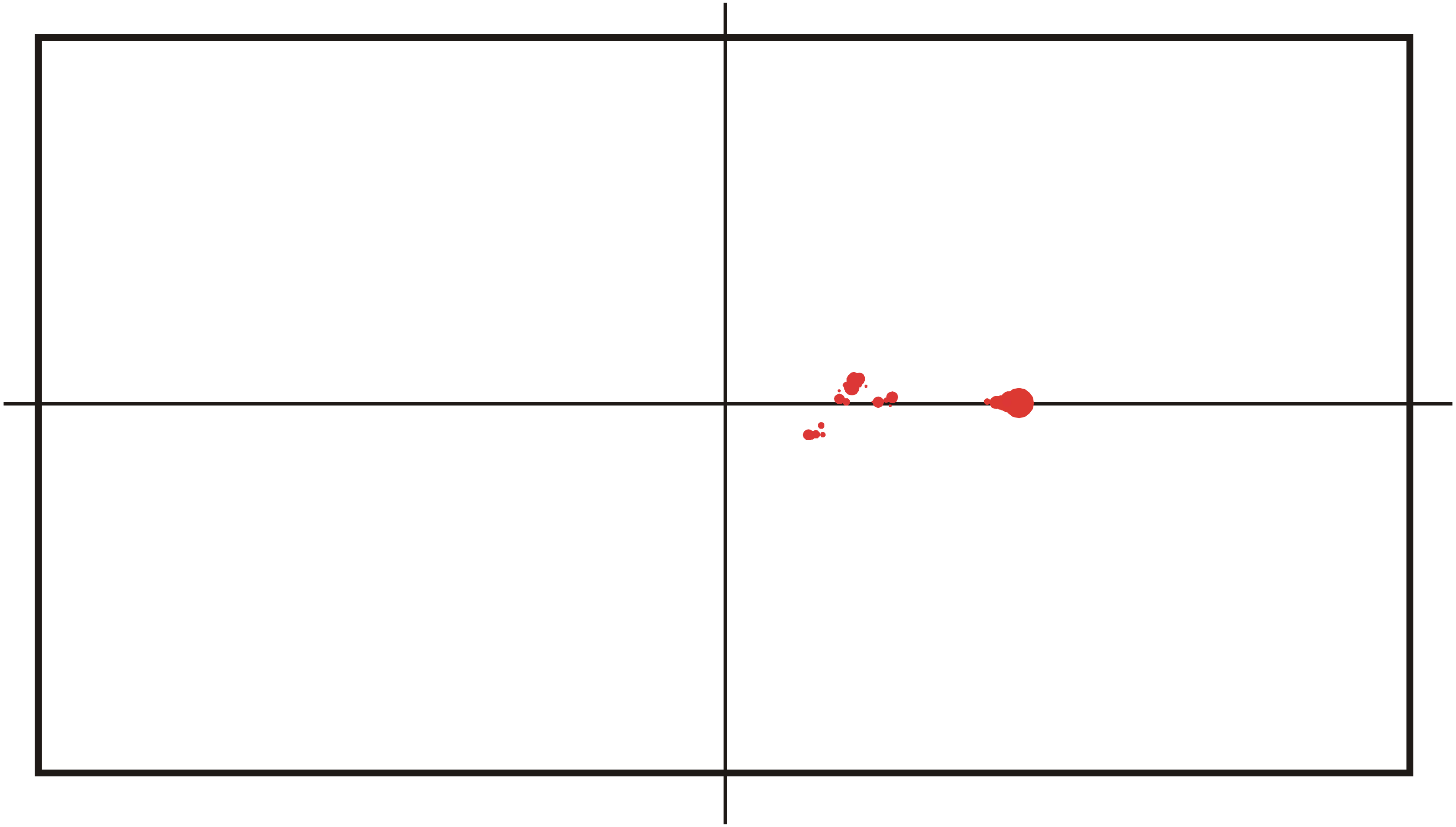
Spread of population in sequence space during a quasistationary epoch: t = 825



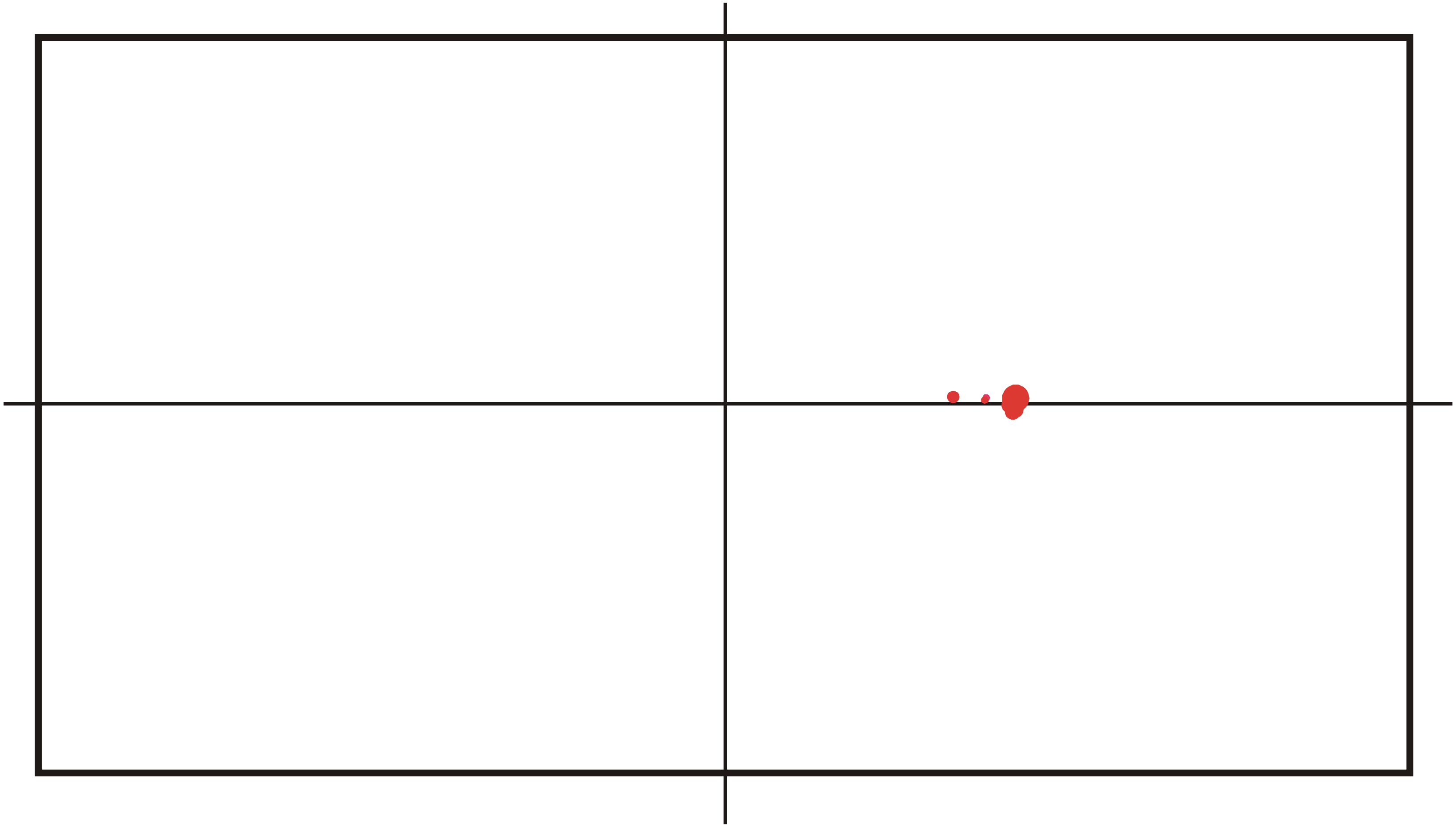
Spread of population in sequence space during a quasistationary epoch: t = 830



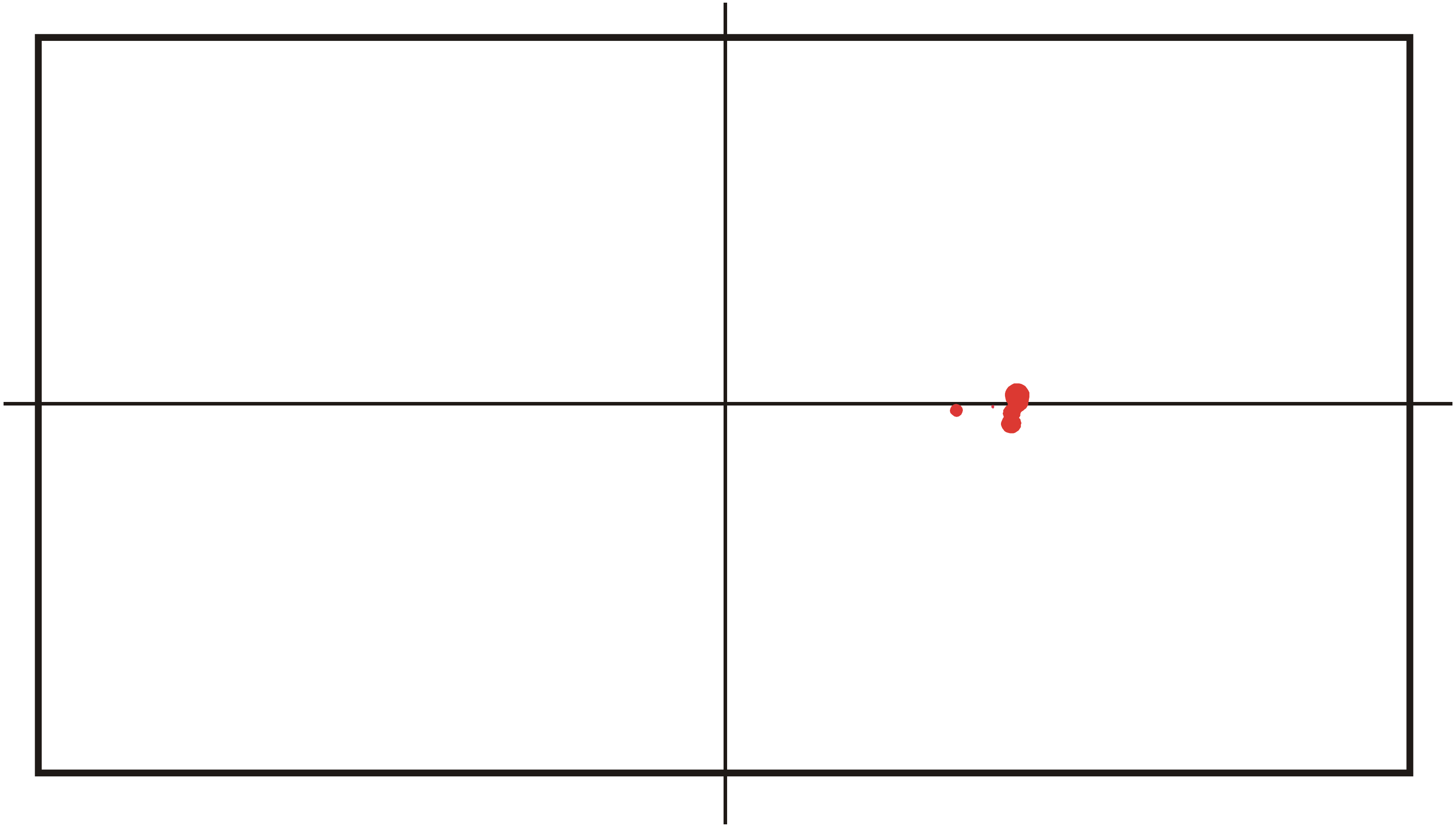
Spread of population in sequence space during a quasistationary epoch: t = 835



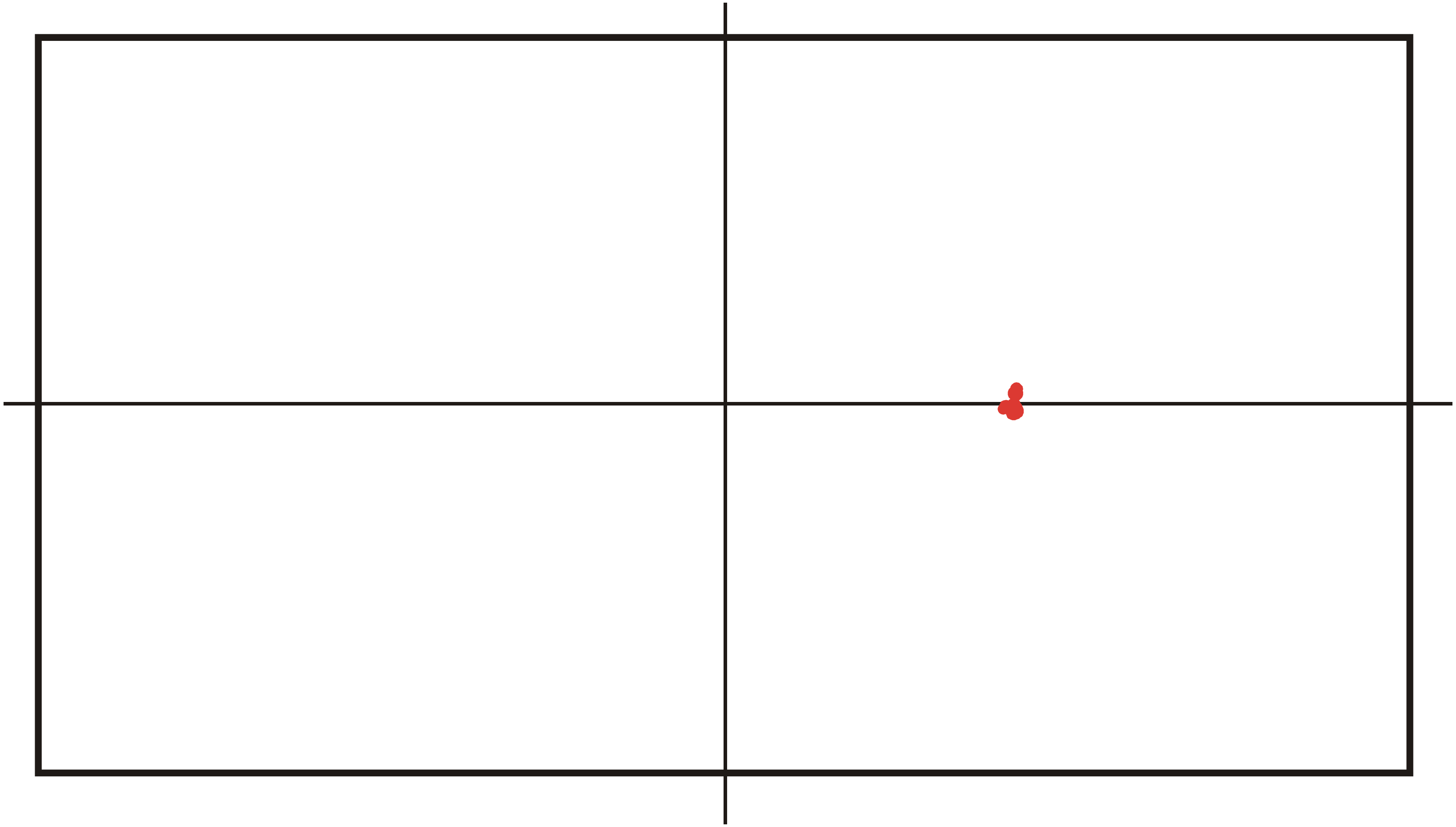
Spread of population in sequence space during a quasistationary epoch: t = 840



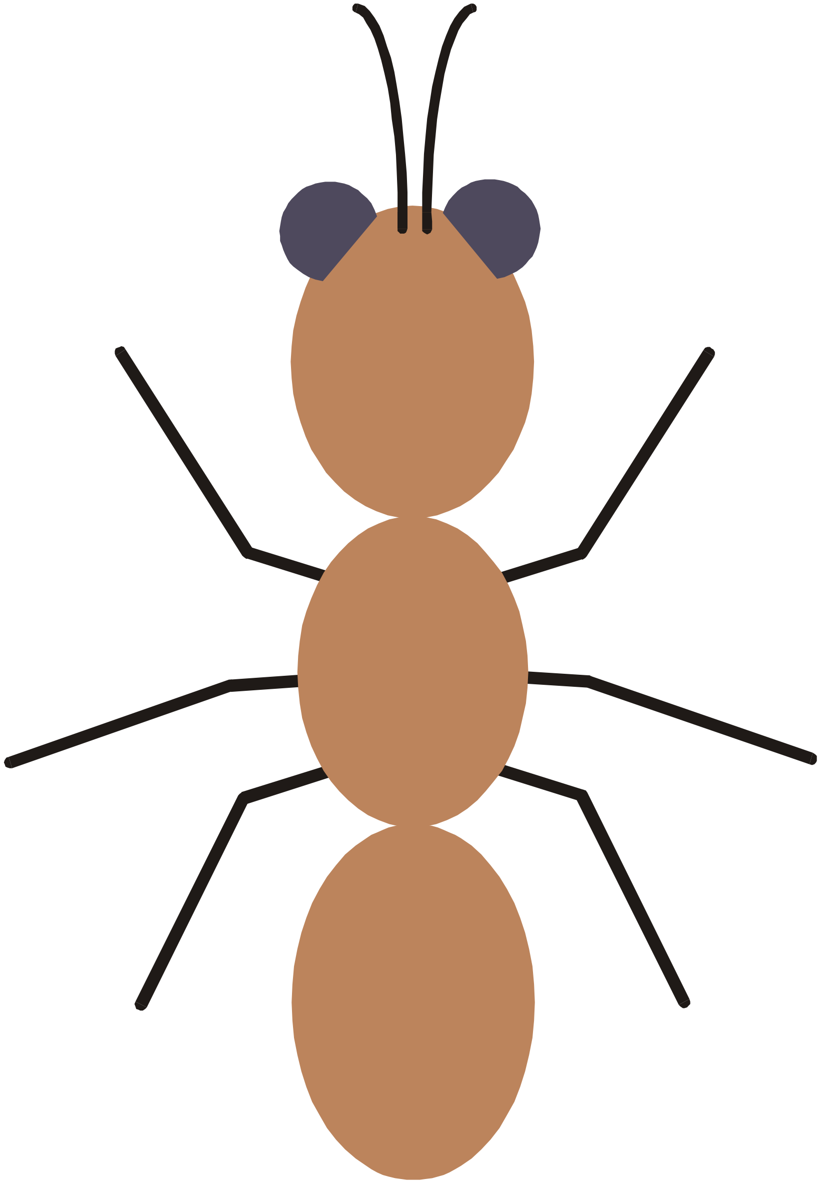
Spread of population in sequence space during a quasistationary epoch: t = 845



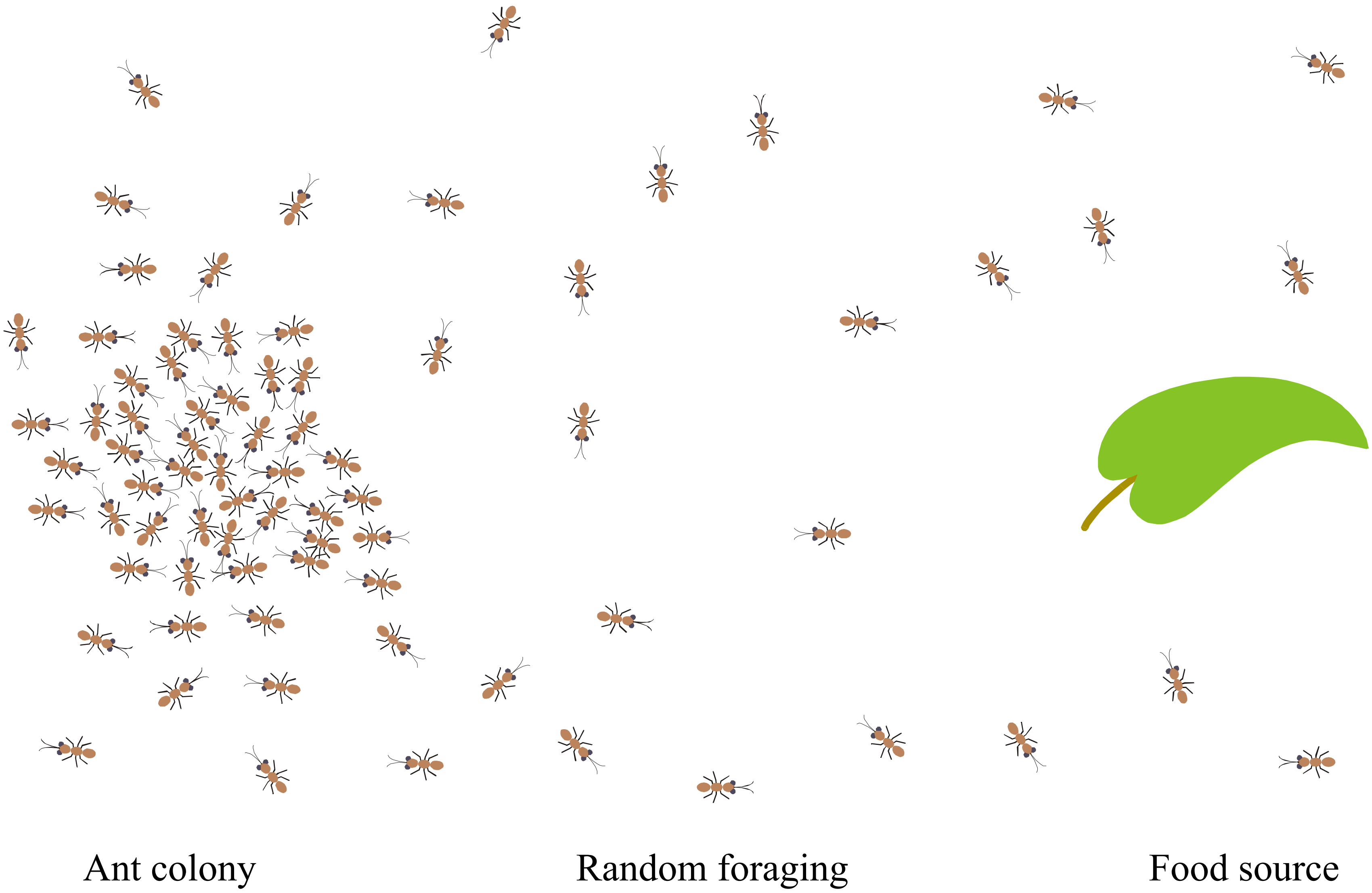
Spread of population in sequence space during a quasistationary epoch: t = 850



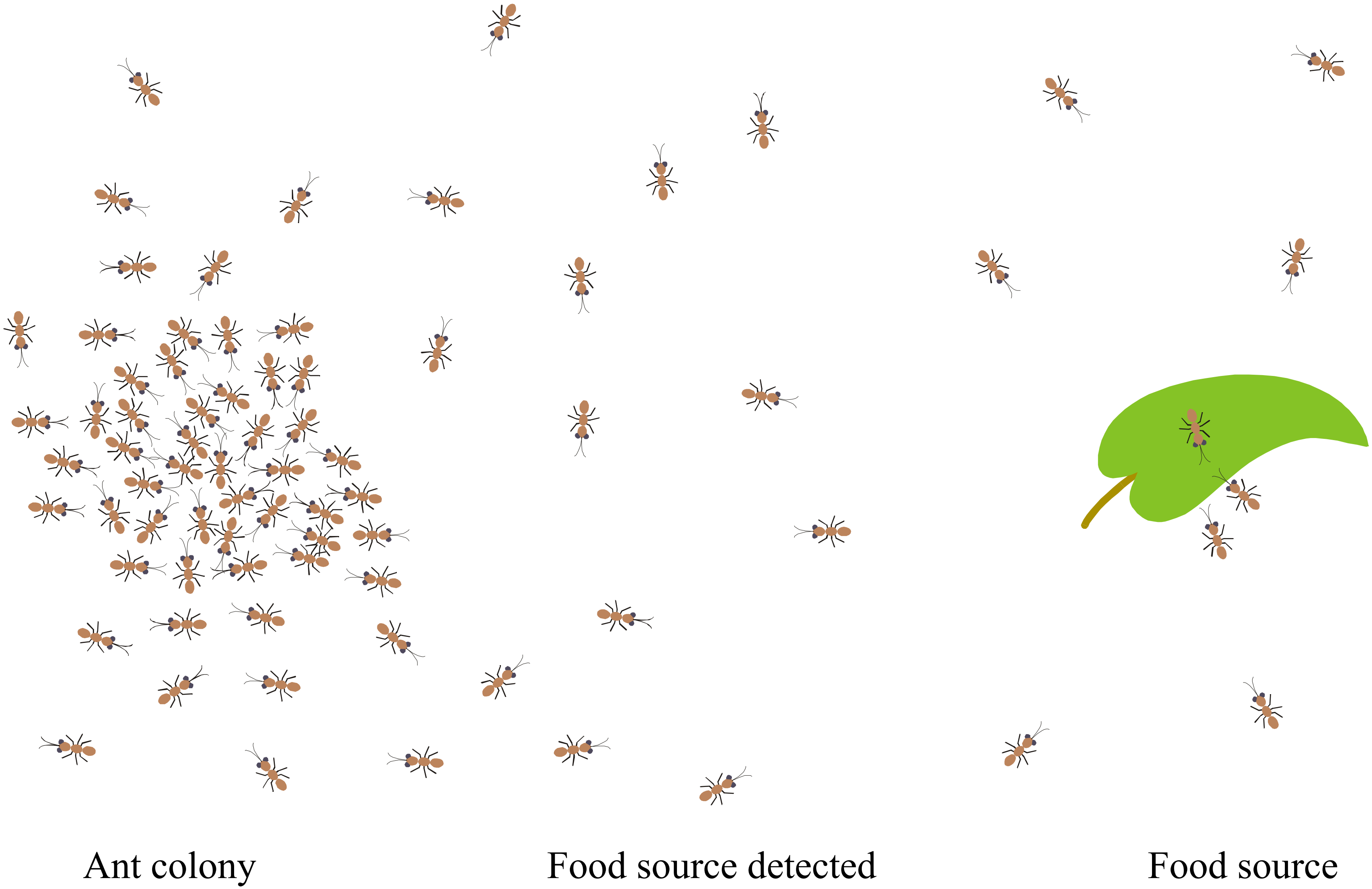
Spread of population in sequence space during a quasistationary epoch: t = 855



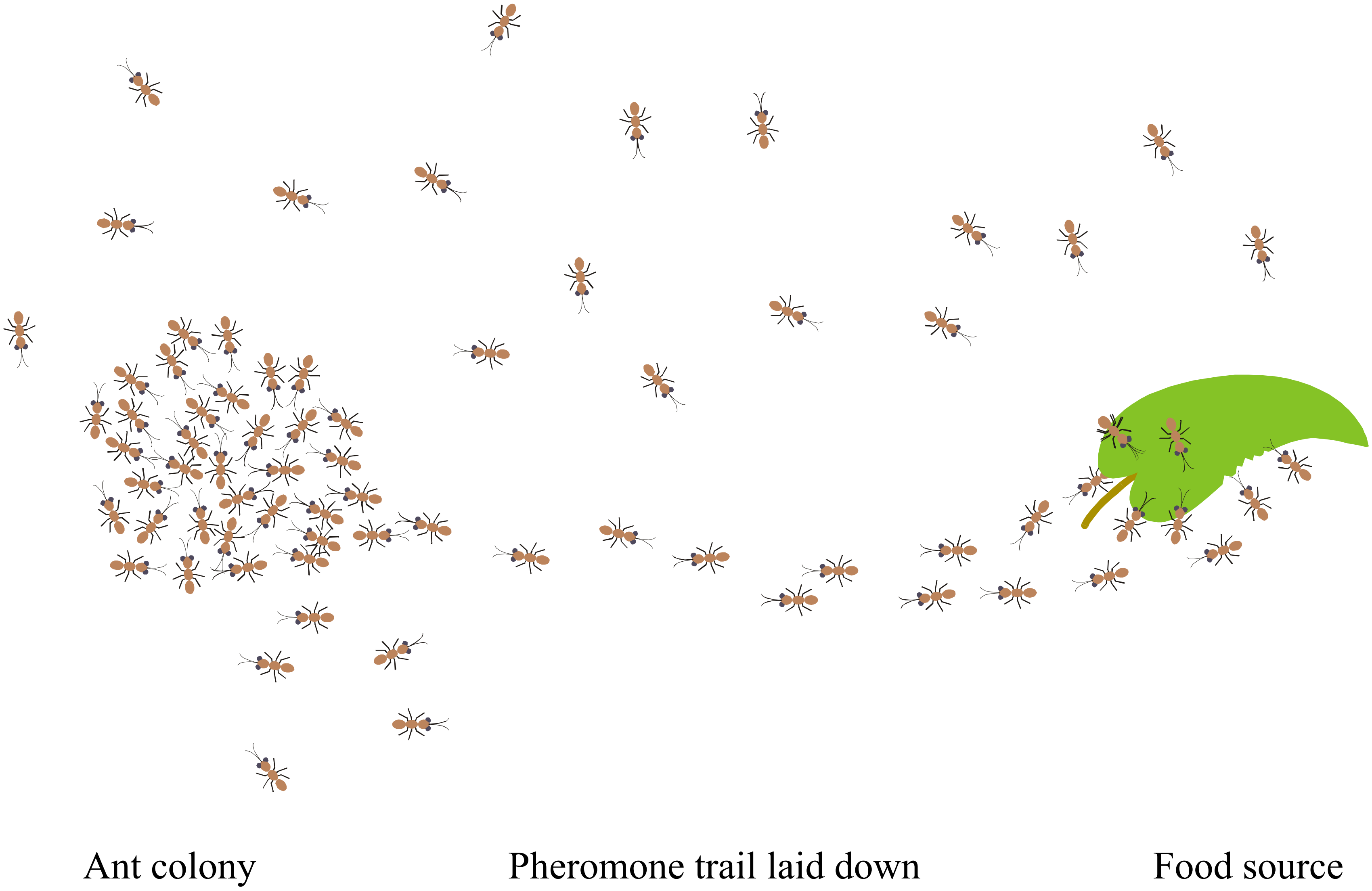
Element of class 2: The ant worker



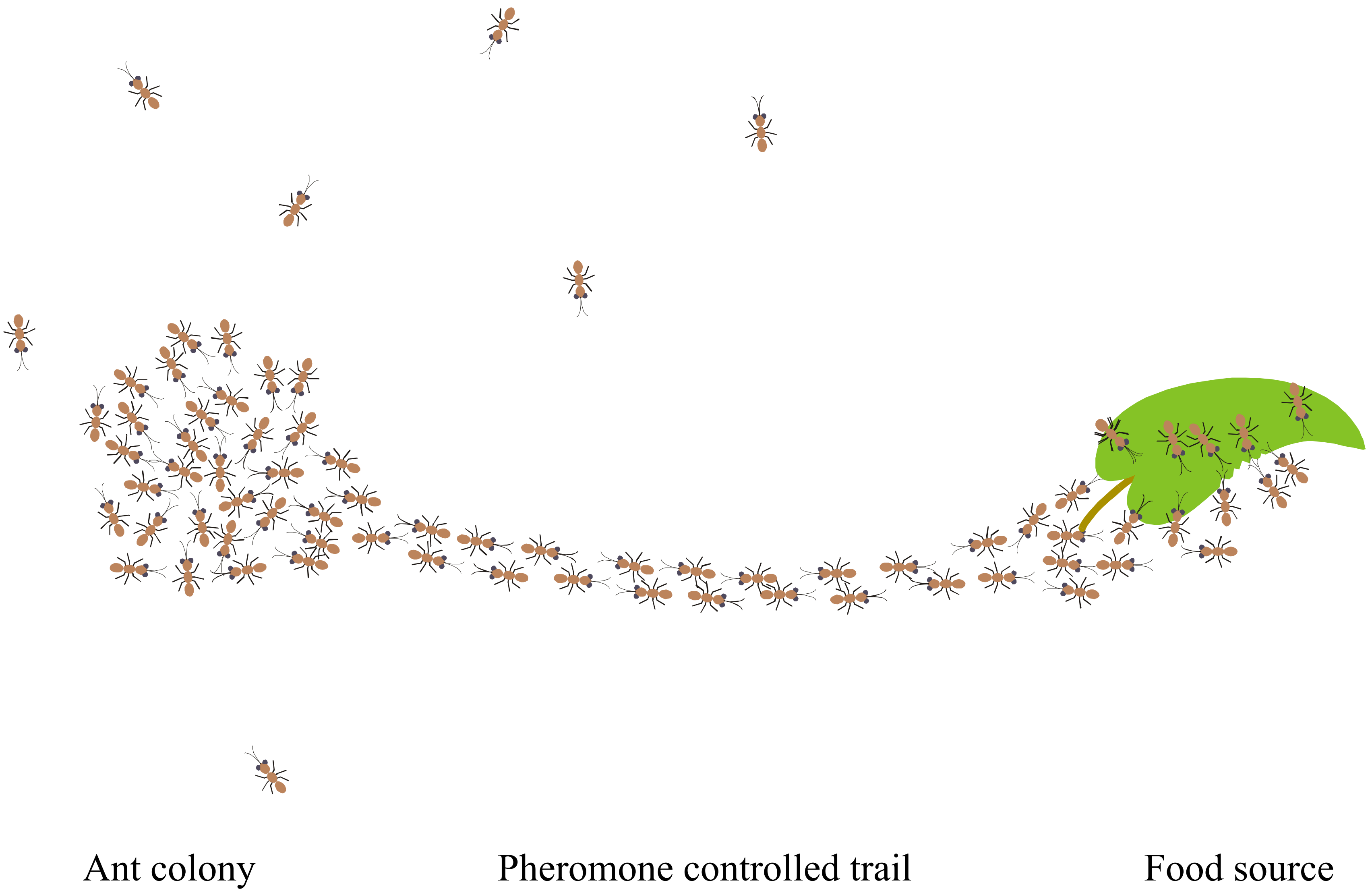
Foraging behavior of ant colonies



Foraging behavior of ant colonies



Foraging behavior of ant colonies



Foraging behavior of ant colonies

	RNA model	**Foraging behavior of ant colonies**
Element	**RNA molecule**	**Individual worker ant**
Mechanism relating elements	Mutation in quasi-species	Genetics of kinship
Search process	Optimization of RNA structure	Recruiting of food
Search space	Sequence space	Three-dimensional space
Random step	Mutation	Element of ant walk
Self-enhancing process	Replication	Secretion of pheromone
Interaction between elements	Mean replication rate	Mean pheromone concentration
Goal of the search	Target structure	Food source
Temporary memory	RNA sequences in population	Pheromone trail
‘Learning’ entity	**Population of molecules**	**Ant colony**

Learning at population or colony level by trial and error

Two examples: (i) RNA model and (ii) ant colony

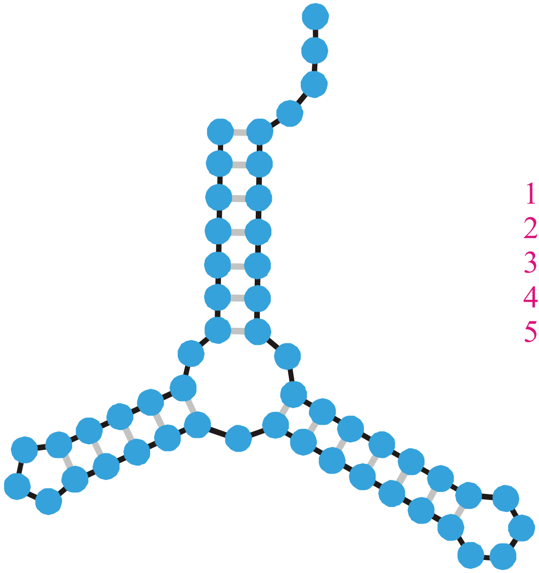
1. Autocatalytic chemical reactions in the flow reactor

2. Replication, mutation, selection and Shannon information

3. Evolution *in silico* and optimization of RNA structures

4. Random walks and ‚ensemble learning‘

5. **Sequence-structure maps, neutral networks, and intersections**



Minimum free energy
criterion

1st
2nd
3rd trial
4th
5th

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

UGGUUACGCGUUGGGGUAACGAAGAUUCCGAGAGGAGUUUAGUGACUAGAGG

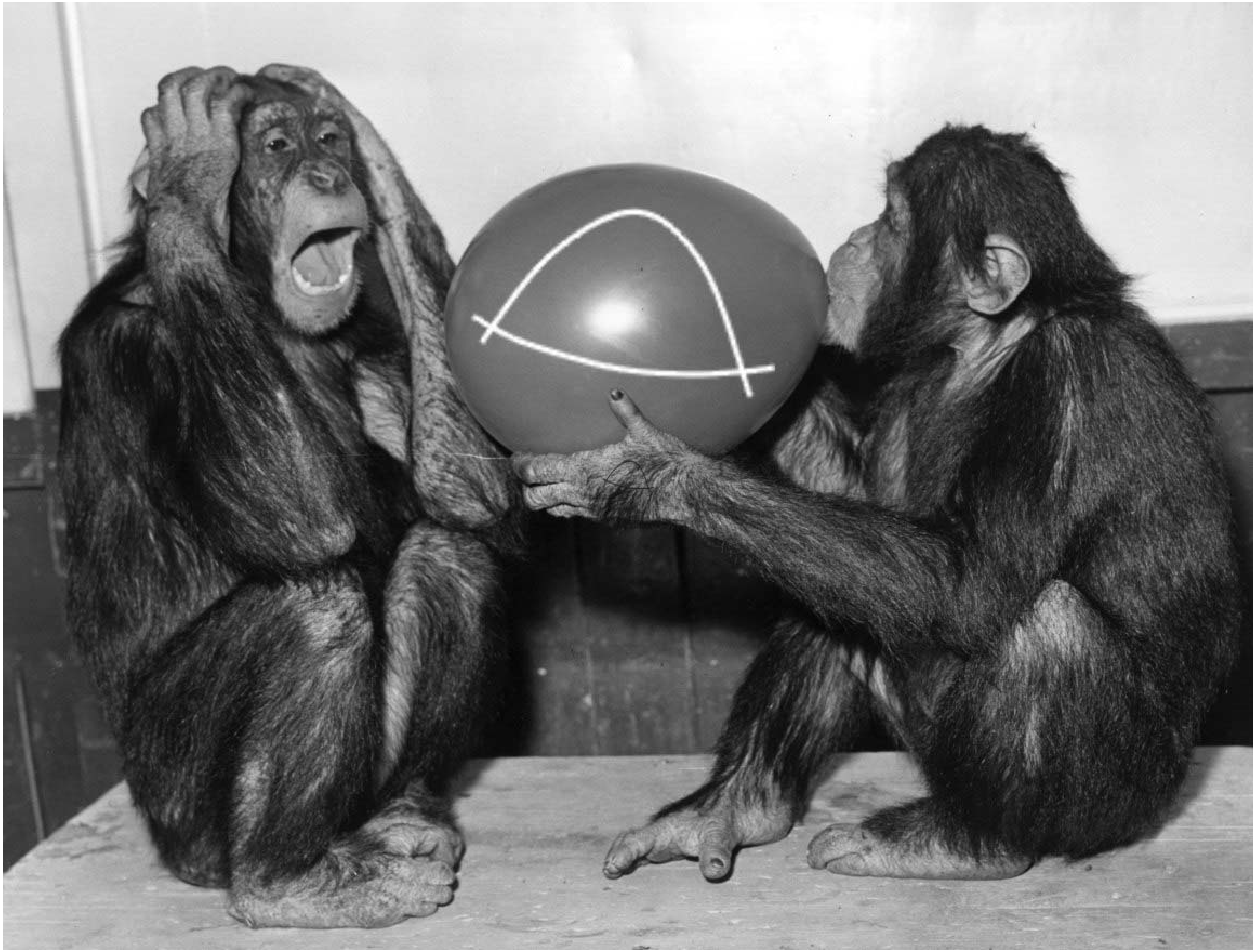
CUUCUUGAGCUAGUACCUAGUCGGAUAGGAUUUCCUAUCUCCAGGGAGGAUG

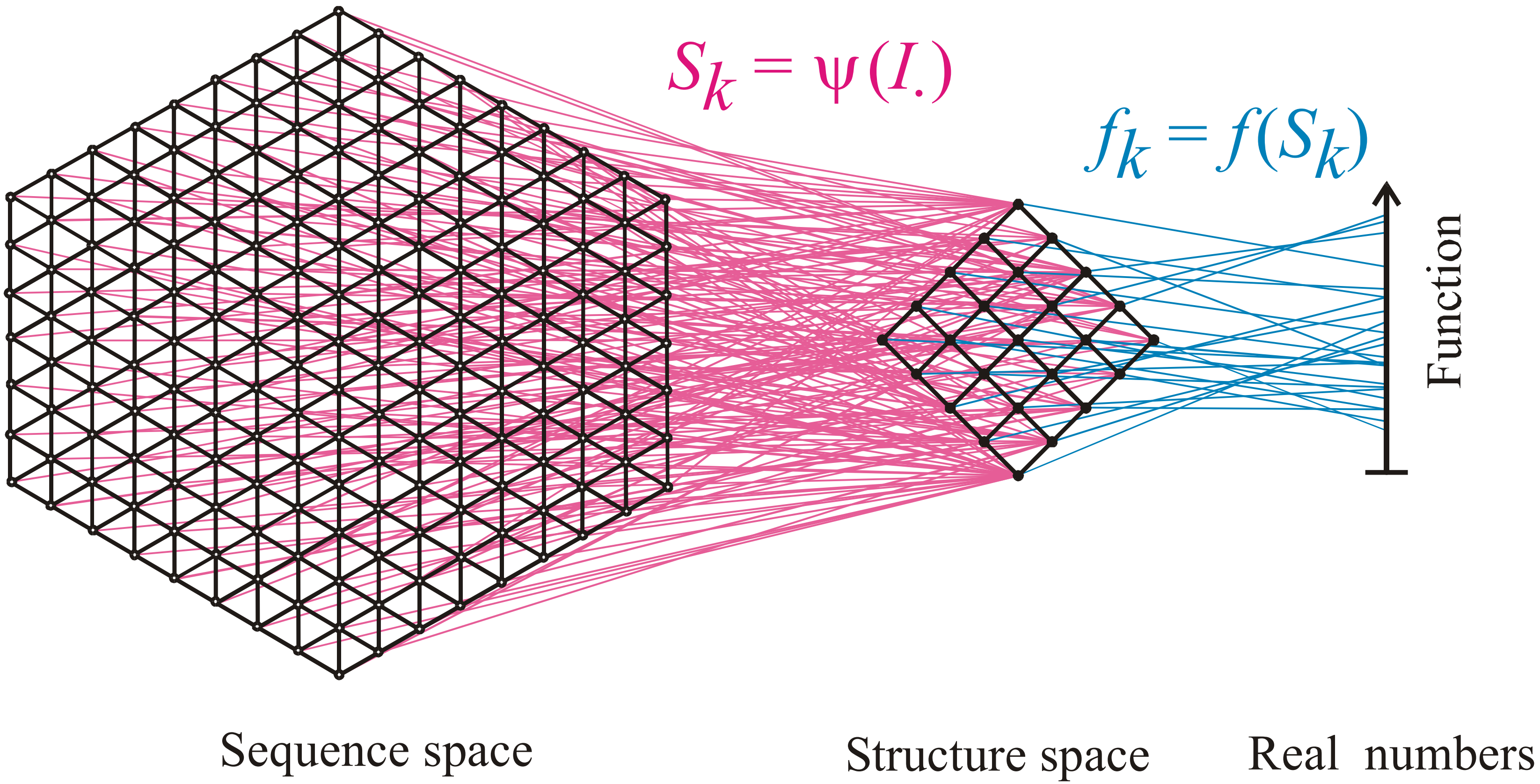
CUUUUCUUCACGUUAGAUGUGUAAUGGACAUGUGUUUAUUUAGGAAAGGCGC

AUAACGUGAGUGUCUAAUACUGAUCGCUCCGGAGGGUGGUGGCGUUGUUAAU

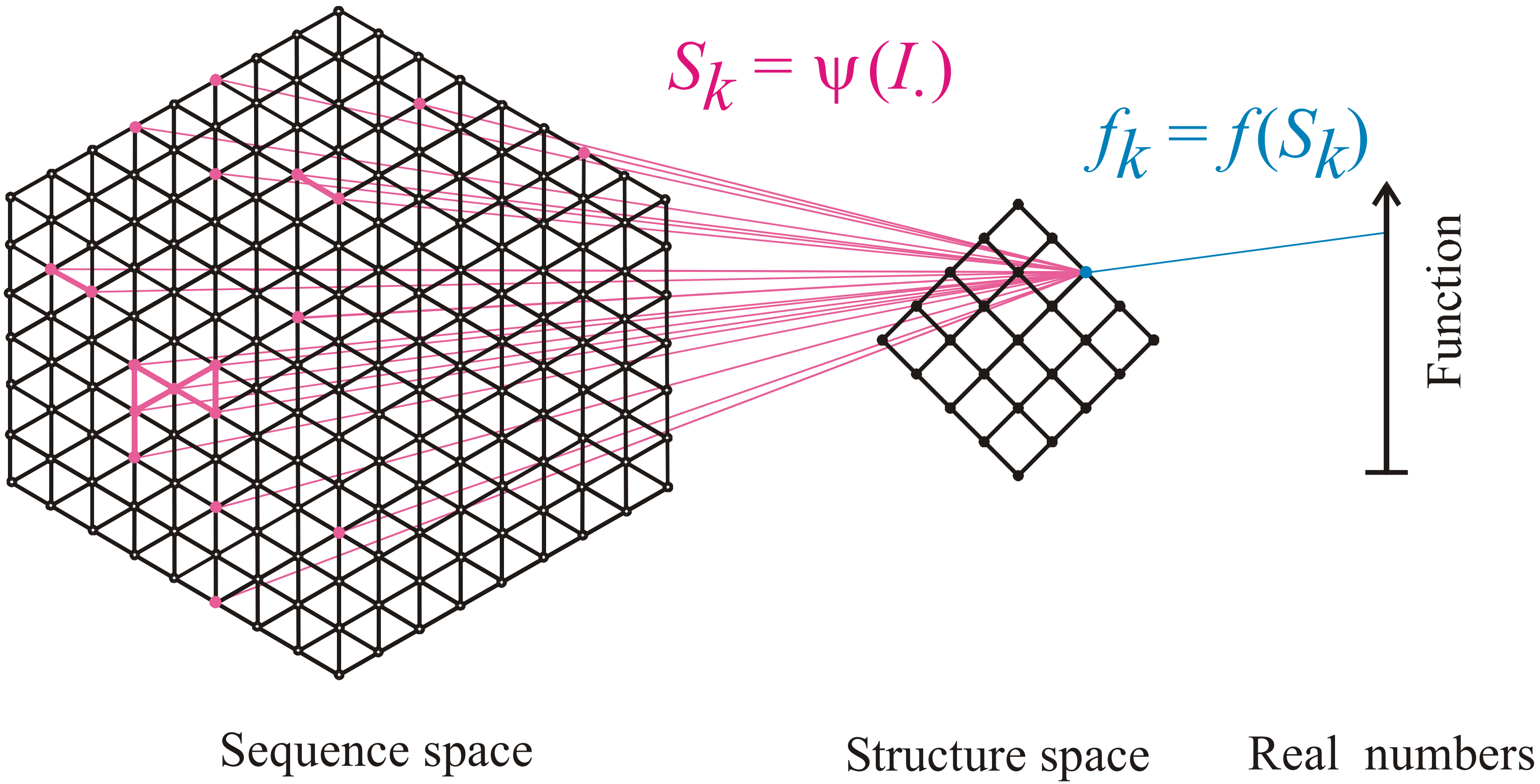
Inverse folding of RNA secondary structures

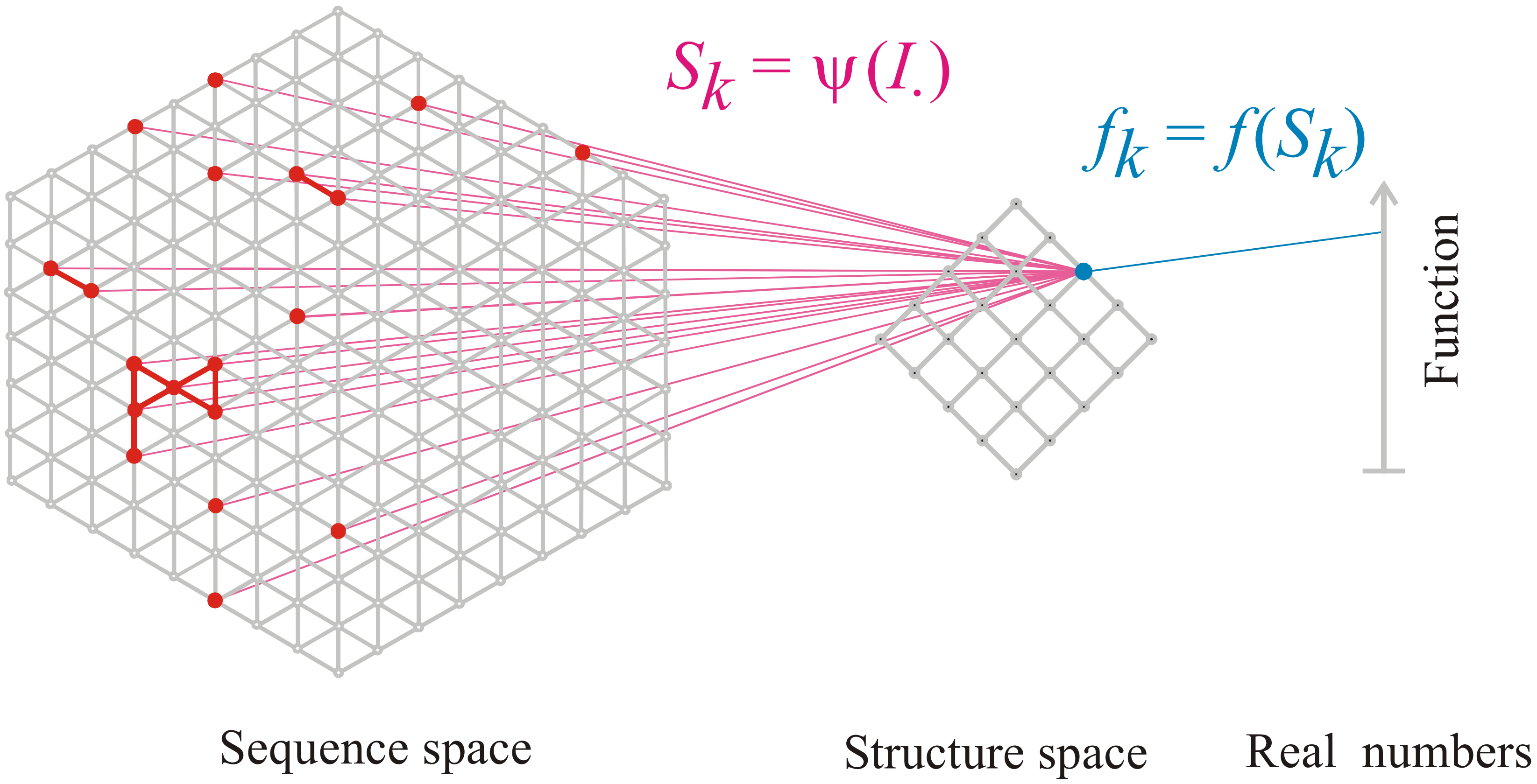
The inverse folding algorithm searches for sequences that form a given RNA secondary structure under the minimum free energy criterion.





Mapping from sequence space into structure space and into function


$S_k = \psi(I.)$
$f_k = f(S_k)$
Function
Sequence space
Structure space
Real numbers



$S_k = \psi(I.)$

$f_k = f(S_k)$

Function

Sequence space

Structure space

Real numbers

The pre-image of the structure S_k in sequence space is the **neutral network G_k**

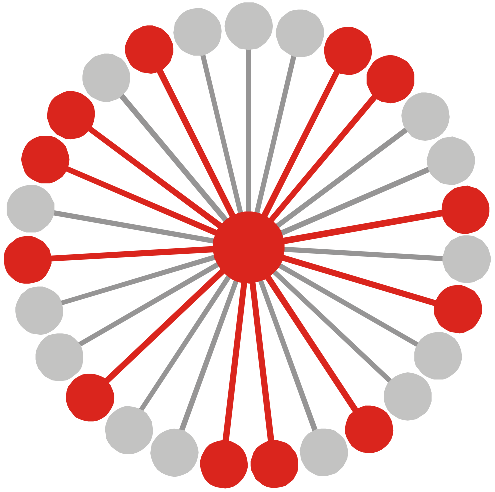
Neutral networks are sets of sequences forming the same structure. $\mathbf{G}_k$ is the pre-image of the structure S_k in sequence space:

$$\mathbf{G}_k = y^{-1}(S_k)\ \pi\ \{y_j \mid y(I_j) = S_k\}$$

The set is converted into a graph by connecting all sequences of Hamming distance one.

Neutral networks of small RNA molecules can be computed by exhaustive folding of complete sequence spaces, i.e. all RNA sequences of a given chain length. This number, $\mathbf{N=4^n}$, becomes very large with increasing length, and is prohibitive for numerical computations.

Neutral networks can be modelled by **random graphs** in sequence space. In this approach, nodes are inserted randomly into sequence space until the size of the pre-image, i.e. the number of neutral sequences, matches the neutral network to be studied.



$$G_k = y^{-1}(S_k) \cup \{ I_j \mid y(I_j) = S_k \}$$

$$\lambda_j = 12 / 27 = 0.444\,, \quad \bar{\lambda}_k = \frac{\underset{j \in |G_k|}{\text{ø}} \; l_j(k)}{|G_k|}$$

Connectivity threshold: $\lambda_{cr} = 1 - \kappa^{-1/(\kappa-1)}$

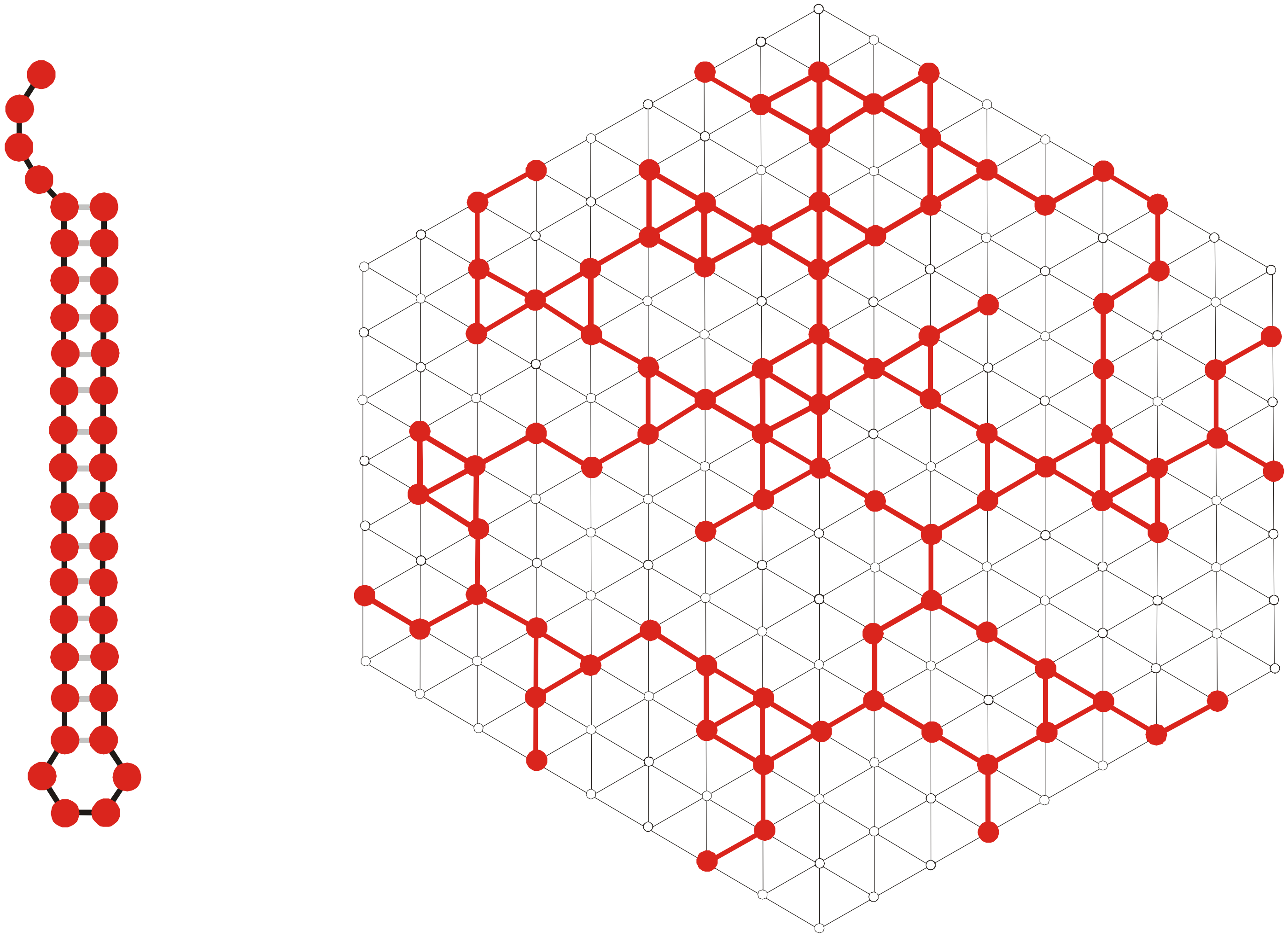
Alphabet size k : **AUGC** ñ k = 4

$\bar{\lambda}_k > \lambda_{cr}$ . . . . network G_k is connected

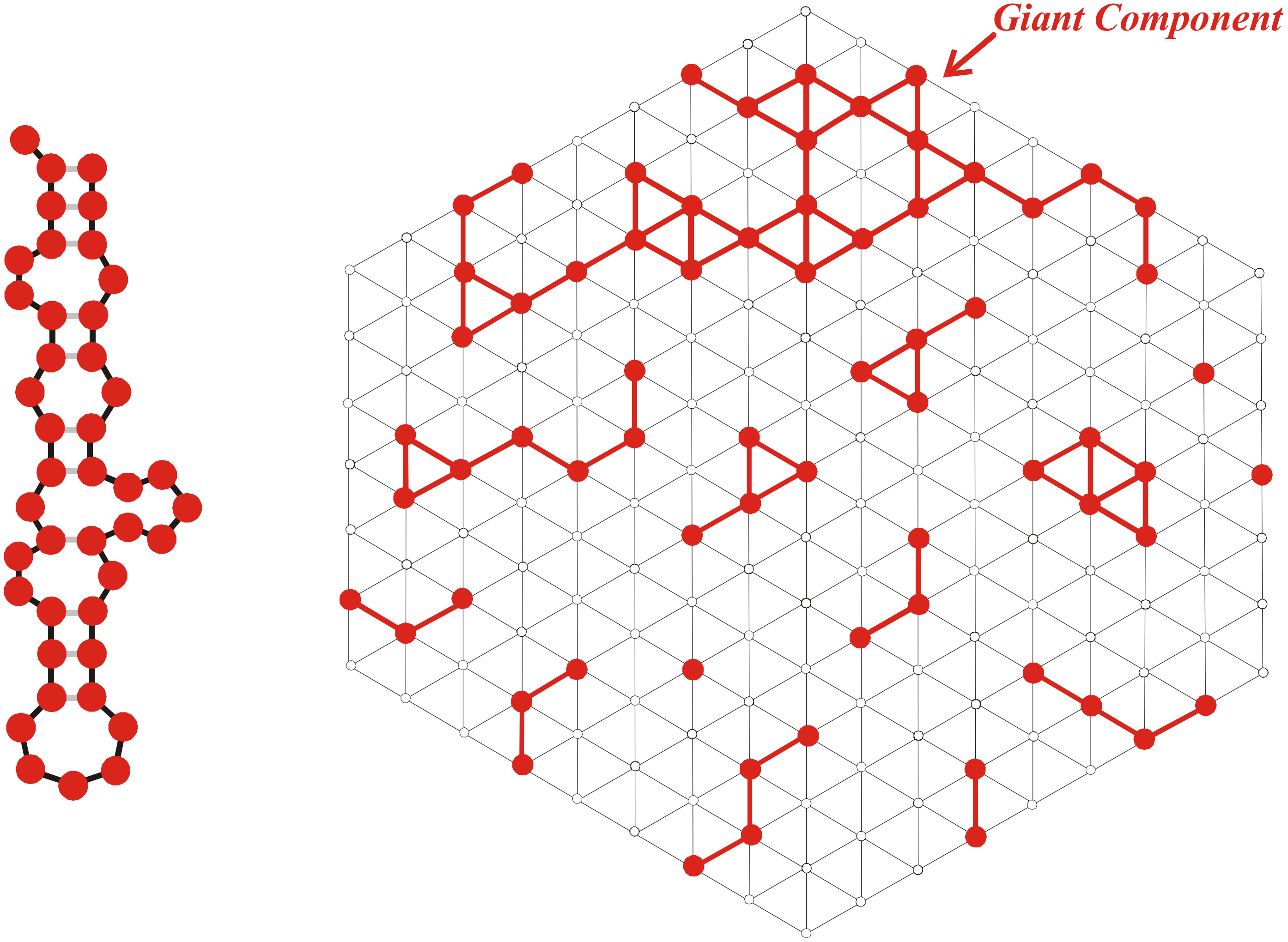
$\bar{\lambda}_k < \lambda_{cr}$ . . . . network G_k is **not** connected

k	l_{cr}	
2	0.5	**GC,AU**
3	0.423	**GUC,AUG**
4	0.370	**AUGC**

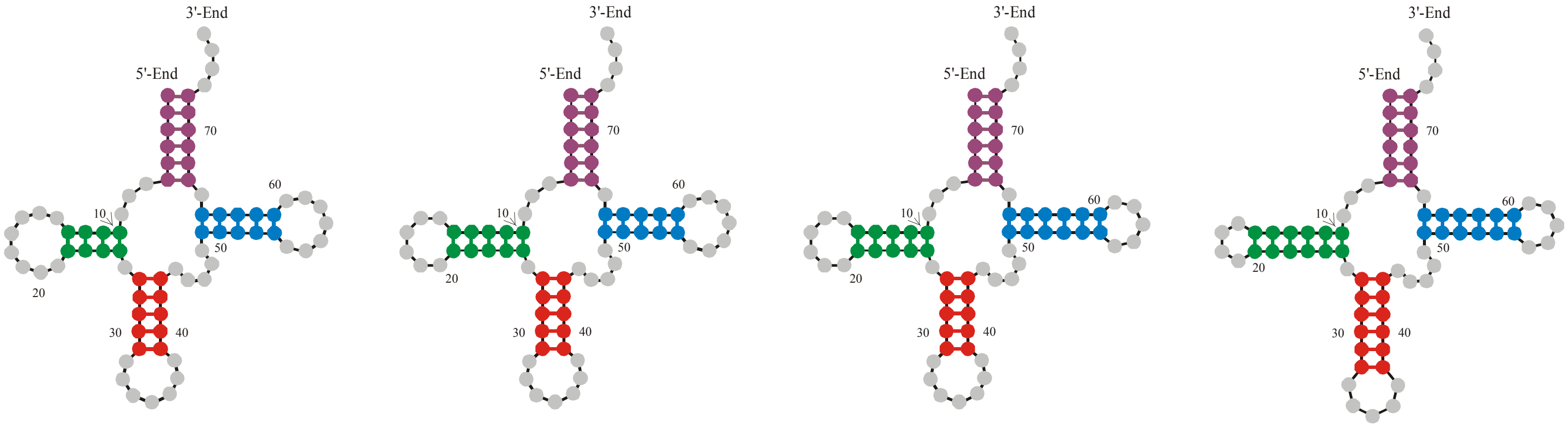
Mean degree of neutrality and connectivity of neutral networks



A connected neutral network

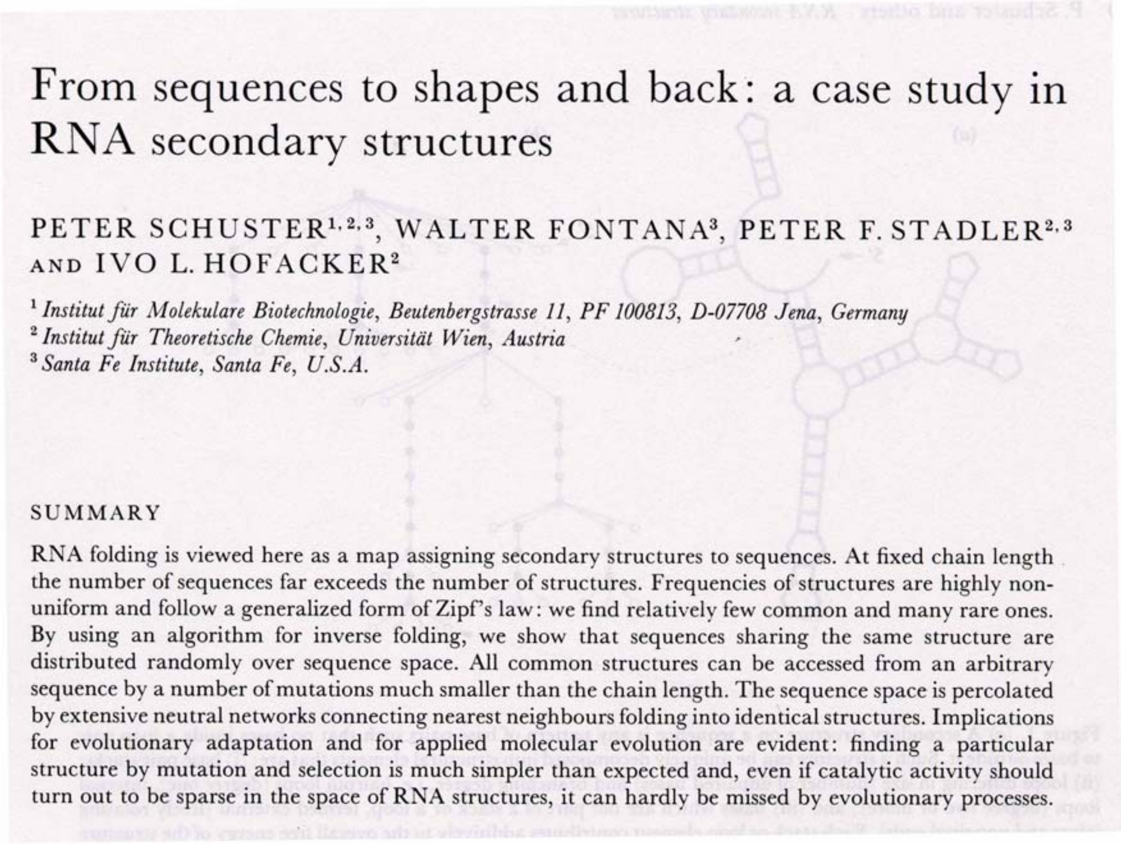


A multi-component neutral network



Alphabet		Degree of neutrality `l		
AU	- -	- -	- -	0.073 Ä 0.032
AUG	- -	0.217 Ä 0.051	0.207 ± 0.055	0.201 Ä 0.056
AUGC	0.275 Ä 0.064	0.279 Ä 0.063	0.289 ± 0.062	0.313 Ä 0.058
UGC	0.263 Ä 0.071	0.257 Ä 0.070	0.251 ± 0.068	0.250 Ä 0.064
GC	0.052 Ä 0.033	0.057 Ä 0.034	0.060 ± 0.033	0.068 Ä 0.034

Degree of neutrality of cloverleaf RNA secondary structures over different alphabets



From sequences to shapes and back: a case study in RNA secondary structures

PETER SCHUSTER[1,2,3], WALTER FONTANA[3], PETER F. STADLER[2,3] AND IVO L. HOFACKER[2]

[1] *Institut für Molekulare Biotechnologie, Beutenbergstrasse 11, PF 100813, D-07708 Jena, Germany*
[2] *Institut für Theoretische Chemie, Universität Wien, Austria*
[3] *Santa Fe Institute, Santa Fe, U.S.A.*

SUMMARY

RNA folding is viewed here as a map assigning secondary structures to sequences. At fixed chain length the number of sequences far exceeds the number of structures. Frequencies of structures are highly non-uniform and follow a generalized form of Zipf's law: we find relatively few common and many rare ones. By using an algorithm for inverse folding, we show that sequences sharing the same structure are distributed randomly over sequence space. All common structures can be accessed from an arbitrary sequence by a number of mutations much smaller than the chain length. The sequence space is percolated by extensive neutral networks connecting nearest neighbours folding into identical structures. Implications for evolutionary adaptation and for applied molecular evolution are evident: finding a particular structure by mutation and selection is much simpler than expected and, even if catalytic activity should turn out to be sparse in the space of RNA structures, it can hardly be missed by evolutionary processes.

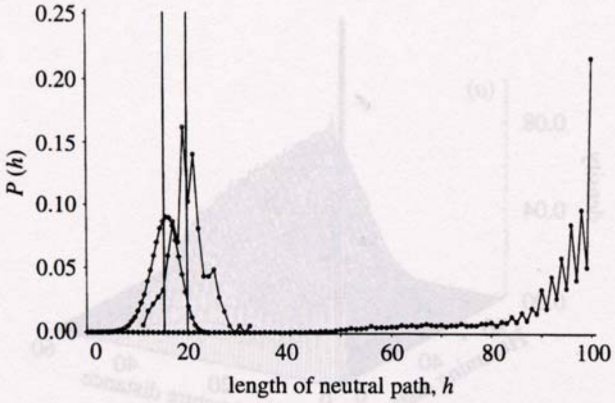
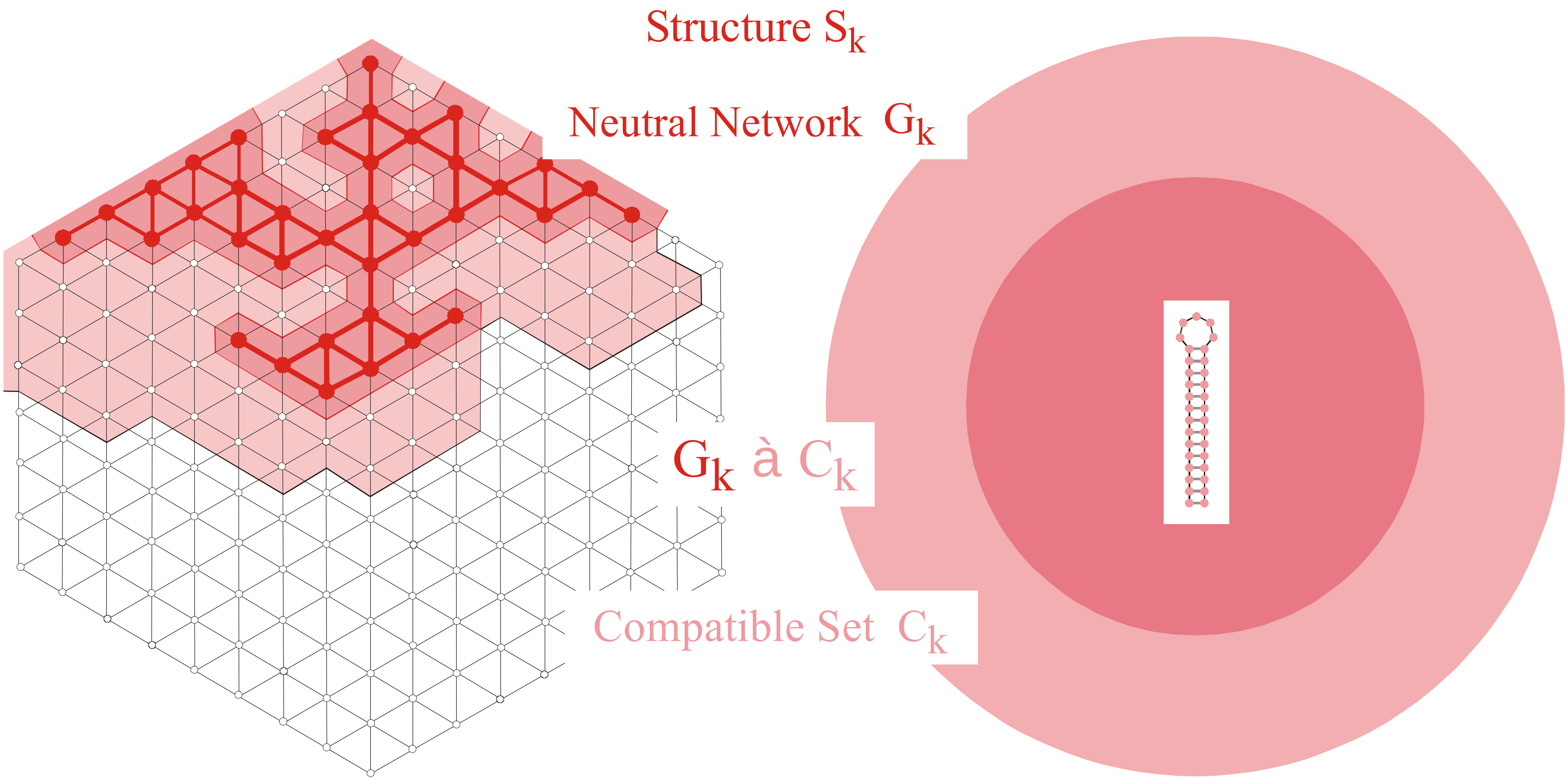
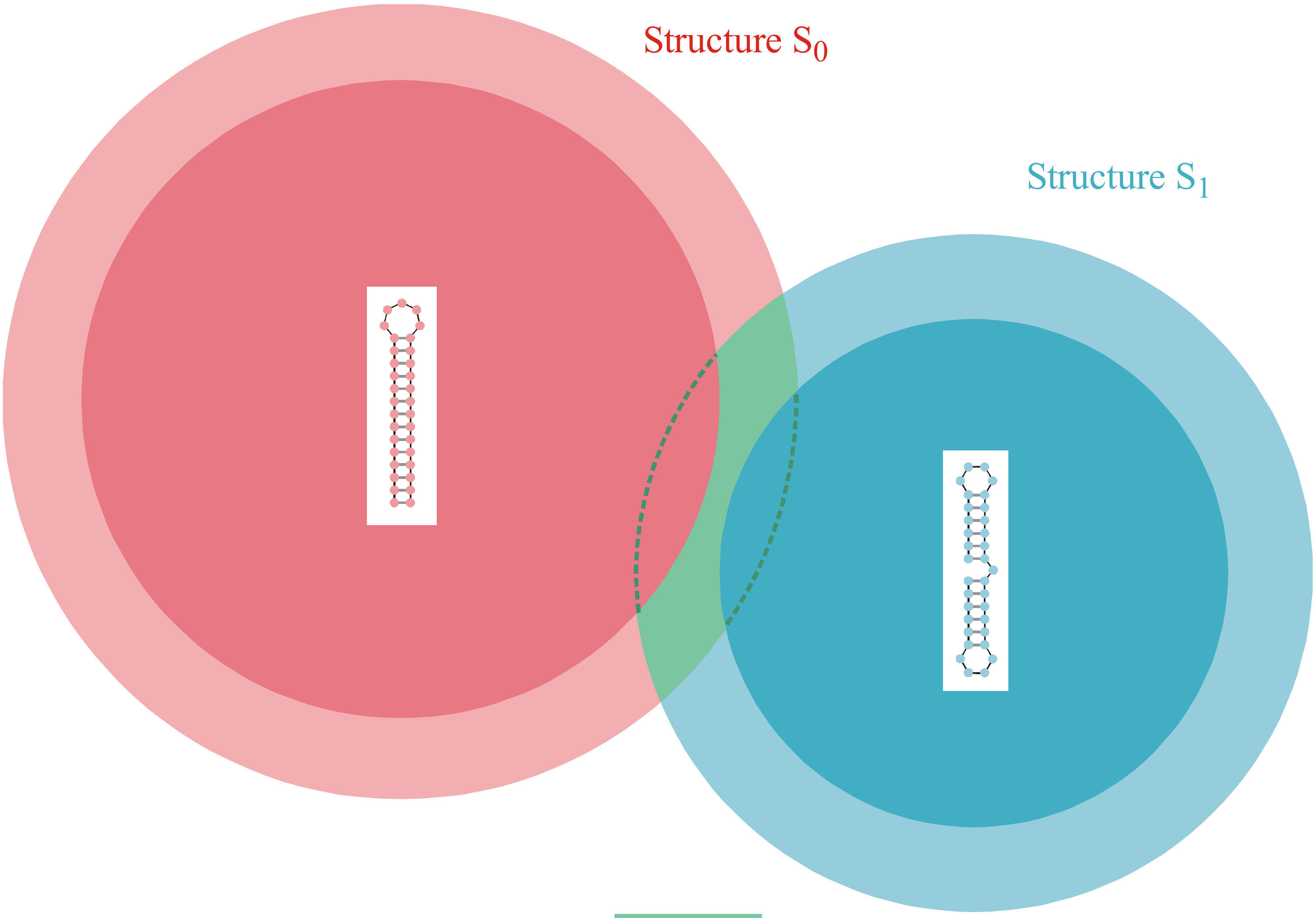


Figure 4. Neutral paths. A neutral path is defined by a series of nearest neighbour sequences that fold into identical structures. Two classes of nearest neighbours are admitted: neighbours of Hamming distance 1, which are obtained by single base exchanges in unpaired stretches of the structure, and neighbours of Hamming distance 2, resulting from base pair exchanges in stacks. Two probability densities of Hamming distances are shown that were obtained by searching for neutral paths in sequence space: (i) an upper bound for the closest approach of trial and target sequences (open circles) obtained as endpoints of neutral paths approaching the target from a random trial sequence (185 targets and 100 trials for each were used); (ii) a lower bound for the closest approach of trial and target sequences (open diamonds) derived from secondary structure statistics (Fontana *et al.* 1993*a*; see this paper, §4); and (iii) longest distances between the reference and the endpoints of monotonously diverging neutral paths (filled circles) (500 reference sequences were used).

Reference for postulation and ***in silico*** verification of ***neutral networks***

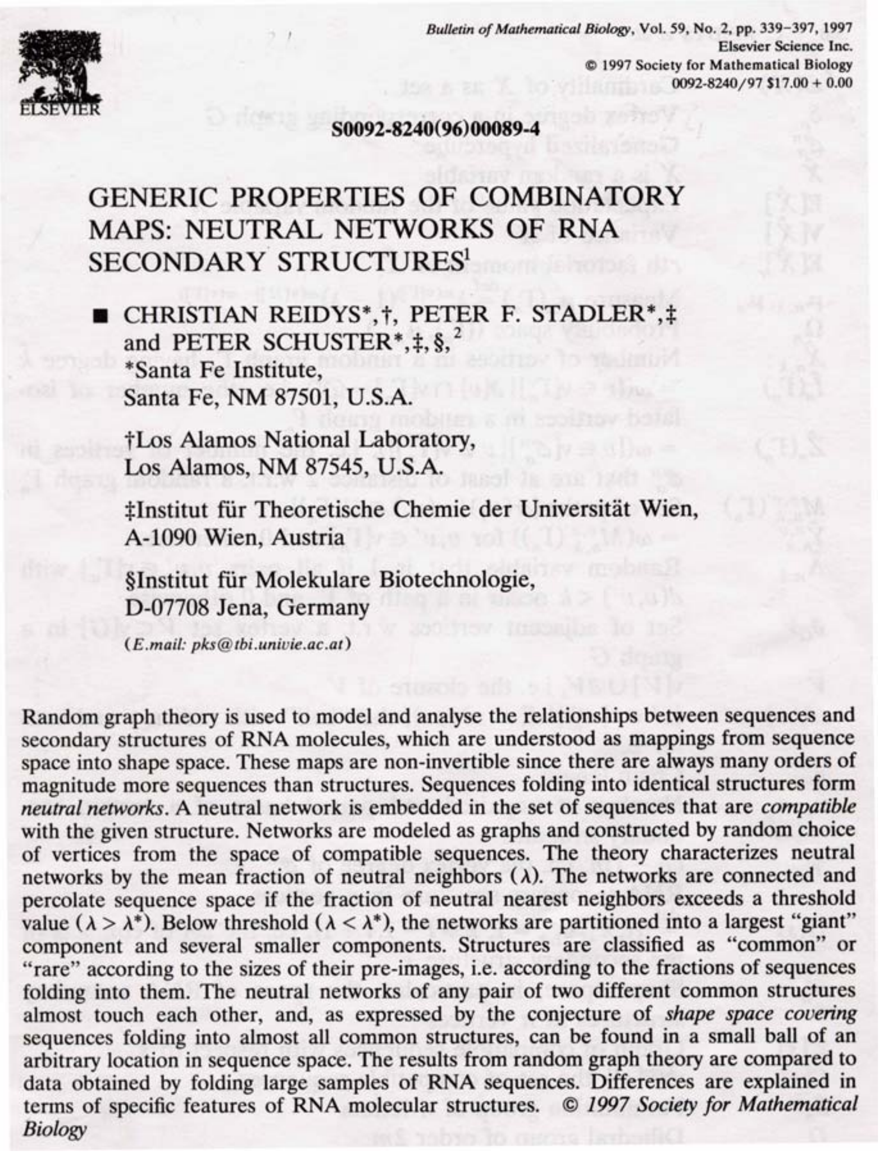


The **compatible set C_k** of a structure **S_k** consists of all sequences which form **S_k** as its minimum free energy structure (the neutral network **G_k**) or one of its suboptimal structures.



Intersection of two compatible sets: $C_0 \cap C_1$

The intersection of two compatible sets is always non empty: C_0 Ú C_1 â Ù



Bulletin of Mathematical Biology, Vol. 59, No. 2, pp. 339–397, 1997
Elsevier Science Inc.
© 1997 Society for Mathematical Biology
0092-8240/97 $17.00 + 0.00

S0092-8240(96)00089-4

GENERIC PROPERTIES OF COMBINATORY MAPS: NEUTRAL NETWORKS OF RNA SECONDARY STRUCTURES[1]

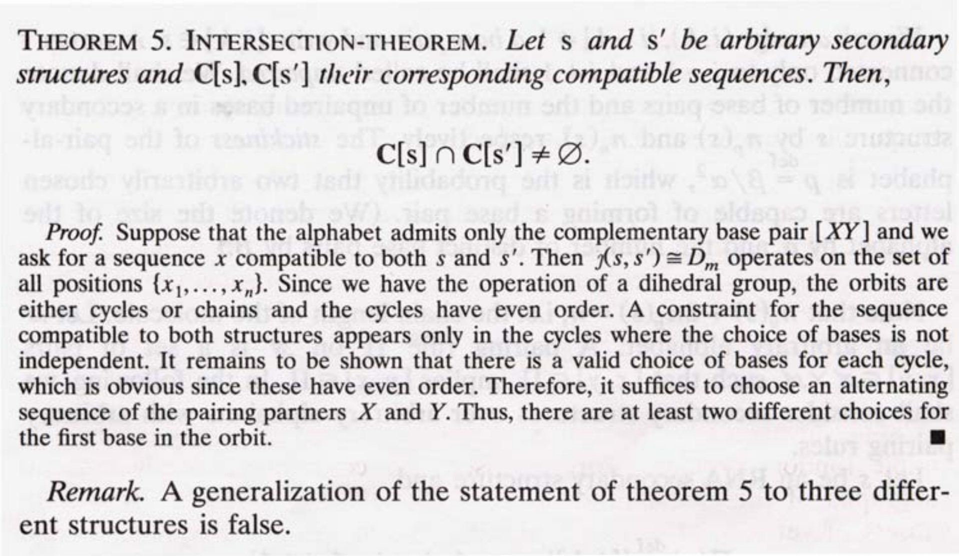
■ CHRISTIAN REIDYS*,†, PETER F. STADLER*,‡ and PETER SCHUSTER*,‡,§,[2]
*Santa Fe Institute,
Santa Fe, NM 87501, U.S.A.

†Los Alamos National Laboratory,
Los Alamos, NM 87545, U.S.A.

‡Institut für Theoretische Chemie der Universität Wien,
A-1090 Wien, Austria

§Institut für Molekulare Biotechnologie,
D-07708 Jena, Germany
(*E.mail: pks@tbi.univie.ac.at*)

Random graph theory is used to model and analyse the relationships between sequences and secondary structures of RNA molecules, which are understood as mappings from sequence space into shape space. These maps are non-invertible since there are always many orders of magnitude more sequences than structures. Sequences folding into identical structures form *neutral networks*. A neutral network is embedded in the set of sequences that are *compatible* with the given structure. Networks are modeled as graphs and constructed by random choice of vertices from the space of compatible sequences. The theory characterizes neutral networks by the mean fraction of neutral neighbors (λ). The networks are connected and percolate sequence space if the fraction of neutral nearest neighbors exceeds a threshold value ($\lambda > \lambda^*$). Below threshold ($\lambda < \lambda^*$), the networks are partitioned into a largest "giant" component and several smaller components. Structures are classified as "common" or "rare" according to the sizes of their pre-images, i.e. according to the fractions of sequences folding into them. The neutral networks of any pair of two different common structures almost touch each other, and, as expressed by the conjecture of *shape space covering* sequences folding into almost all common structures, can be found in a small ball of an arbitrary location in sequence space. The results from random graph theory are compared to data obtained by folding large samples of RNA sequences. Differences are explained in terms of specific features of RNA molecular structures. © *1997 Society for Mathematical Biology*



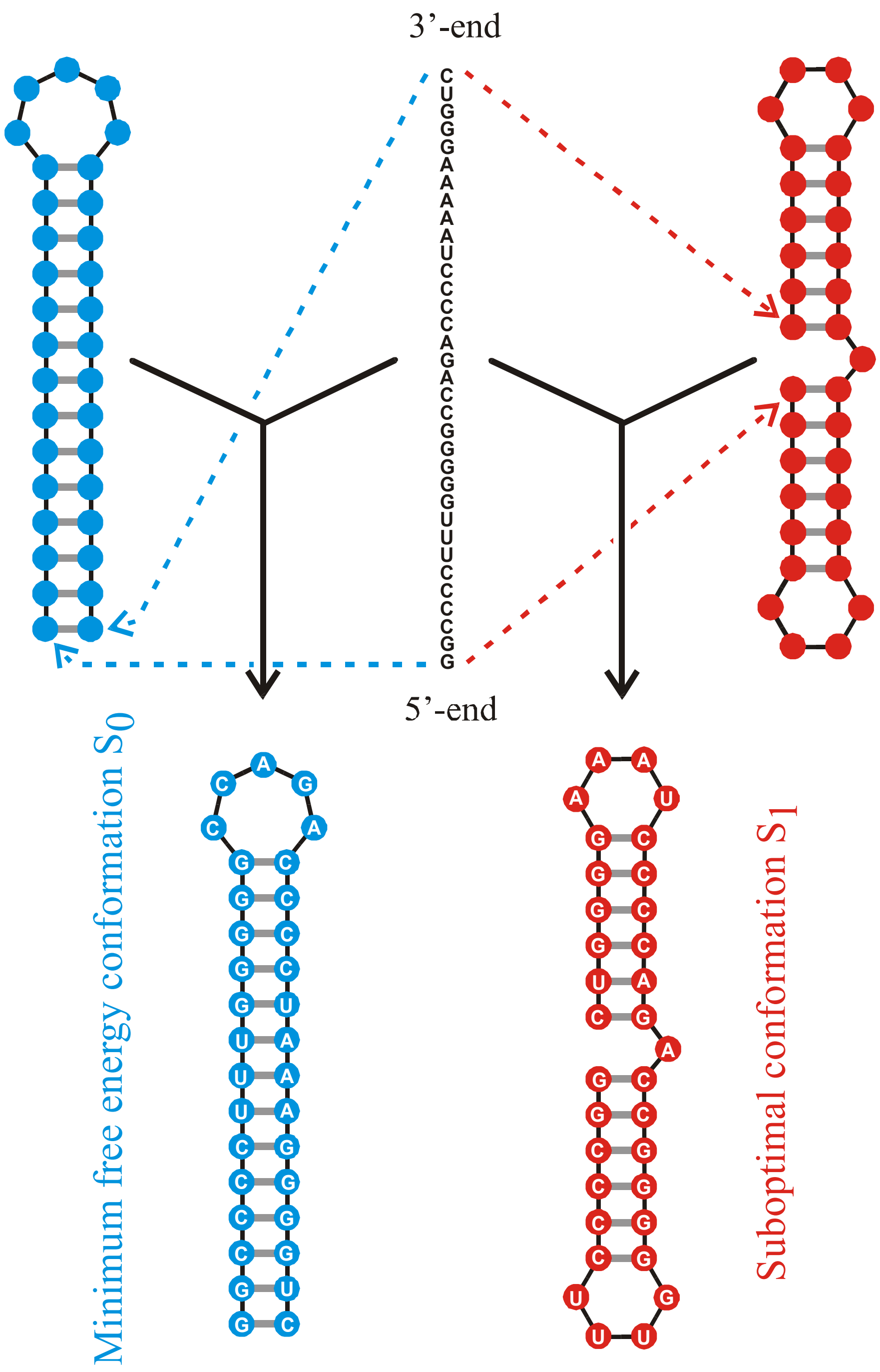
THEOREM 5. INTERSECTION-THEOREM. *Let* s *and* s′ *be arbitrary secondary structures and* C[s]. C[s′] *their corresponding compatible sequences. Then,*

$$\mathrm{C}[\mathrm{s}] \cap \mathrm{C}[\mathrm{s}'] \neq \varnothing.$$

Proof. Suppose that the alphabet admits only the complementary base pair $[XY]$ and we ask for a sequence x compatible to both s and s'. Then $\mathcal{J}(s, s') \cong D_m$ operates on the set of all positions $\{x_1, \ldots, x_n\}$. Since we have the operation of a dihedral group, the orbits are either cycles or chains and the cycles have even order. A constraint for the sequence compatible to both structures appears only in the cycles where the choice of bases is not independent. It remains to be shown that there is a valid choice of bases for each cycle, which is obvious since these have even order. Therefore, it suffices to choose an alternating sequence of the pairing partners X and Y. Thus, there are at least two different choices for the first base in the orbit. ■

Remark. A generalization of the statement of theorem 5 to three different structures is false.

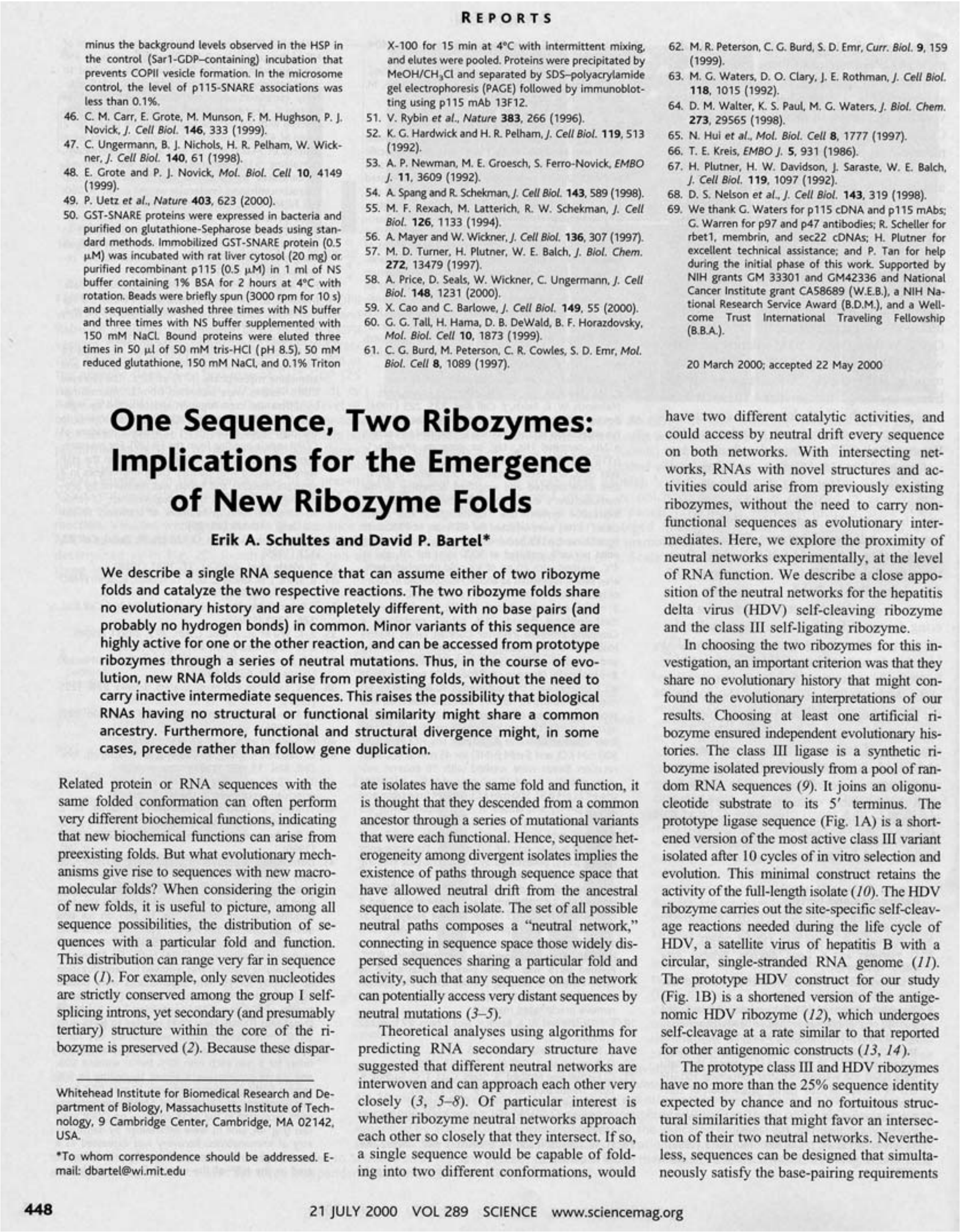
Reference for the definition of the intersection and the proof of the **intersection theorem**


3’-end
CUGGGAAAAAUCCCCAGACCGGGGUUUCCCCGG
5’-end
Minimum free energy conformation S_0
Suboptimal conformation S_1

A sequence at the **intersection** of two neutral networks is compatible with both structures

A ribozyme switch

E.A.Schultes, D.B.Bartel,
Science **289** (2000), 448-452



minus the background levels observed in the HSP in the control (Sar1-GDP–containing) incubation that prevents COPII vesicle formation. In the microsome control, the level of p115-SNARE associations was less than 0.1%.

46. C. M. Carr, E. Grote, M. Munson, F. M. Hughson, P. J. Novick, *J. Cell Biol.* **146**, 333 (1999).
47. C. Ungermann, B. J. Nichols, H. R. Pelham, W. Wickner, *J. Cell Biol.* **140**, 61 (1998).
48. E. Grote and P. J. Novick, *Mol. Biol. Cell* **10**, 4149 (1999).
49. P. Uetz *et al.*, *Nature* **403**, 623 (2000).
50. GST-SNARE proteins were expressed in bacteria and purified on glutathione-Sepharose beads using standard methods. Immobilized GST-SNARE protein (0.5 μM) was incubated with rat liver cytosol (20 mg) or purified recombinant p115 (0.5 μM) in 1 ml of NS buffer containing 1% BSA for 2 hours at 4°C with rotation. Beads were briefly spun (3000 rpm for 10 s) and sequentially washed three times with NS buffer and three times with NS buffer supplemented with 150 mM NaCl. Bound proteins were eluted three times in 50 μl of 50 mM tris-HCl (pH 8.5), 50 mM reduced glutathione, 150 mM NaCl, and 0.1% Triton X-100 for 15 min at 4°C with intermittent mixing, and elutes were pooled. Proteins were precipitated by MeOH/CH_3Cl and separated by SDS–polyacrylamide gel electrophoresis (PAGE) followed by immunoblotting using p115 mAb 13F12.
51. V. Rybin *et al.*, *Nature* **383**, 266 (1996).
52. K. G. Hardwick and H. R. Pelham, *J. Cell Biol.* **119**, 513 (1992).
53. A. P. Newman, M. E. Groesch, S. Ferro-Novick, *EMBO J.* **11**, 3609 (1992).
54. A. Spang and R. Schekman, *J. Cell Biol.* **143**, 589 (1998).
55. M. F. Rexach, M. Latterich, R. W. Schekman, *J. Cell Biol.* **126**, 1133 (1994).
56. A. Mayer and W. Wickner, *J. Cell Biol.* **136**, 307 (1997).
57. M. D. Turner, H. Plutner, W. E. Balch, *J. Biol. Chem.* **272**, 13479 (1997).
58. A. Price, D. Seals, W. Wickner, C. Ungermann, *J. Cell Biol.* **148**, 1231 (2000).
59. X. Cao and C. Barlowe, *J. Cell Biol.* **149**, 55 (2000).
60. G. G. Tall, H. Hama, D. B. DeWald, B. F. Horazdovsky, *Mol. Biol. Cell* **10**, 1873 (1999).
61. C. G. Burd, M. Peterson, C. R. Cowles, S. D. Emr, *Mol. Biol. Cell* **8**, 1089 (1997).
62. M. R. Peterson, C. G. Burd, S. D. Emr, *Curr. Biol.* **9**, 159 (1999).
63. M. G. Waters, D. O. Clary, J. E. Rothman, *J. Cell Biol.* **118**, 1015 (1992).
64. D. M. Walter, K. S. Paul, M. G. Waters, *J. Biol. Chem.* **273**, 29565 (1998).
65. N. Hui *et al.*, *Mol. Biol. Cell* **8**, 1777 (1997).
66. T. E. Kreis, *EMBO J.* **5**, 931 (1986).
67. H. Plutner, H. W. Davidson, J. Saraste, W. E. Balch, *J. Cell Biol.* **119**, 1097 (1992).
68. D. S. Nelson *et al.*, *J. Cell Biol.* **143**, 319 (1998).

69. We thank G. Waters for p115 cDNA and p115 mAbs; G. Warren for p97 and p47 antibodies; R. Scheller for rbet1, membrin, and sec22 cDNAs; H. Plutner for excellent technical assistance; and P. Tan for help during the initial phase of this work. Supported by NIH grants GM 33301 and GM42336 and National Cancer Institute grant CA58689 (W.E.B.), a NIH National Research Service Award (B.D.M.), and a Wellcome Trust International Traveling Fellowship (B.B.A.).

20 March 2000; accepted 22 May 2000

One Sequence, Two Ribozymes: Implications for the Emergence of New Ribozyme Folds

Erik A. Schultes and David P. Bartel*

We describe a single RNA sequence that can assume either of two ribozyme folds and catalyze the two respective reactions. The two ribozyme folds share no evolutionary history and are completely different, with no base pairs (and probably no hydrogen bonds) in common. Minor variants of this sequence are highly active for one or the other reaction, and can be accessed from prototype ribozymes through a series of neutral mutations. Thus, in the course of evolution, new RNA folds could arise from preexisting folds, without the need to carry inactive intermediate sequences. This raises the possibility that biological RNAs having no structural or functional similarity might share a common ancestry. Furthermore, functional and structural divergence might, in some cases, precede rather than follow gene duplication.

Related protein or RNA sequences with the same folded conformation can often perform very different biochemical functions, indicating that new biochemical functions can arise from preexisting folds. But what evolutionary mechanisms give rise to sequences with new macromolecular folds? When considering the origin of new folds, it is useful to picture, among all sequence possibilities, the distribution of sequences with a particular fold and function. This distribution can range very far in sequence space (*1*). For example, only seven nucleotides are strictly conserved among the group I self-splicing introns, yet secondary (and presumably tertiary) structure within the core of the ribozyme is preserved (*2*). Because these disparate isolates have the same fold and function, it is thought that they descended from a common ancestor through a series of mutational variants that were each functional. Hence, sequence heterogeneity among divergent isolates implies the existence of paths through sequence space that have allowed neutral drift from the ancestral sequence to each isolate. The set of all possible neutral paths composes a "neutral network," connecting in sequence space those widely dispersed sequences sharing a particular fold and activity, such that any sequence on the network can potentially access very distant sequences by neutral mutations (*3–5*).

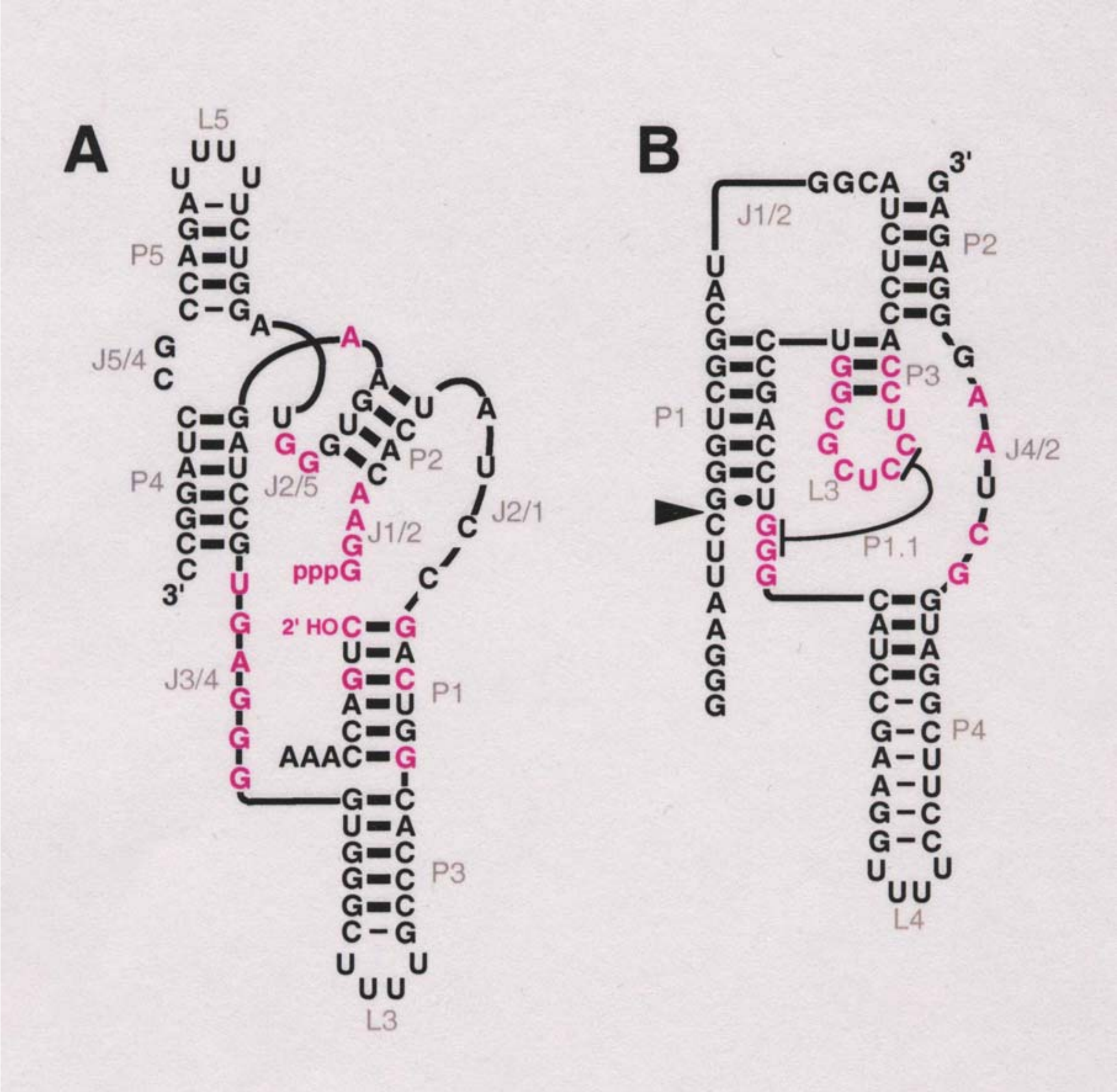
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect. If so, a single sequence would be capable of folding into two different conformations, would have two different catalytic activities, and could access by neutral drift every sequence on both networks. With intersecting networks, RNAs with novel structures and activities could arise from previously existing ribozymes, without the need to carry nonfunctional sequences as evolutionary intermediates. Here, we explore the proximity of neutral networks experimentally, at the level of RNA function. We describe a close apposition of the neutral networks for the hepatitis delta virus (HDV) self-cleaving ribozyme and the class III self-ligating ribozyme.

In choosing the two ribozymes for this investigation, an important criterion was that they share no evolutionary history that might confound the evolutionary interpretations of our results. Choosing at least one artificial ribozyme ensured independent evolutionary histories. The class III ligase is a synthetic ribozyme isolated previously from a pool of random RNA sequences (*9*). It joins an oligonucleotide substrate to its 5′ terminus. The prototype ligase sequence (Fig. 1A) is a shortened version of the most active class III variant isolated after 10 cycles of in vitro selection and evolution. This minimal construct retains the activity of the full-length isolate (*10*). The HDV ribozyme carries out the site-specific self-cleavage reactions needed during the life cycle of HDV, a satellite virus of hepatitis B with a circular, single-stranded RNA genome (*11*). The prototype HDV construct for our study (Fig. 1B) is a shortened version of the antigenomic HDV ribozyme (*12*), which undergoes self-cleavage at a rate similar to that reported for other antigenomic constructs (*13*, *14*).

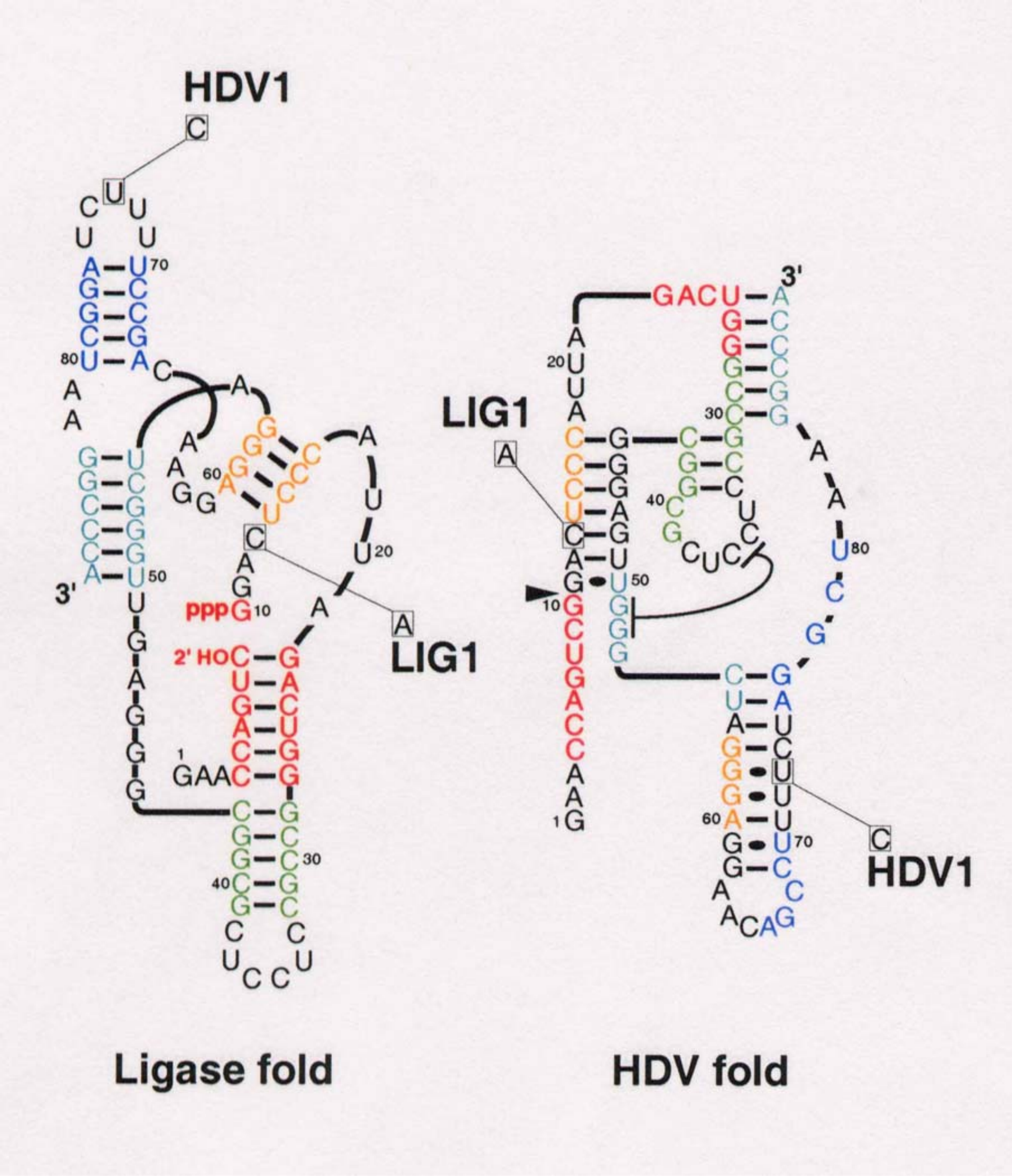
The prototype class III and HDV ribozymes have no more than the 25% sequence identity expected by chance and no fortuitous structural similarities that might favor an intersection of their two neutral networks. Nevertheless, sequences can be designed that simultaneously satisfy the base-pairing requirements

Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA.

*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

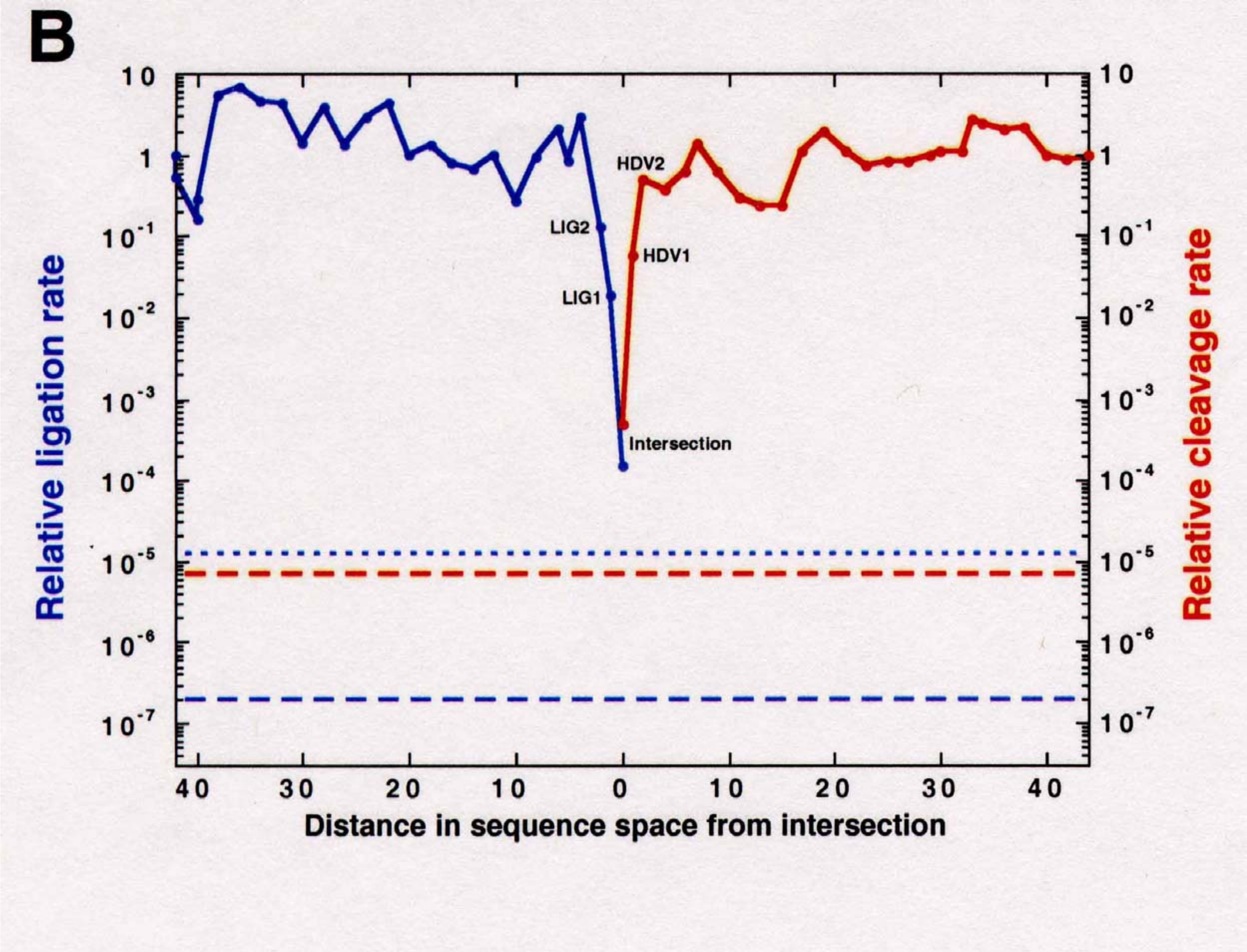


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-d-virus (**B**)



The sequence at the *intersection*:

An RNA molecules which is 88 nucleotides long and can form both structures



Two neutral walks through sequence space with conservation of structure and catalytic activity

Evolution of RNA molecules based on Qβ phage

D.R.Mills, R.L.Peterson, S.Spiegelman, ***An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule***. Proc.Natl.Acad.Sci.USA **58** (1967), 217-224

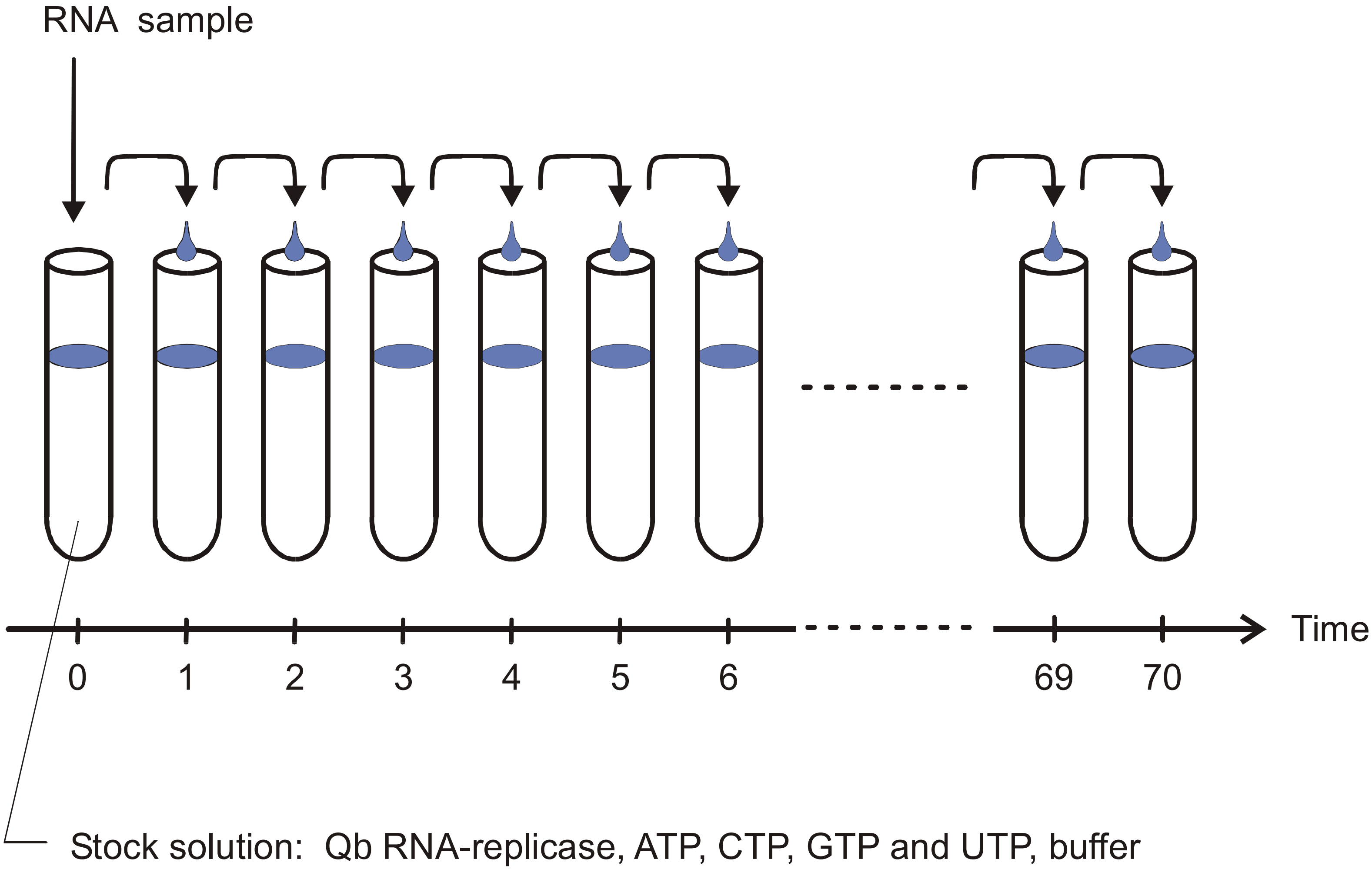
S.Spiegelman, ***An approach to the experimental analysis of precellular evolution***. Quart.Rev.Biophys. **4** (1971), 213-253

C.K.Biebricher, ***Darwinian selection of self-replicating RNA molecules***. Evolutionary Biology **16** (1983), 1-52

G.Bauer, H.Otten, J.S.McCaskill, ***Travelling waves of*** **in vitro** ***evolving RNA.*** *Proc.Natl.Acad.Sci.USA* **86** (1989), 7937-7941

C.K.Biebricher, W.C.Gardiner, ***Molecular evolution of RNA*** **in vitro**. Biophysical Chemistry **66** (1997), 179-192

G.Strunk, T.Ederhof, ***Machines for automated evolution experiments*** **in vitro** ***based on the serial transfer concept***. Biophysical Chemistry **66** (1997), 193-202


RNA sample
0
1
2
3
4
5
6
69
70
Time
Stock solution: Qb RNA-replicase, ATP, CTP, GTP and UTP, buffer

The serial transfer technique applied to RNA evolution ***in vitro***

Reproduction of the original figure of the serial transfer experiment with Qβ RNA

D.R.Mills, R,L,Peterson, S.Spiegelman, *An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule*. Proc.Natl.Acad.Sci.USA **58** (1967), 217-224

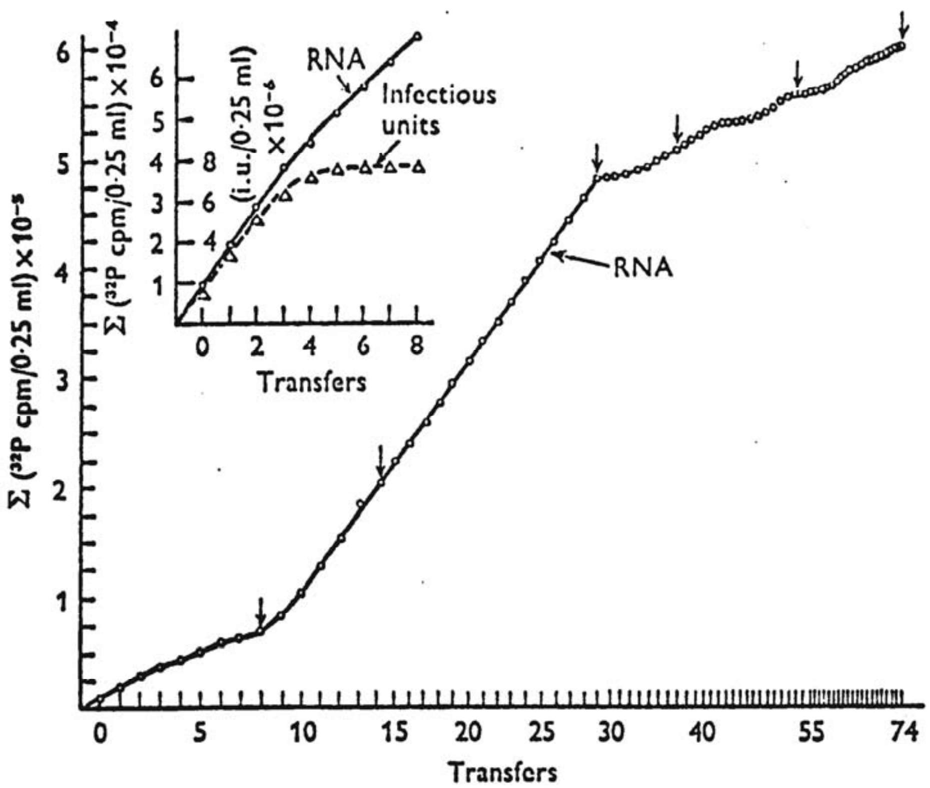
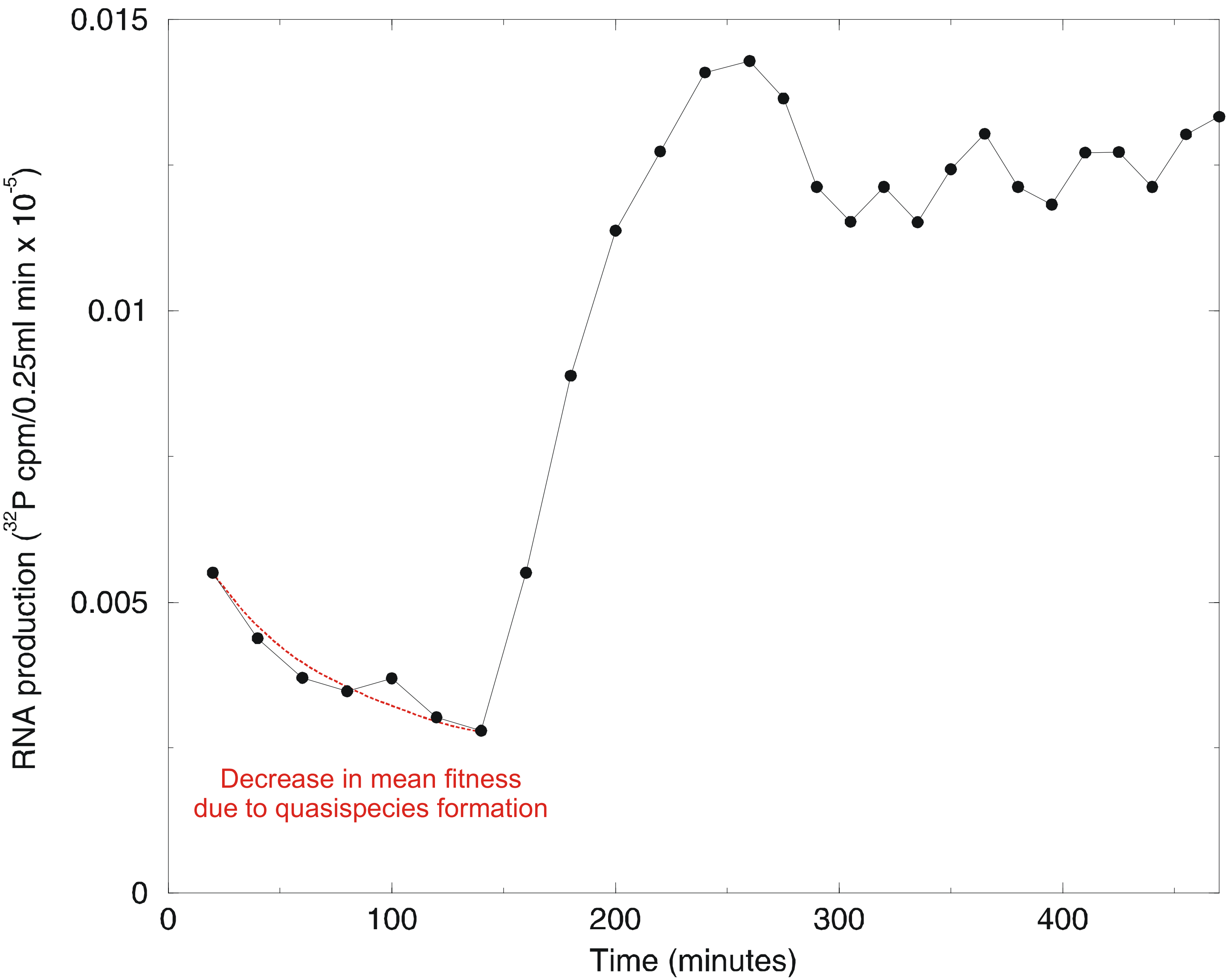


Fig. 9. Serial transfer experiment. Each 0·25 ml standard reaction mixture contained 40 μg of Qβ replicase and ^{32}P-UTP. The first reaction (0 transfer) was initiated by the addition of 0·2 μg ts-1 (temperature-sensitive RNA) and incubated at 35 °C for 20 min, whereupon 0·02 ml was drawn for counting and 0·02 ml was used to prime the second reaction (first transfer), and so on. After the first 13 reactions, the incubation periods were reduced to 15 min (transfers 14–29). Transfers 30–38 were incubated for 10 min. Transfers 39–52 were incubated for 7 min, and transfers 53–74 were incubated for 5 min. The arrows above certain transfers (0, 8, 14, 29, 37, 53, and 73) indicate where 0·001–0·1 ml of product was removed and used to prime reactions for sedimentation analysis on sucrose. The inset examines both infectious and total RNA. The results show that biologically competent RNA ceases to appear after the 4th transfer (Mills *et al.* 1967).



The increase in RNA production rate during a serial transfer experiment

Evolutionary design of RNA molecules

D.B.Bartel, J.W.Szostak, **In vitro *selection of RNA molecules that bind specific ligands***. Nature **346** (1990), 818-822

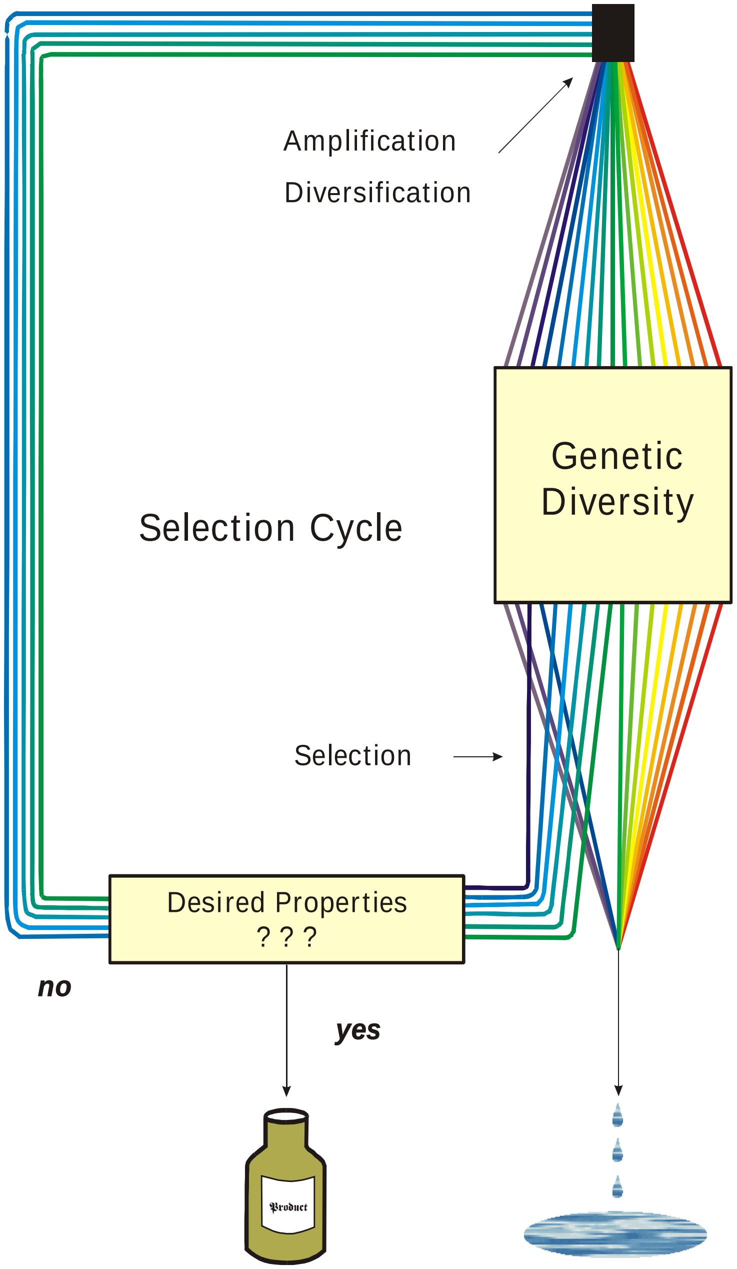
C.Tuerk, L.Gold, **SELEX - *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage* T4 *DNA polymerase***. Science **249** (1990), 505-510

D.P.Bartel, J.W.Szostak, ***Isolation of new ribozymes from a large pool of random sequences***. Science **261** (1993), 1411-1418

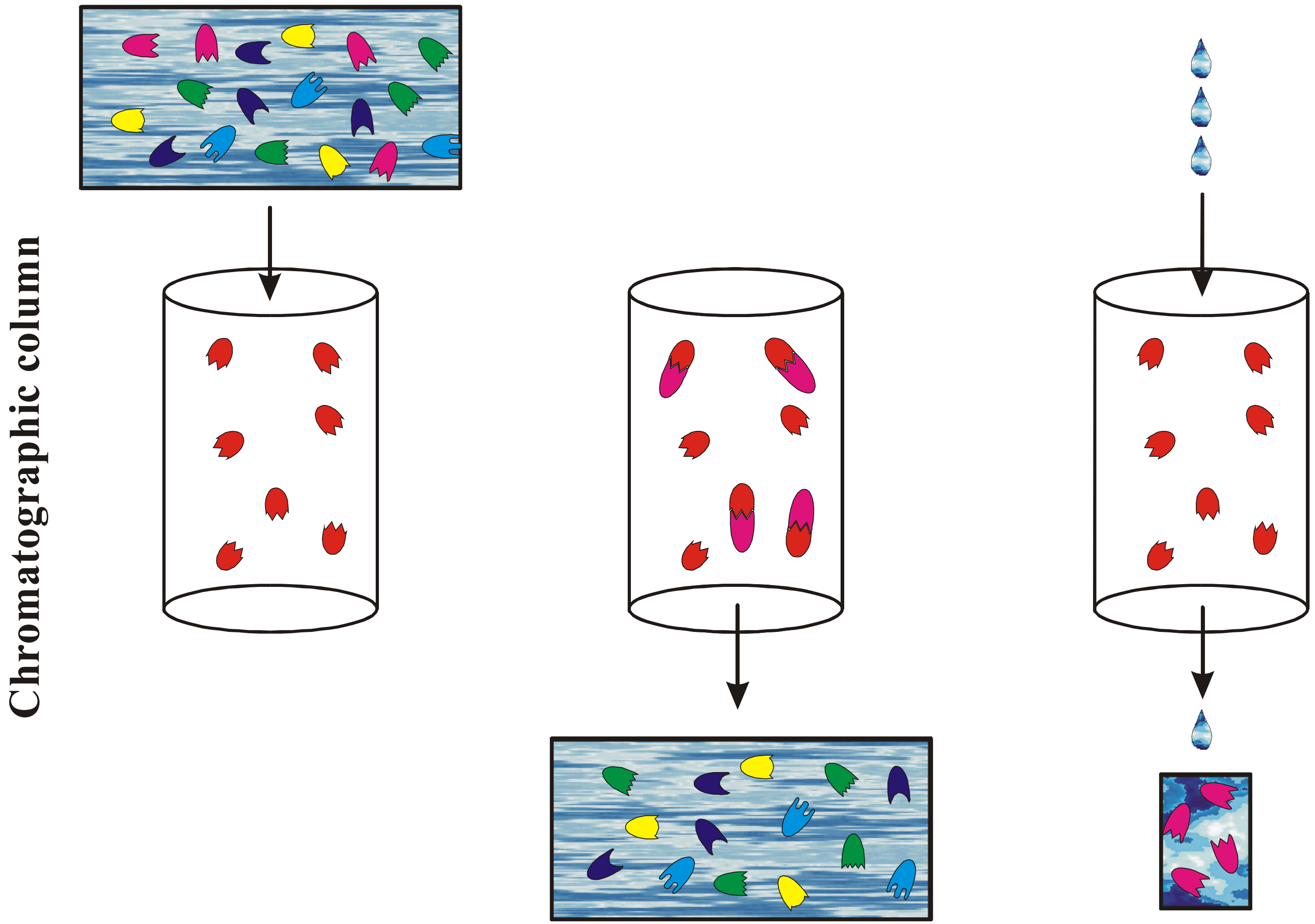
R.D.Jenison, S.C.Gill, A.Pardi, B.Poliski, ***High-resolution molecular discrimination by RNA***. Science **263** (1994), 1425-1429

Y. Wang, R.R.Rando, ***Specific binding of aminoglycoside antibiotics to RNA***. Chemistry & Biology **2** (1995), 281-290

Jiang, A. K. Suri, R. Fiala, D. J. Patel, ***Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex***. Chemistry & Biology **4** (1997), 35-50



Selection cycle used in applied molecular evolution to design molecules with predefined properties

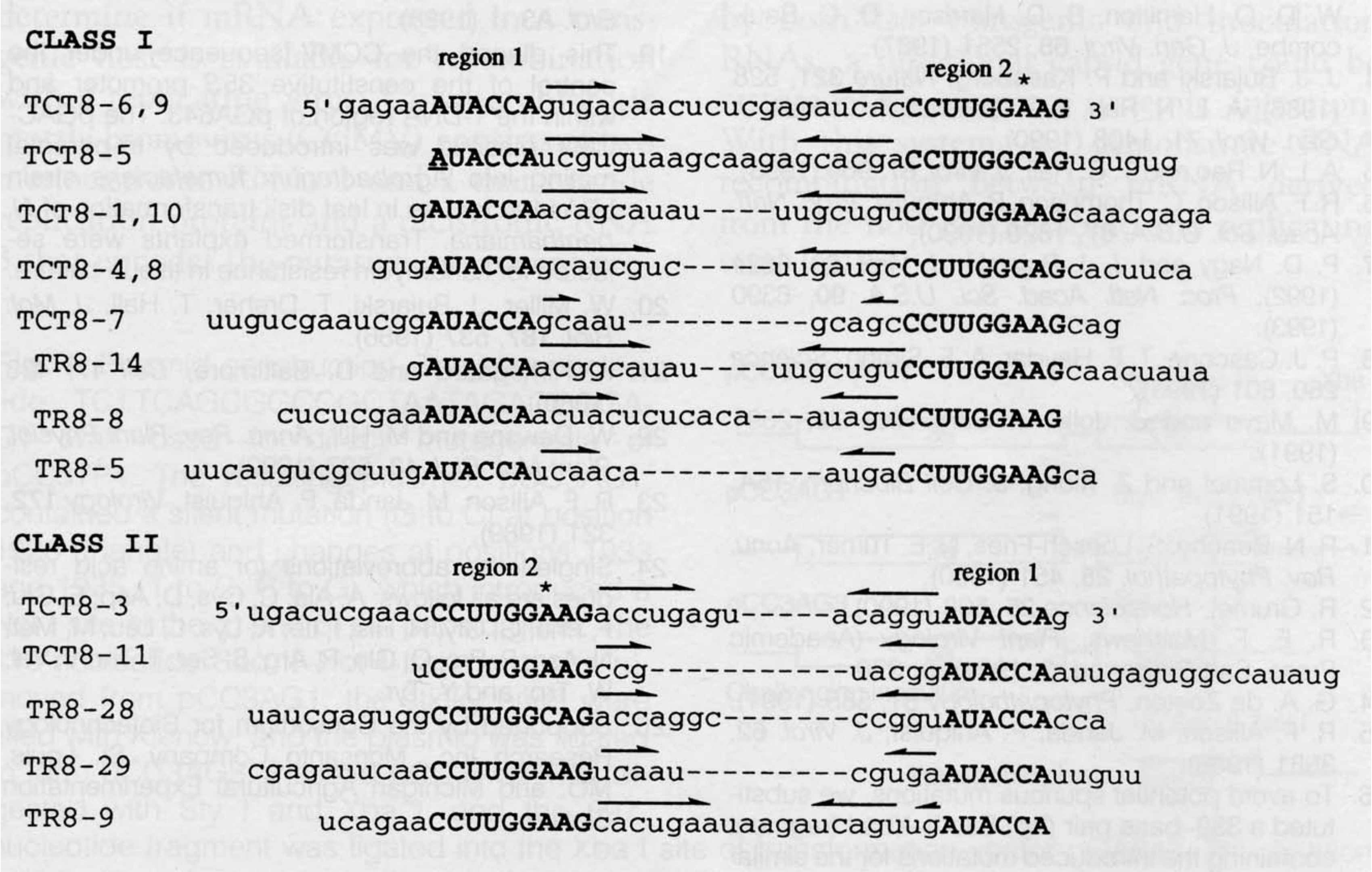


Retention of binders

Elution of binders

Chromatographic column

The SELEX technique for the evolutionary design of ***aptamers***



CLASS I

region 1 region 2 *

TCT8-6,9 5'gagaa**AUACCA**gugacaacucucgagaucac**CCUUGGAAG** 3'

TCT8-5 **AUACCA**ucguguaagcaagagcacga**CCUUGGCAG**ugugug

TCT8-1,10 g**AUACCA**acagcauau----uugcugu**CCUUGGAAG**caacgaga

TCT8-4,8 gug**AUACCA**gcaucguc----uugaugc**CCUUGGCAG**cacuuca

TCT8-7 uugucgaaucgg**AUACCA**gcaau----------gcagc**CCUUGGAAG**cag

TR8-14 g**AUACCA**acggcauau----uugcugu**CCUUGGAAG**caacuaua

TR8-8 cucucgaa**AUACCA**acuacucucaca---auagu**CCUUGGAAG**

TR8-5 uucaugucgcuug**AUACCA**ucaaca-----------auga**CCUUGGAAG**ca

CLASS II

region 2 * region 1

TCT8-3 5'ugacucgaac**CCUUGGAAG**accugagu-----acaggu**AUACCA**g 3'

TCT8-11 u**CCUUGGAAG**ccg-----------uacgg**AUACCA**auugaguggccauaug

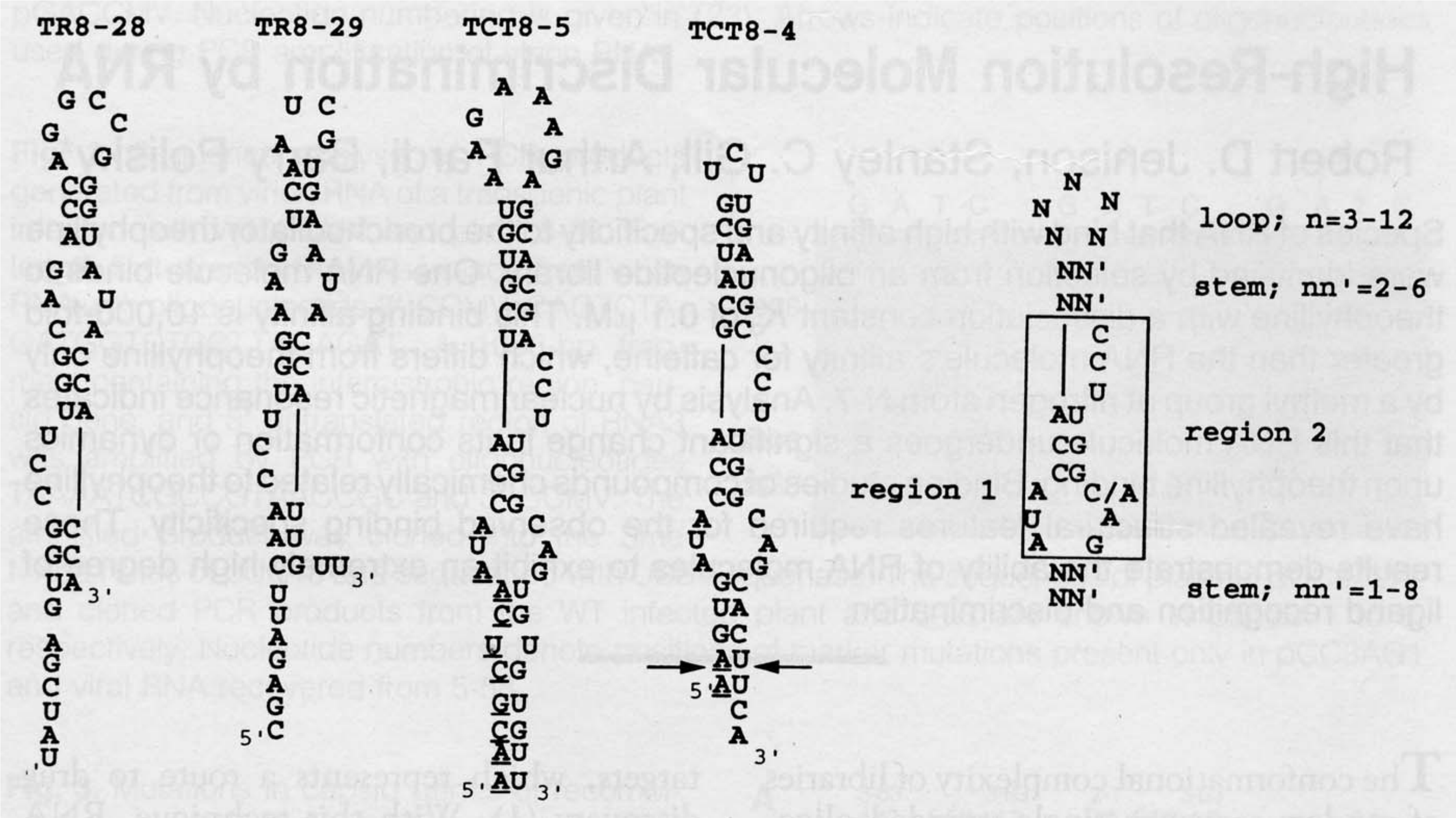
TR8-28 uaucgagugg**CCUUGGCAG**accaggc-------ccggu**AUACCA**cca

TR8-29 cgagauucaa**CCUUGGAAG**ucaau---------cguga**AUACCA**uuguu

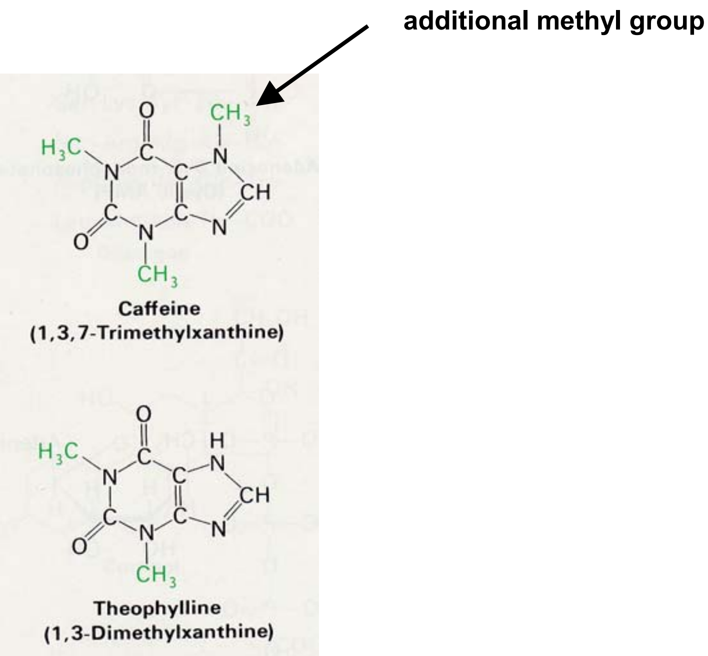
TR8-9 ucagaa**CCUUGGAAG**cacugaauaagaucaguug**AUACCA**

Sequences of aptamers binding theophyllin, caffeine, and related compounds

R.D.Jenison, S.C.Gill, A.Pardi, B.Poliski, ***High-resolution molecular discrimination by RNA***. Science **263** (1994), 1425-1429



Secondary structures of aptamers binding theophyllin, caffeine, and related compounds



Caffeine
(1,3,7-Trimethylxanthine)

Theophylline
(1,3-Dimethylxanthine)

Dissociation constants and specificity of theophylline, caffeine, and related derivatives of uric acid for binding to a discriminating aptamer TCT8-4

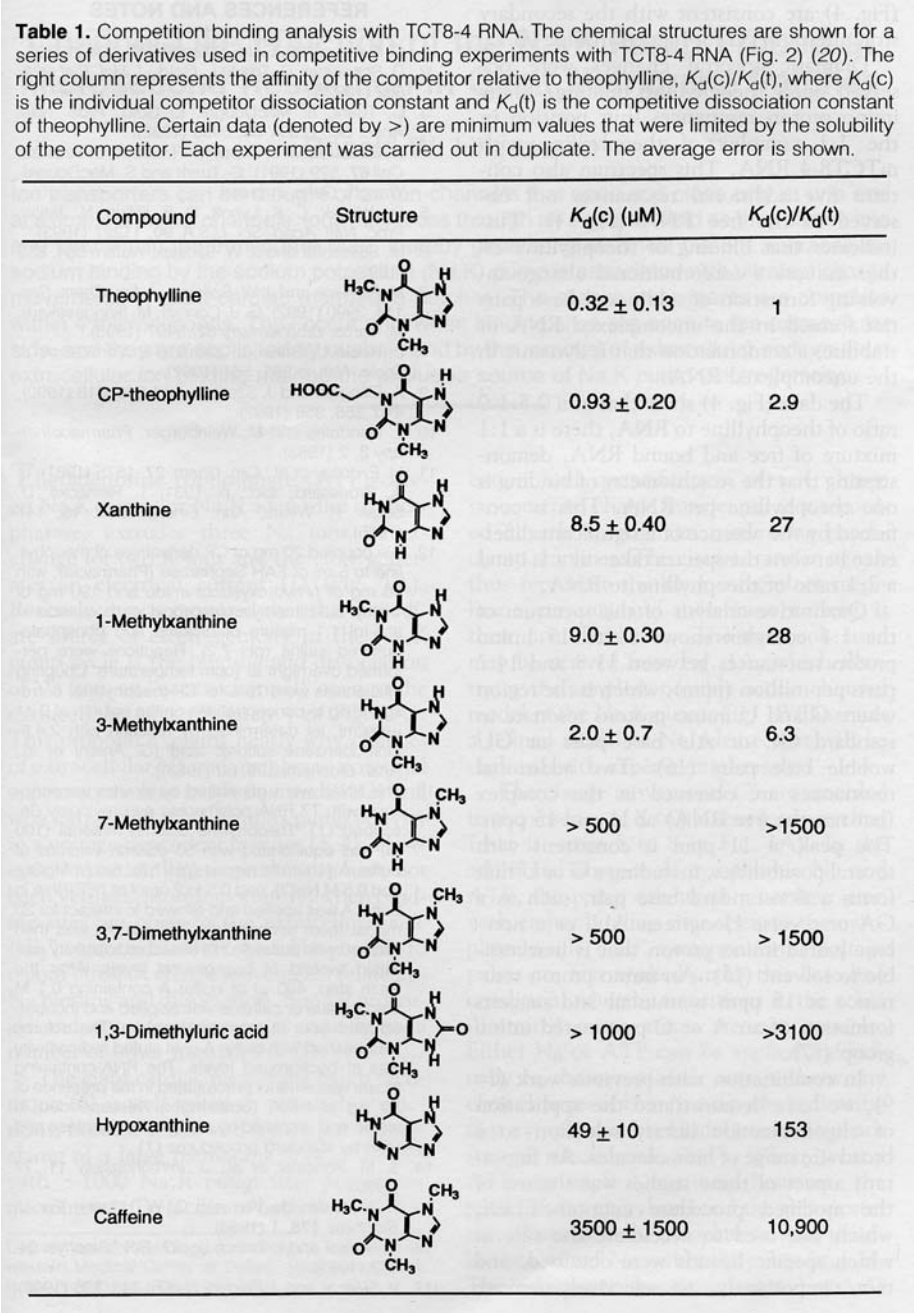


Table 1. Competition binding analysis with TCT8-4 RNA. The chemical structures are shown for a series of derivatives used in competitive binding experiments with TCT8-4 RNA (Fig. 2) (*20*). The right column represents the affinity of the competitor relative to theophylline, $K_d(c)/K_d(t)$, where $K_d(c)$ is the individual competitor dissociation constant and $K_d(t)$ is the competitive dissociation constant of theophylline. Certain data (denoted by >) are minimum values that were limited by the solubility of the competitor. Each experiment was carried out in duplicate. The average error is shown.

Compound	Structure	K_d(c) (μM)	$K_d(c)/K_d(t)$
Theophylline		0.32 ± 0.13	1
CP-theophylline		0.93 ± 0.20	2.9
Xanthine		8.5 ± 0.40	27
1-Methylxanthine		9.0 ± 0.30	28
3-Methylxanthine		2.0 ± 0.7	6.3
7-Methylxanthine		> 500	>1500
3,7-Dimethylxanthine		> 500	> 1500
1,3-Dimethyluric acid		> 1000	>3100
Hypoxanthine		49 ± 10	153
Caffeine		3500 ± 1500	10,900

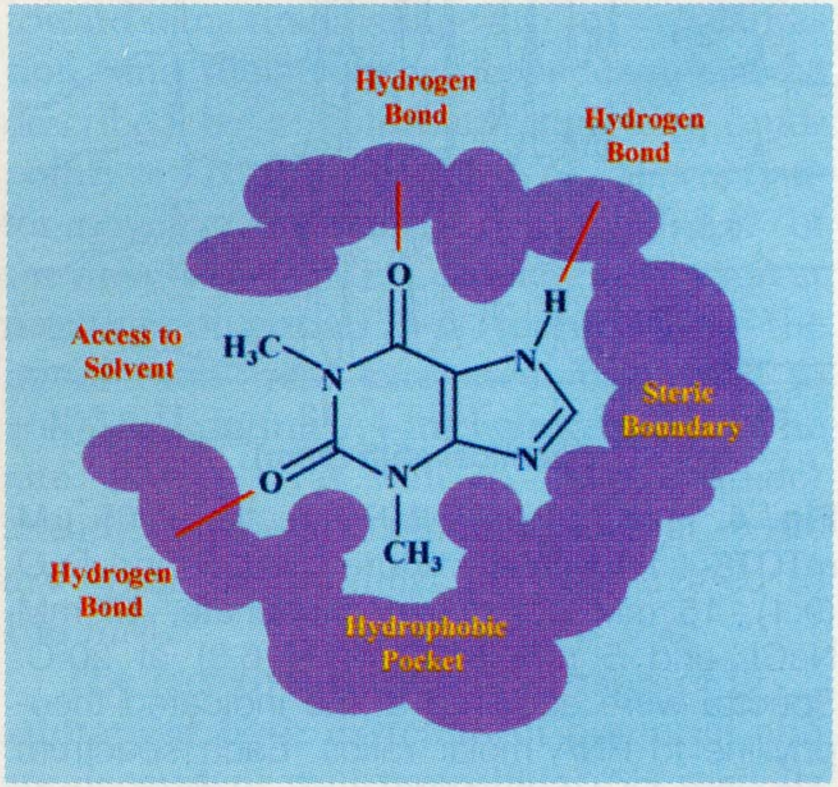
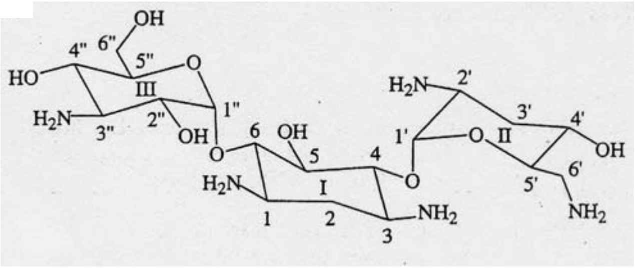
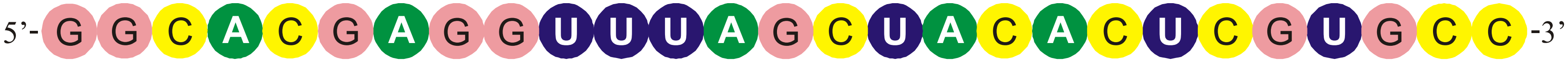


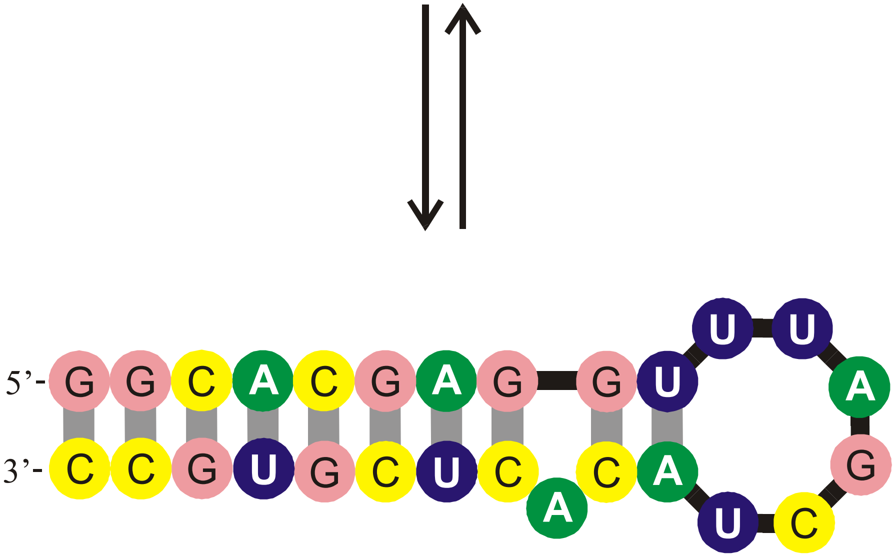
Fig. 3. Schematic representation of the RNA (purple) binding site for theophylline (blue).

Schematic drawing of the aptamer binding site for the theophylline molecule



tobramycin

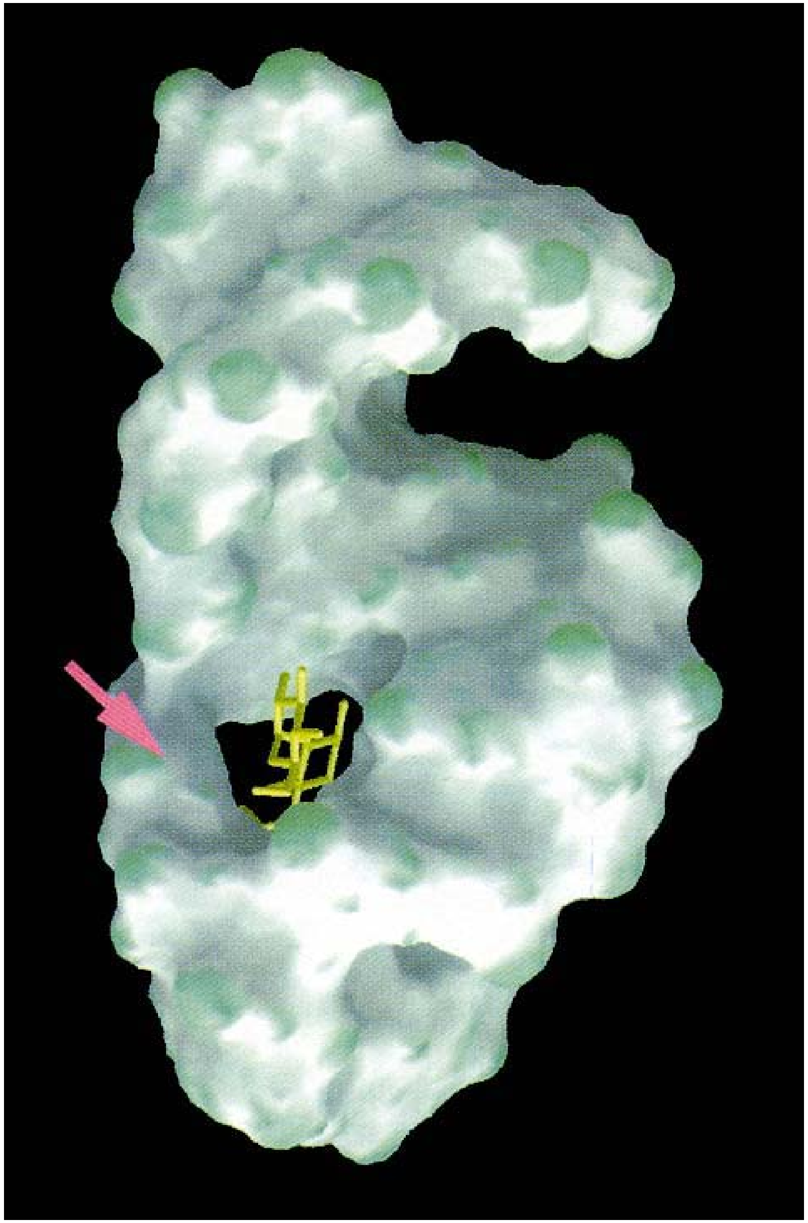




RNA aptamer

Formation of secondary structure of the tobramycin binding RNA aptamer

L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, ***Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex.*** Chemistry & Biology **4**:35-50 (1997)



The three-dimensional structure of the tobramycin aptamer complex

L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, Chemistry & Biology **4**:35-50 (1997)

Universität Wien

Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)

Projects No. 09942, 10578, 11065, 13093
13887, and 14898

Jubiläumsfonds der Österreichischen Nationalbank

Project No. Nat-7813

European Commission: Project No. EU-980189

Siemens AG, Austria

The Santa Fe Institute and the Universität Wien

The software for producing RNA movies was developed by
Robert Giegerich and coworkers at the Universität Bielefeld

Coworkers

Universität Wien

Walter Fontana, Santa Fe Institute, NM

Christian Reidys, Christian Forst, Los Alamos National Laboratory, NM

Peter Stadler, **Bärbel Stadler,** Universität Leipzig, GE

Ivo L.Hofacker, Christoph Flamm, Universität Wien, AT

Andreas Wernitznig, **Michael Kospach,** Universität Wien, AT
Ulrike Langhammer, Ulrike Mückstein, Stefanie Widder
Jan Cupal, Kurt Grünberger, Andreas Svrček-Seiler, Stefan Wuchty,
Stefan Bernhardt, Andreas De Stefani

Ulrike Göbel, Institut für Molekulare Biotechnologie, Jena, GE
Walter Grüner, Stefan Kopp, Jaqueline Weber

Web-Page for further information:

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