Evolution of RNA Molecules

From Neutral Networks of Structures to Complex Interaction Patterns

Peter Schuster

Institut für Theoretische Chemie der Universität Wien, Austria

and the Santa Fe Institute, NM

Collectives formation and specialization in biological and social systems

Santa Fe, 20.– 22.04.2005

Web-Page for further information:

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

5'-End 3'-End

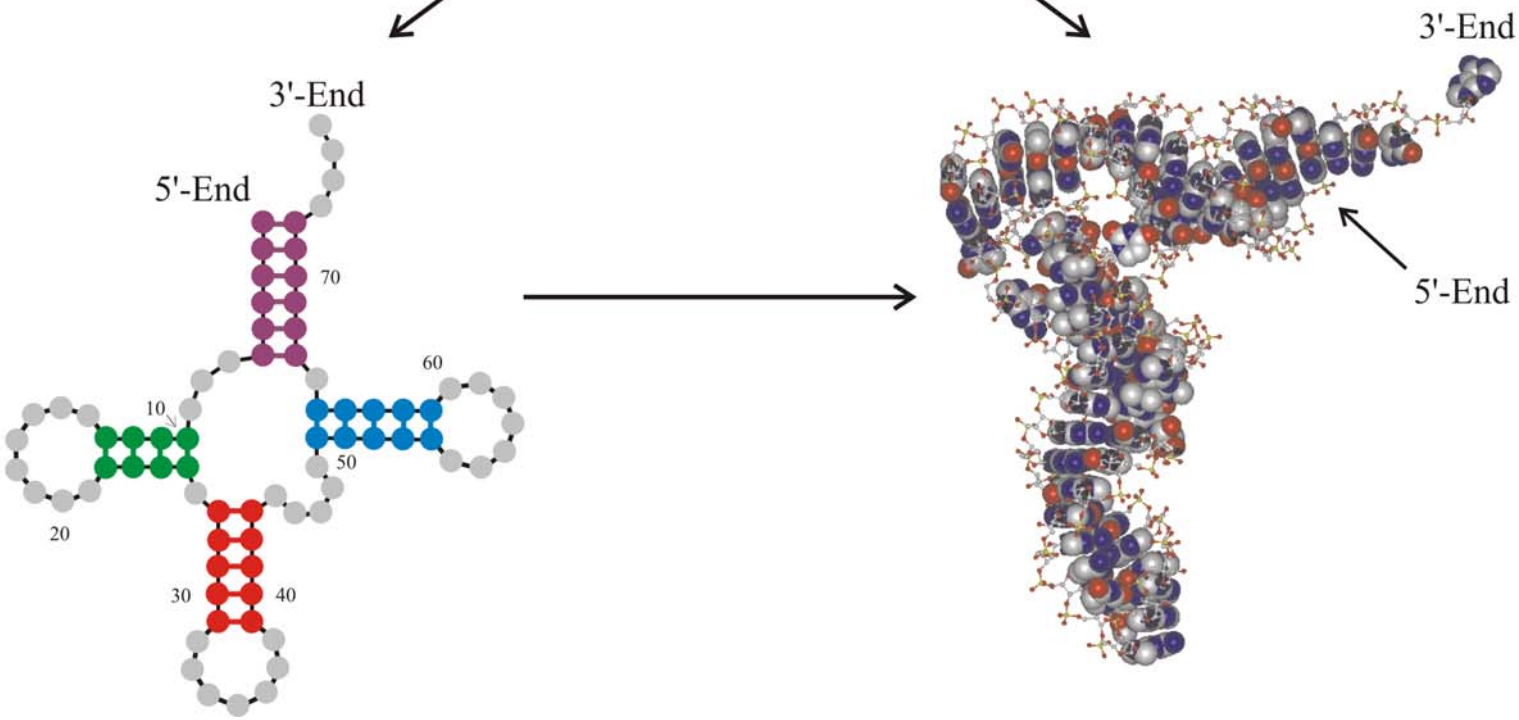
GCGGAUUUAGCUCAGDDGGGAGAGCMCCAGACUGAAYAUCUGGAGMUCCUGUGTPCGAUCCACAGAAUUCGCACCA

Biochemical and chemical probing

Structure prediction

Crystallography

NMR, FRET, ...



1. Folding and inverse folding of RNA
2. Neutral networks
3. Darwinian evolution of RNA
4. Learning by the Darwinian mechanism
5. Folding kinetics and metastable structures
6. Intersections and conformational switches

1. **Folding and inverse folding of RNA**
2. Neutral networks
3. Darwinian evolution of RNA
4. Learning by the Darwinian mechanism
5. Folding kinetics and metastable structures
6. Intersections and conformational switches

RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

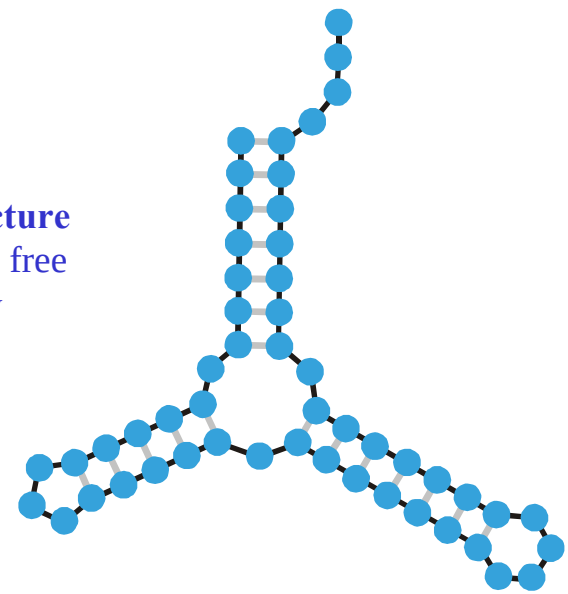
RNA folding:

Structural biology, spectroscopy of biomolecules, understanding **molecular function**

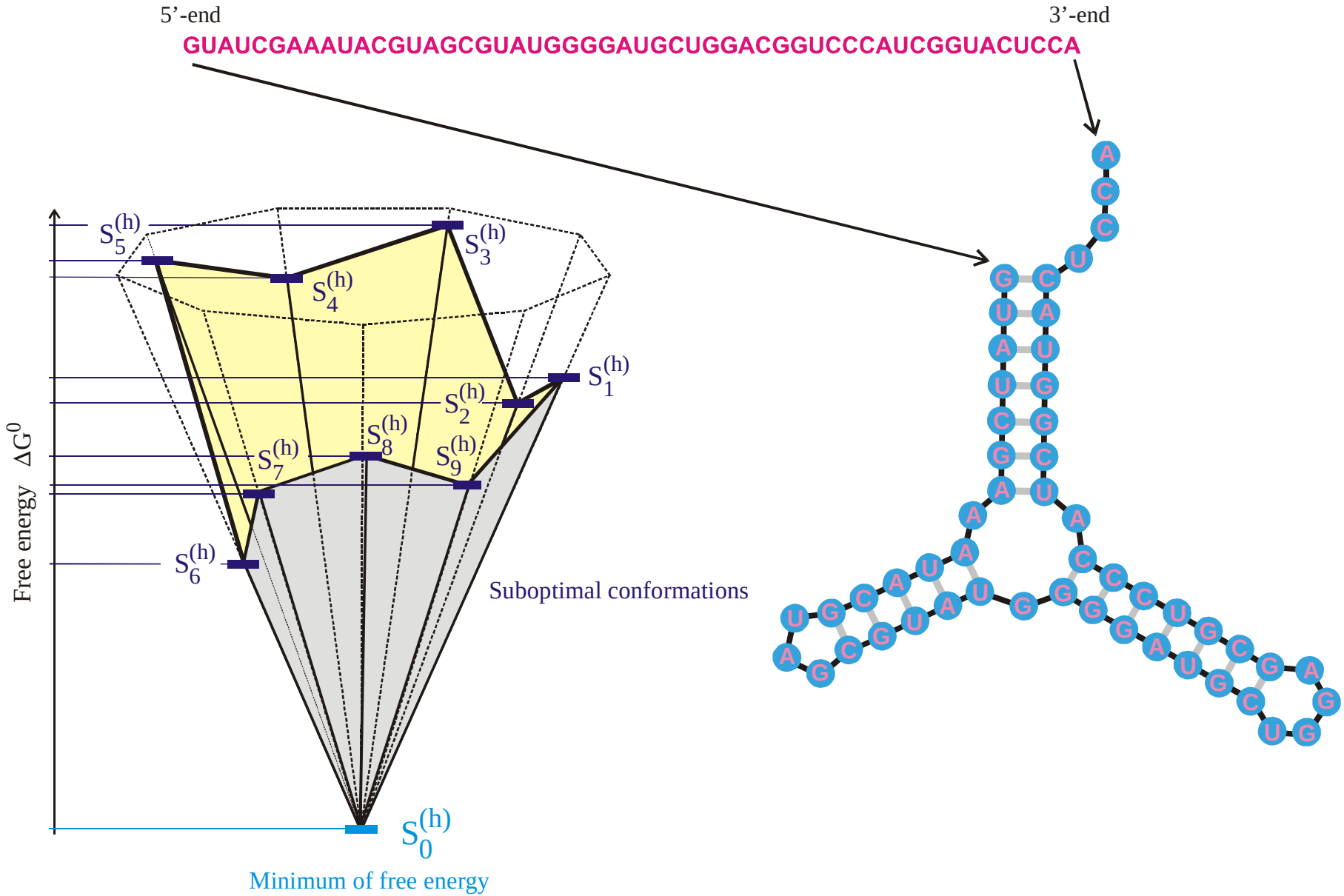
Biophysical chemistry: thermodynamics and kinetics

Empirical parameters

RNA structure of minimal free energy



One sequence – one structure problem



The minimum free energy structures on a discrete space of conformations

RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

RNA folding:

Structural biology, spectroscopy of biomolecules, understanding **molecular function**

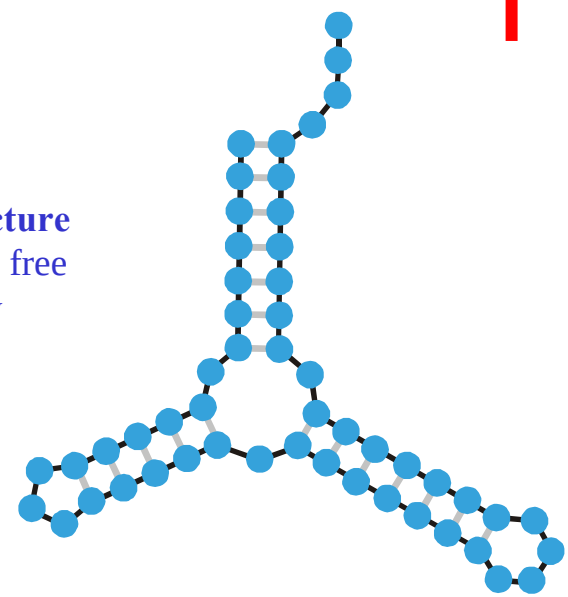
Iterative determination of a sequence for the given secondary structure

Inverse Folding Algorithm

Inverse folding of RNA:

Biotechnology, **design of biomolecules** with predefined structures and functions

RNA structure of minimal free energy



Sequence, structure, and design

1. Folding and inverse folding of RNA

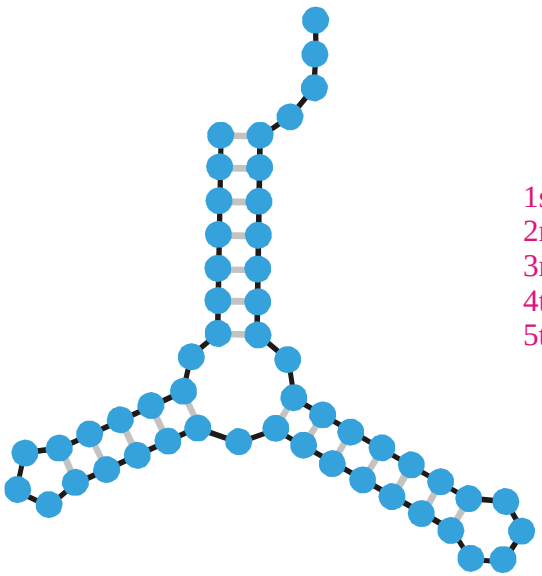
2. Neutral networks

3. Darwinian evolution of RNA

4. Learning by the Darwinian mechanism

5. Folding kinetics and metastable structures

6. Intersections and conformational switches



Minimum free energy criterion

1st
2nd
3rd trial
4th
5th

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

UGGUUACGCGUUGGGGUAACGAAGAUUCCGAGAGGAGUUUAGUGACUAGAGG

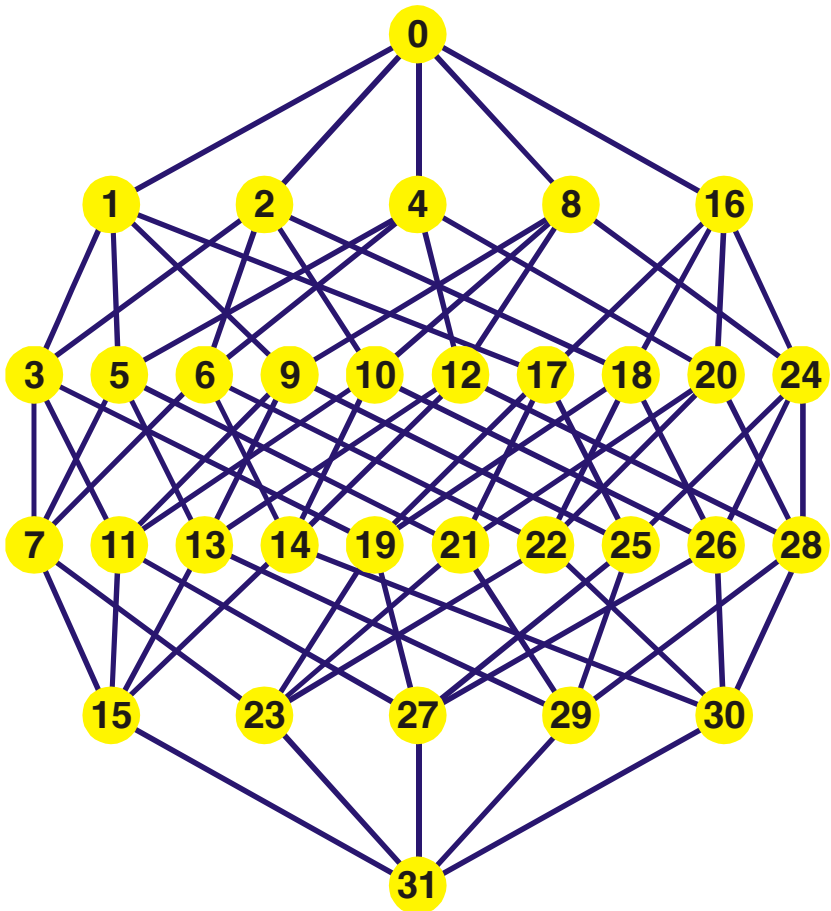
CUUCUUGAGCUAGUACCUAGUCGGAUAGGAUUUCCUAUCUCCAGGGAGGAUG

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.



Hypercube of dimension n = 5

Mutant class

0

1

2

3

4

5

Binary sequences are encoded by their decimal equivalents:

C = 0 and G = 1, for example,

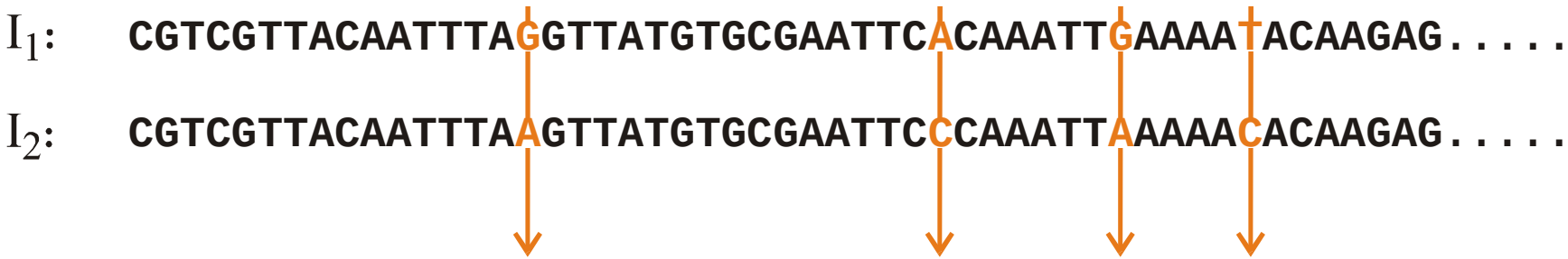
"0" ≡ 00000 = CCCCC,

"14" ≡ 01110 = CGGGC,

"29" ≡ 11101 = GGGCG, etc.

Decimal coding of binary sequences

Sequence space of binary sequences of chain lenght n = 5



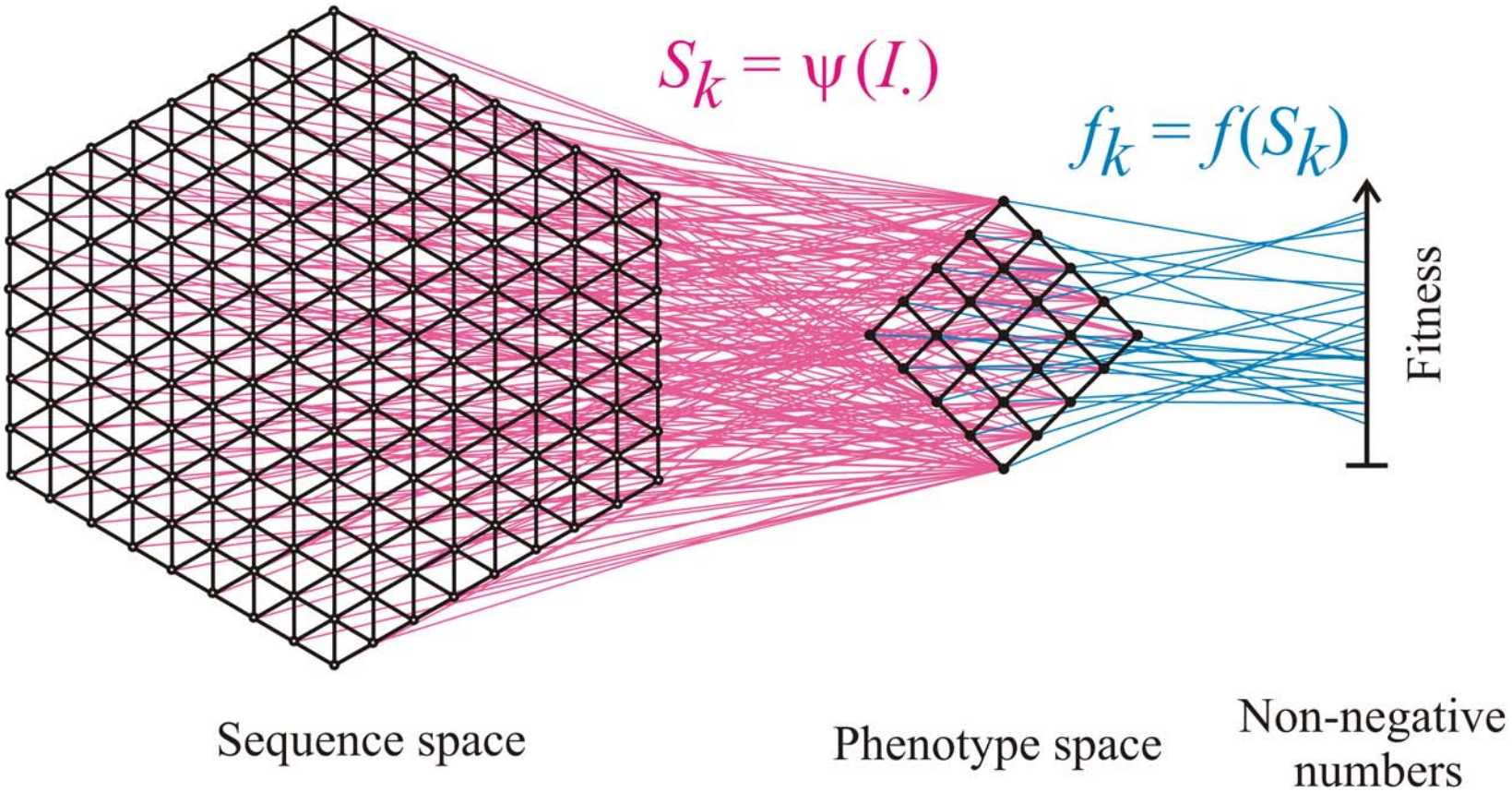
Hamming distance $d_H(I_1,I_2) = 4$

(i) $d_H(I_1,I_1) = 0$

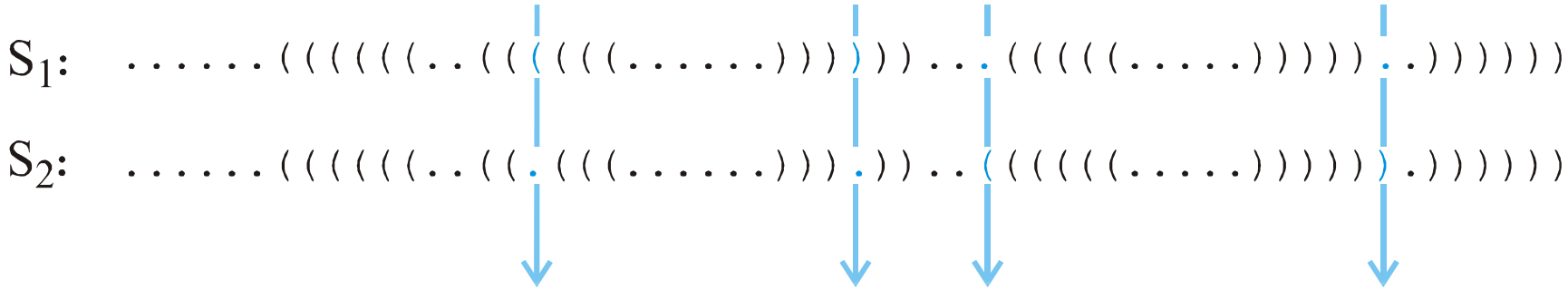
(ii) $d_H(I_1,I_2) = d_H(I_2,I_1)$

(iii) $d_H(I_1,I_3) \leq d_H(I_1,I_2) + d_H(I_2,I_3)$

The Hamming distance between sequences induces a metric in sequence space



Mapping from sequence space into structure space and into function



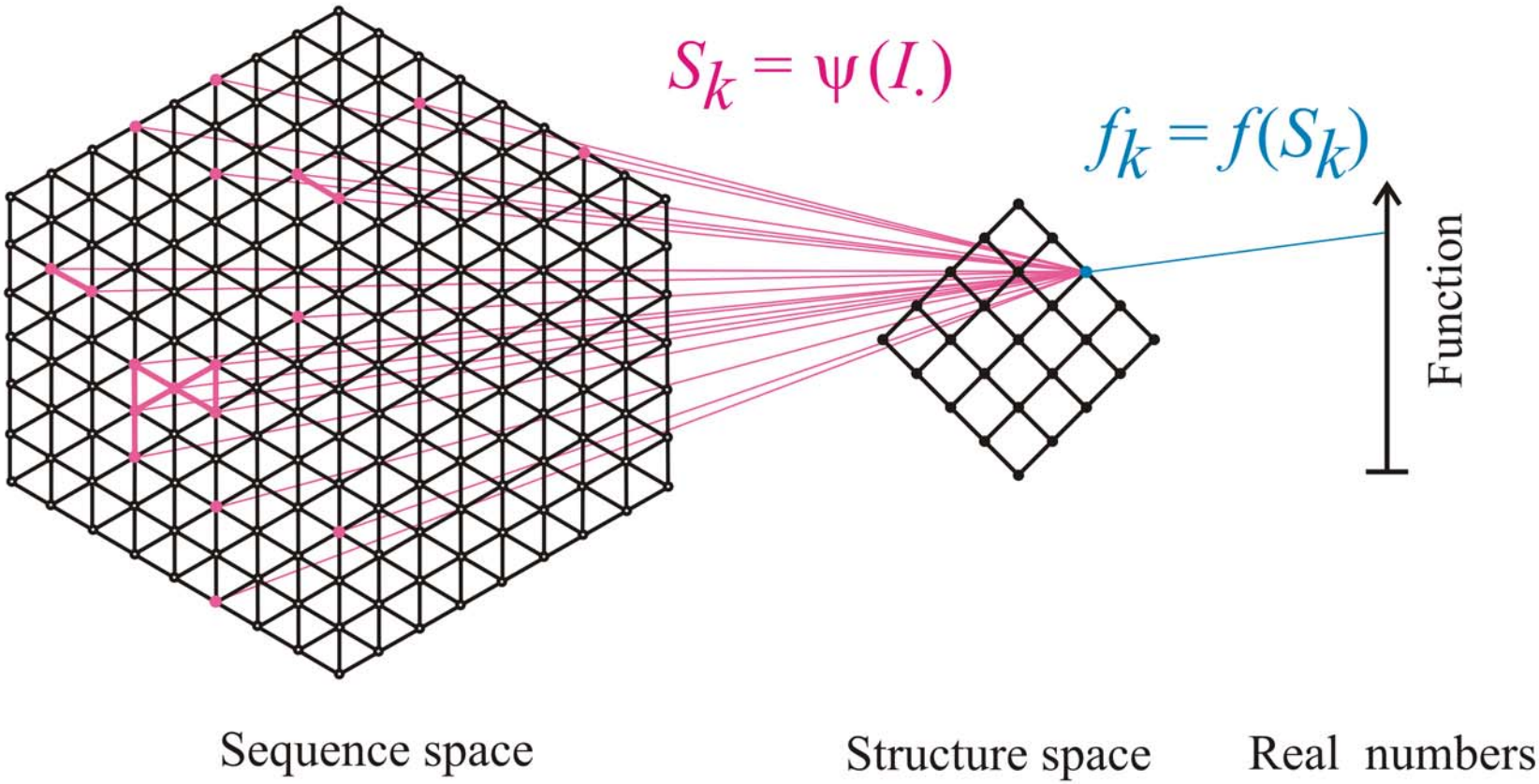
Hamming distance $d_H(S_1,S_2) = 4$

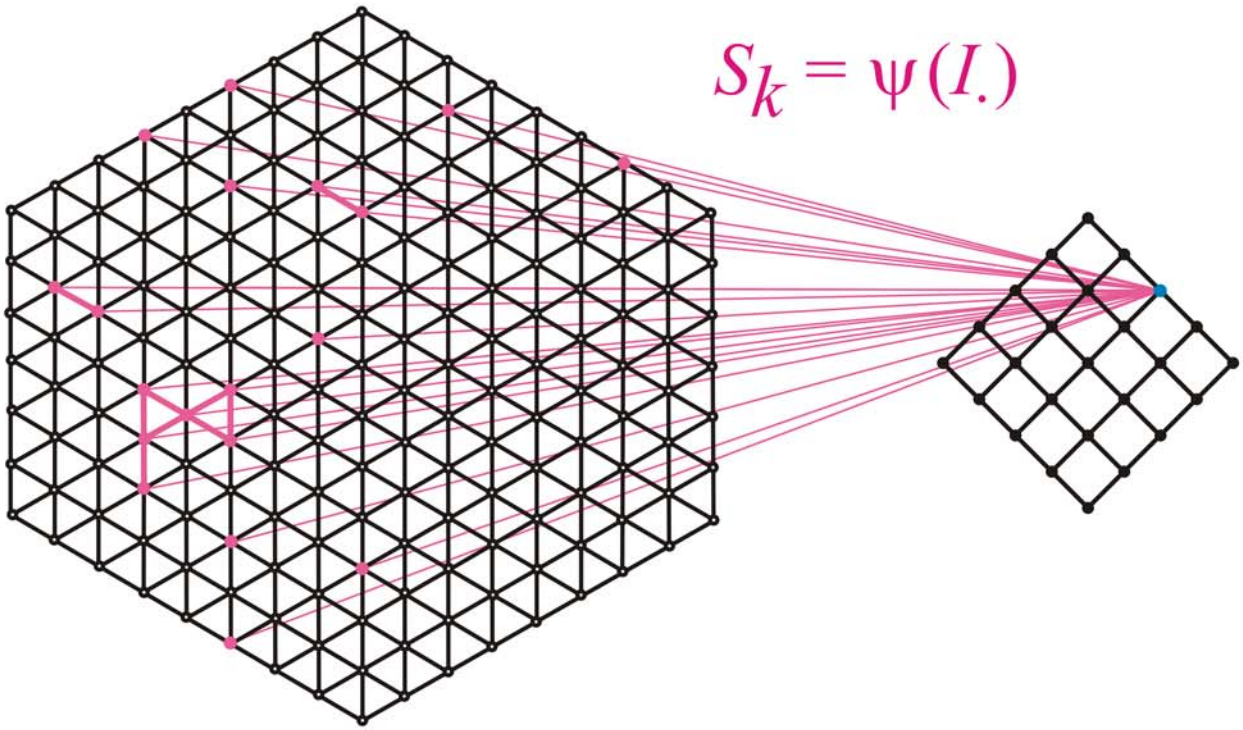
(i) $d_H(S_1,S_1) = 0$

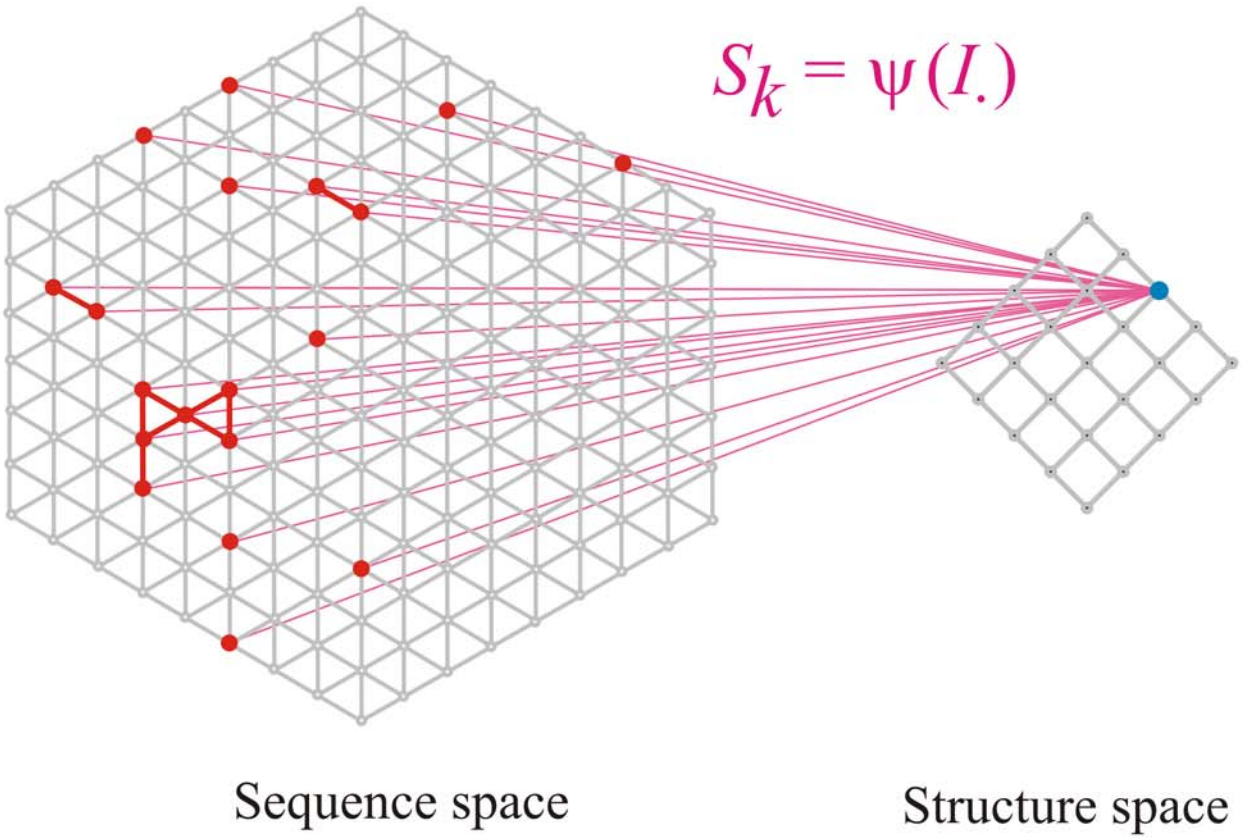
(ii) $d_H(S_1,S_2) = d_H(S_2,S_1)$

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

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


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


$S_k = \psi(I_.)$
Sequence space
Structure space



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

Space of sequences: $\mathcal{I} = \{I_1, I_2, I_3, I_4, \ldots, I_N\}$; Hamming metric

Space of structures: $S = \{S_1, S_2, S_3, S_4, \ldots, S_M\}$; metric (not required)

$$N >> M$$

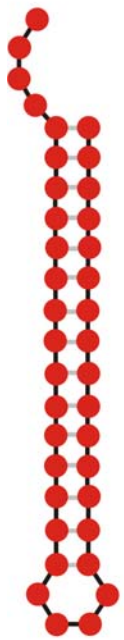
$$\psi(I_j) = S_k$$

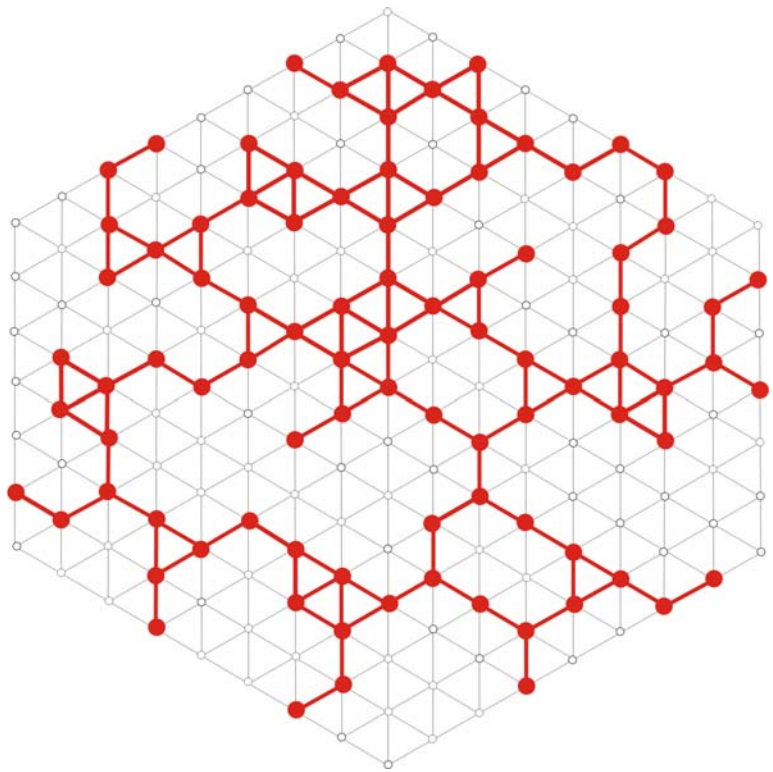
Neutral network: $\mathbf{G}_k = \psi^{-1}(S_k) \doteq \{ I_j \mid \psi(I_j) = S_k \}$

A mapping ψ and its inversion

Properties of RNA sequence to secondary structure mapping

1. More sequences than structures
2. Few common versus many rare structures
3. Shape space covering of common structures
4. Neutral networks of common structures are connected

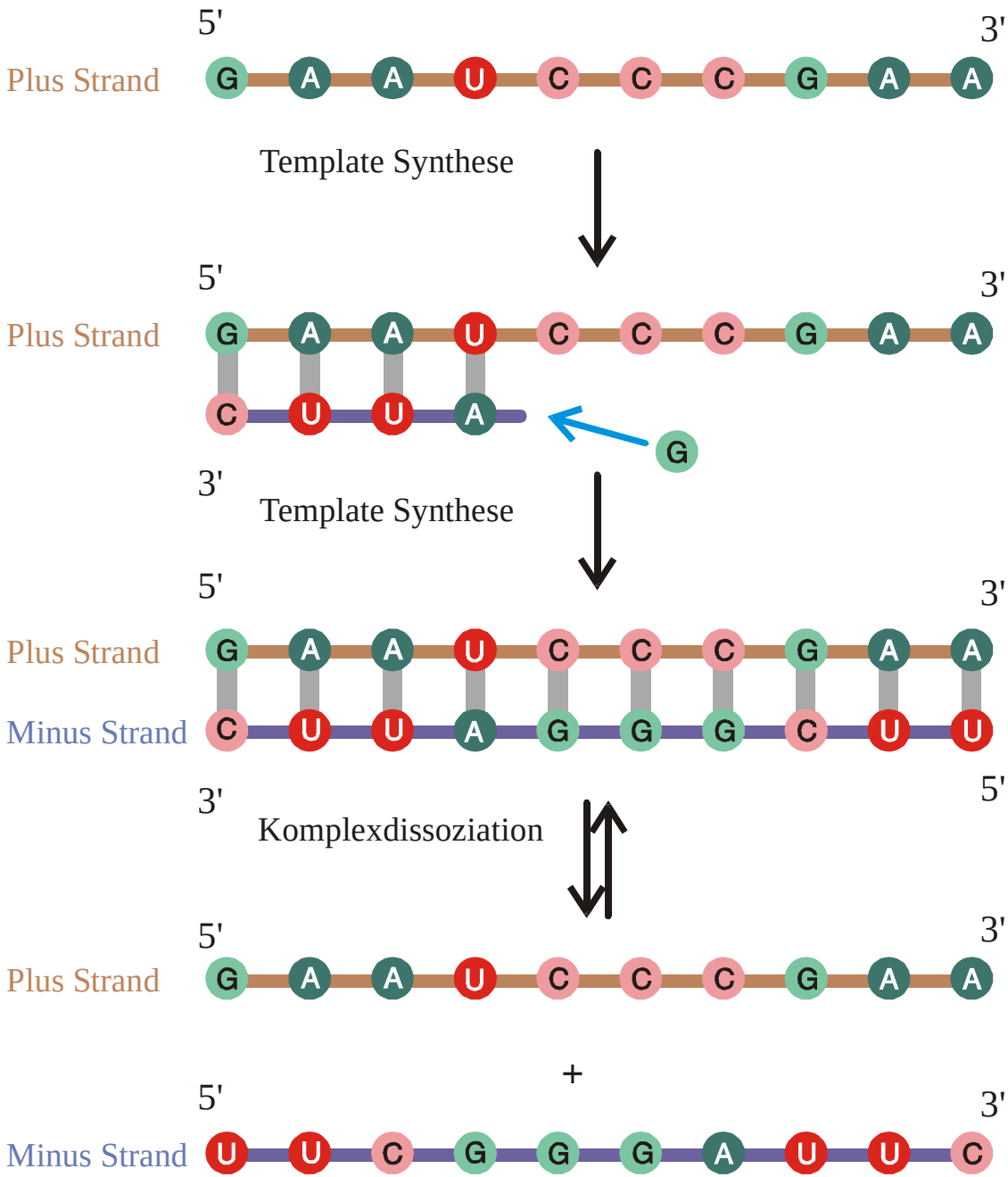


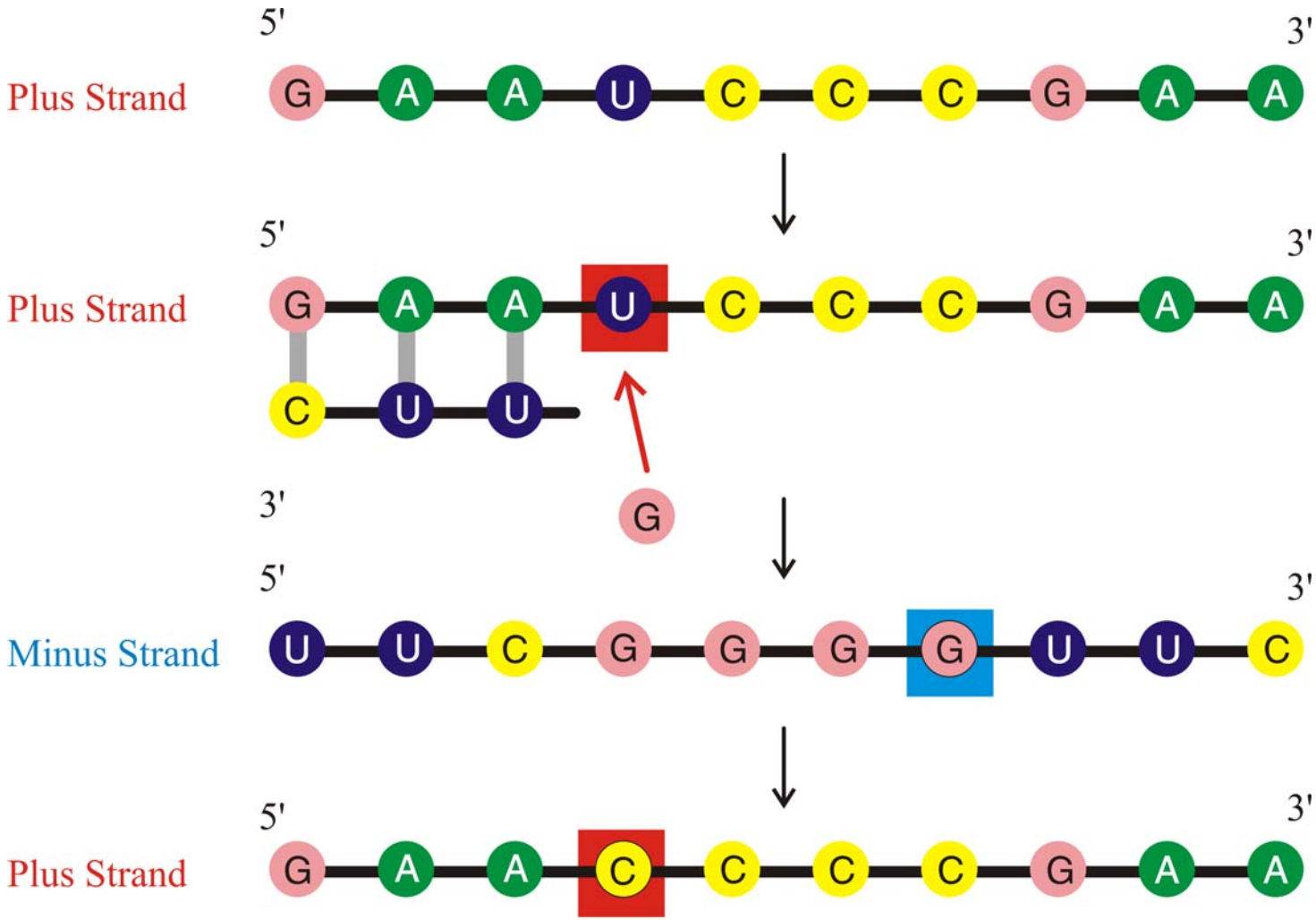


1. Folding and inverse folding of RNA
2. Neutral networks
3. **Darwinian evolution of RNA**
4. Learning by the Darwinian mechanism
5. Folding kinetics and metastable structures
6. Intersections and conformational switches

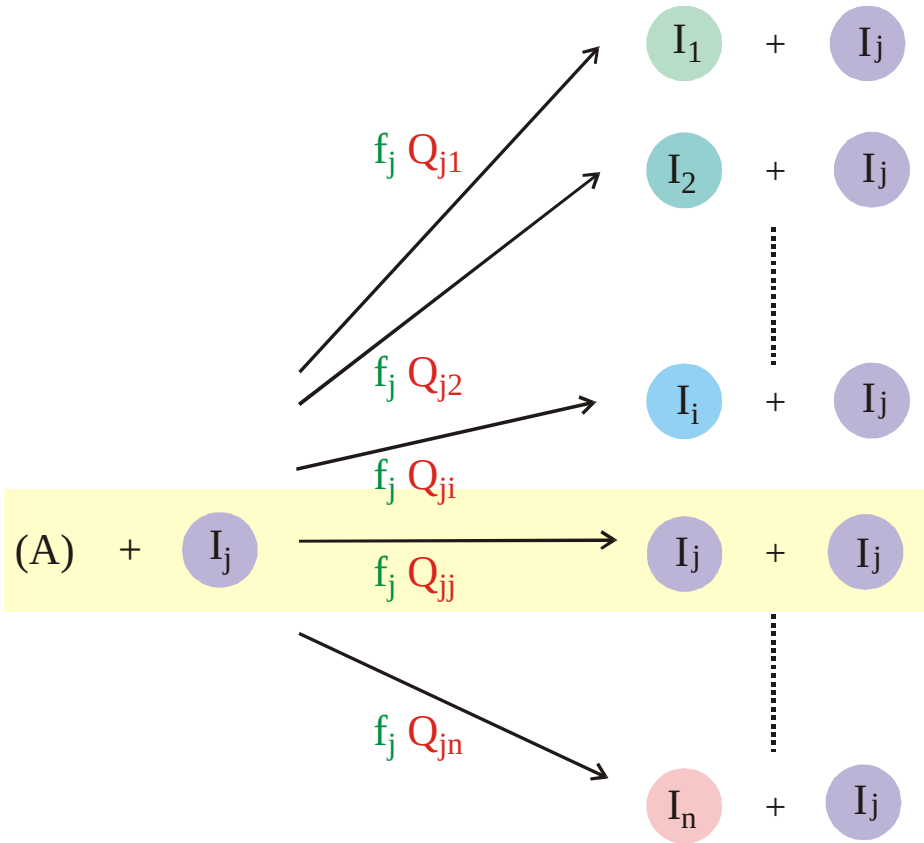
Copying of single-strand RNA-molecules:

Plus-Minus-Replication





Variation of the RNA sequence through copying errors



$$dx_i / dt = \Sigma_j f_j Q_{ji} x_j - x_i \Phi$$

$$\Phi = \Sigma_j f_j x_i ; \quad \Sigma_j x_j = 1 ; \quad \Sigma_i Q_{ij} = 1$$

$$[I_i] = x_i \geq 0 ; \quad i = 1,2,...,n ;$$

$$[A] = a = \text{constant}$$

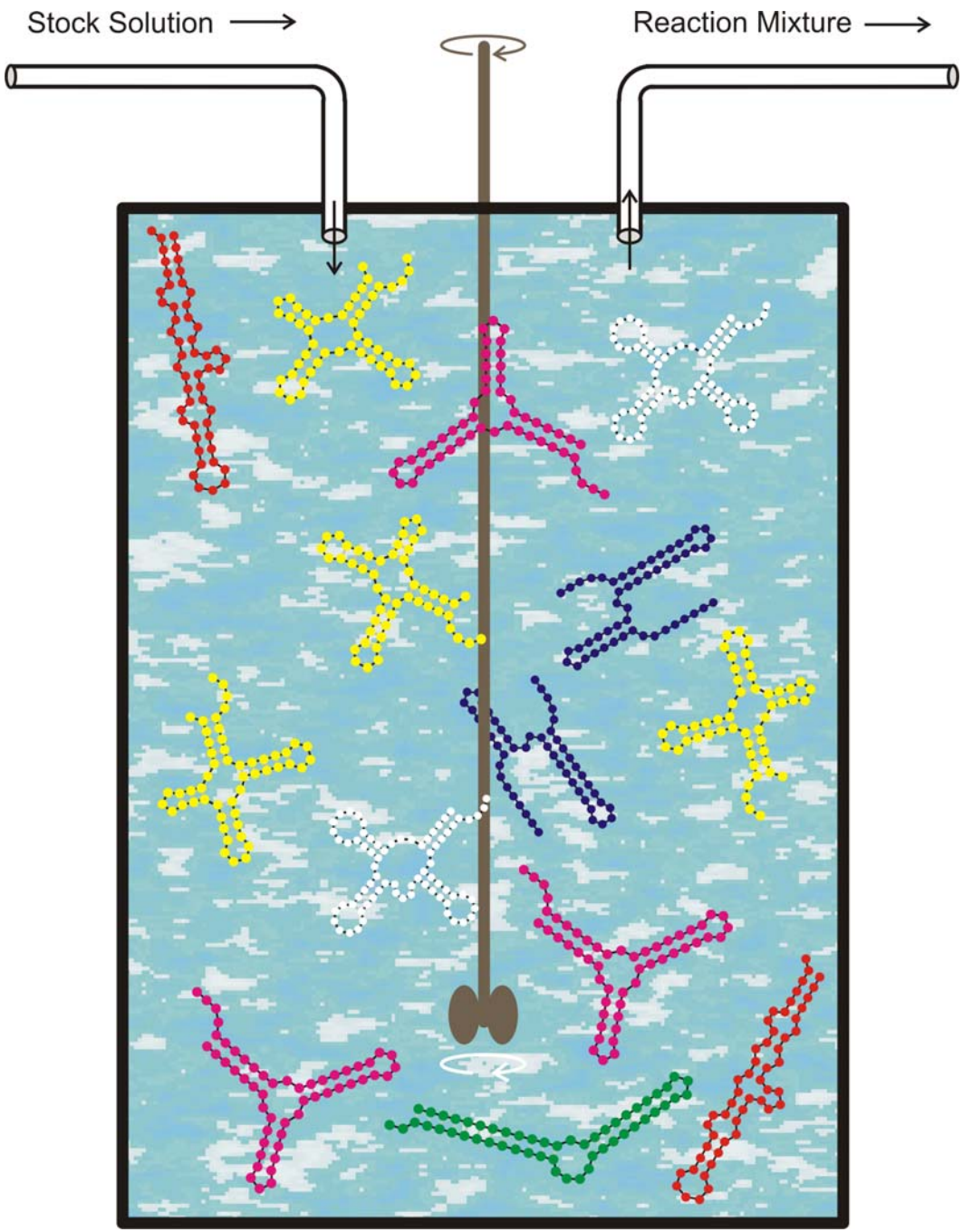
$$Q_{ij} = (1-p)^{\ell - d(i,j)} \, p^{d(i,j)}$$

p ..... Error rate per digit

ℓ ..... Chain length of the polynucleotide

d(i,j) .... Hamming distance between I_i and I_j

Chemical kinetics of replication and mutation as parallel reactions



Replication rate constant:

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

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

Selection constraint:

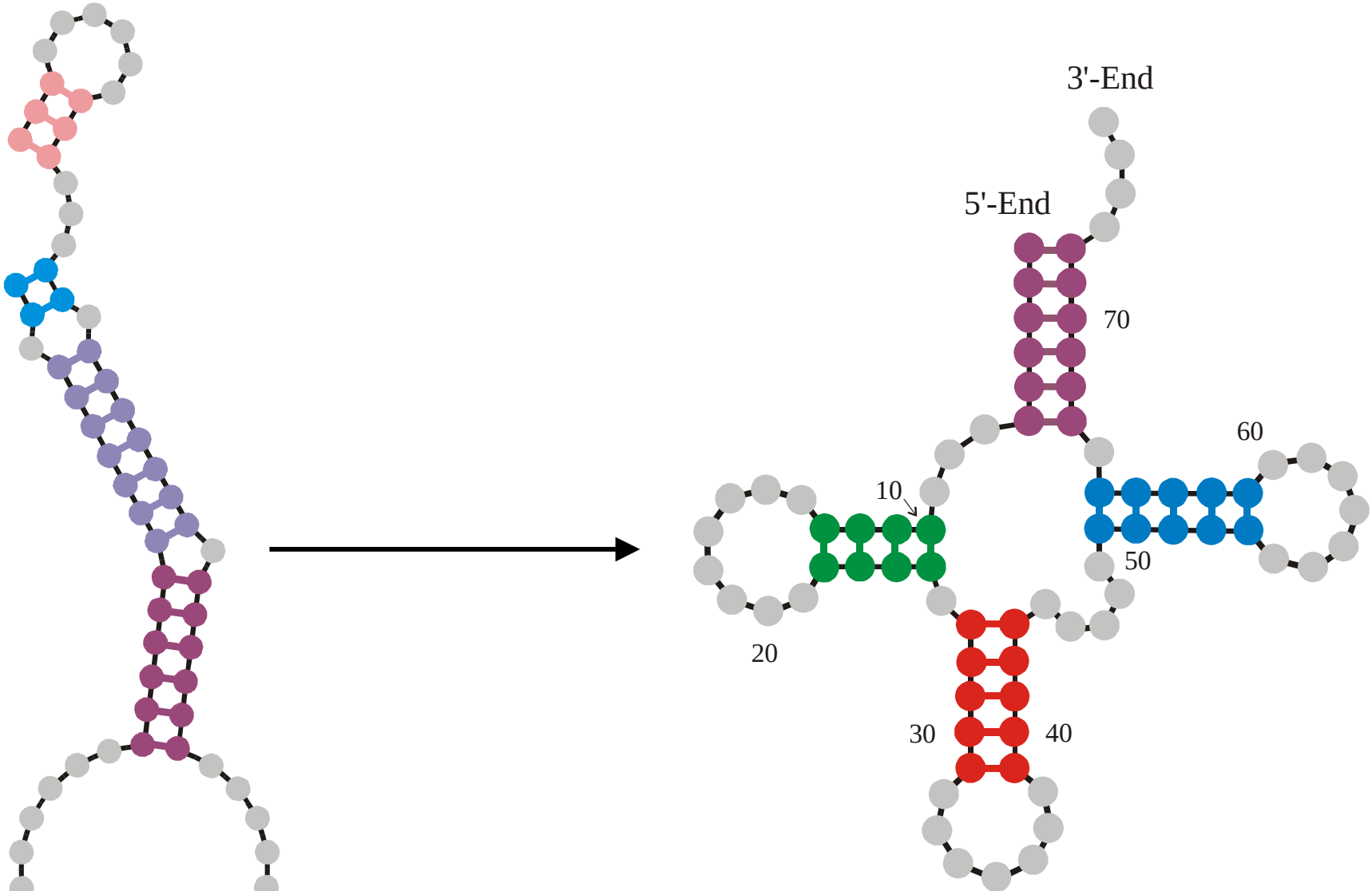
Population size, N = # RNA molecules, is controlled by the flow

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

Mutation rate:

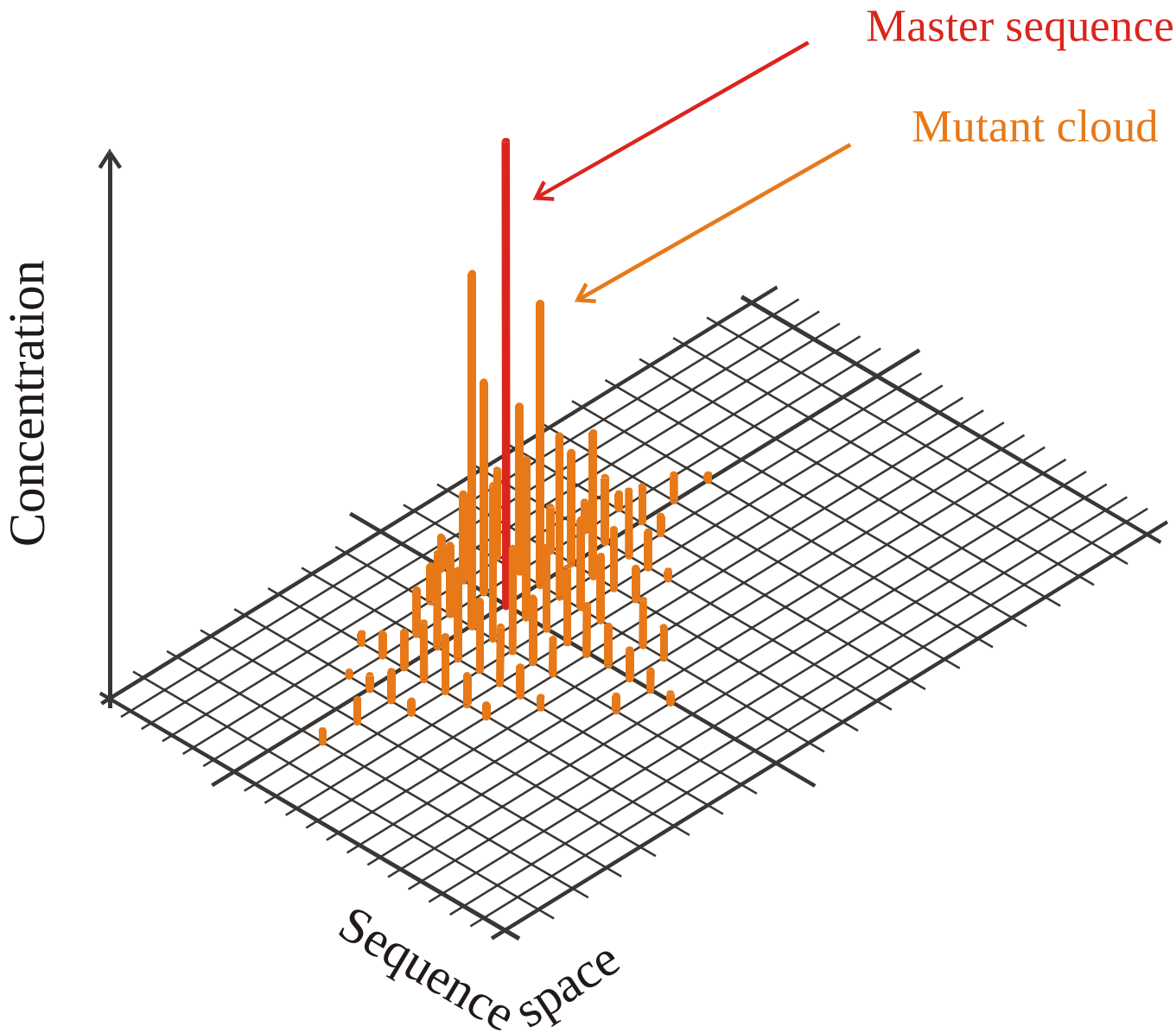
p = 0.001 / site × replication

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

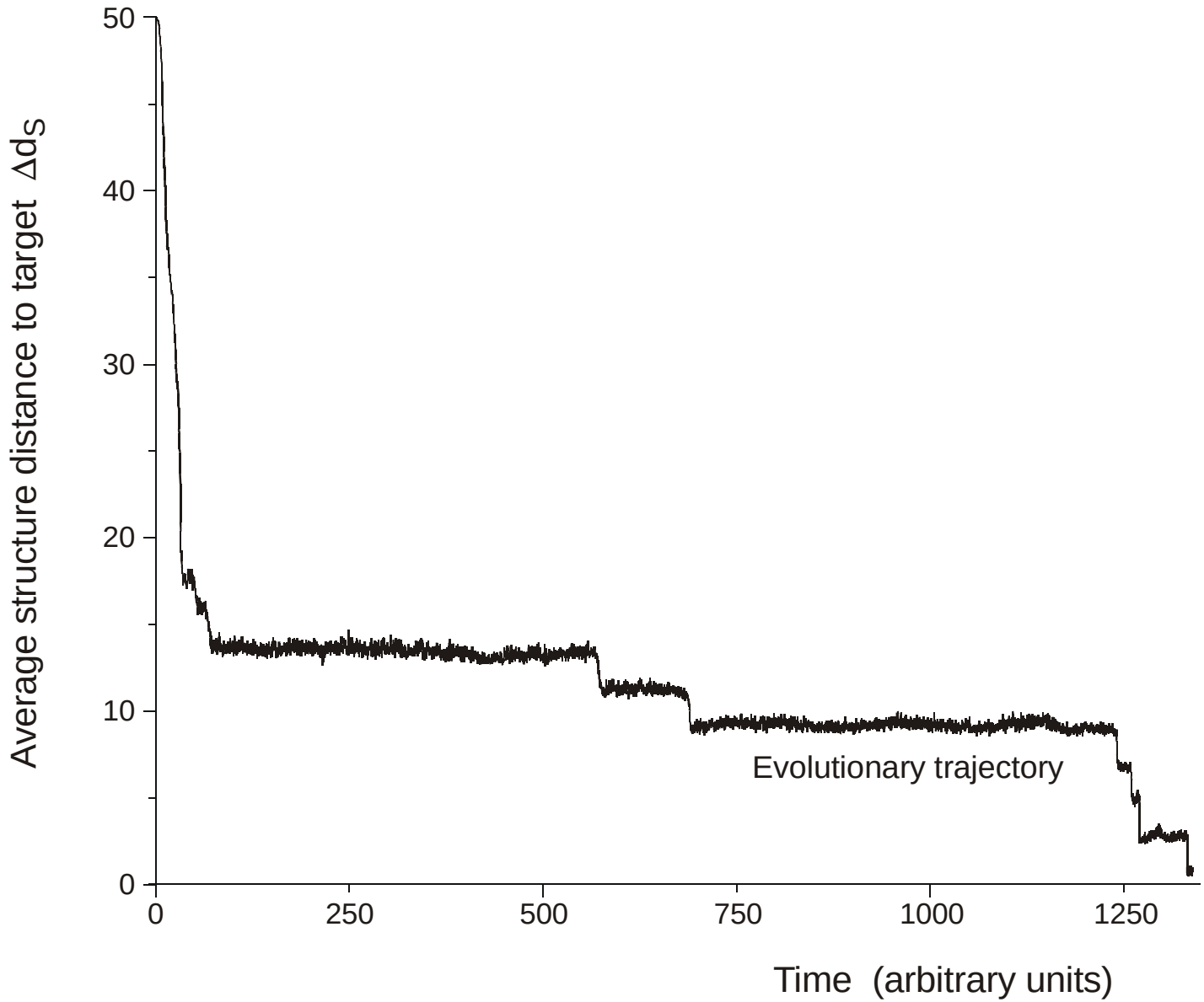


Randomly chosen initial structure

Phenylalanyl-tRNA as target structure

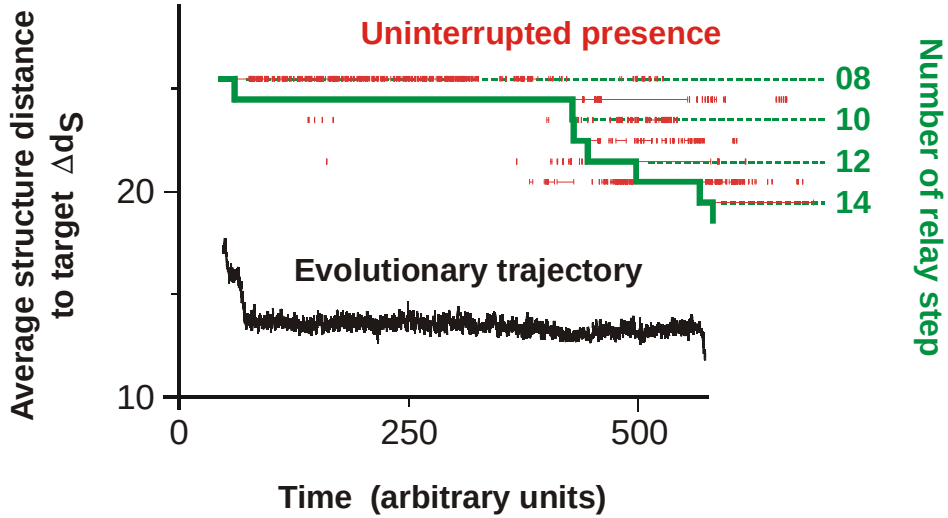


The molecular quasispecies in sequence space



In silico optimization in the flow reactor: Evolutionary trajectory

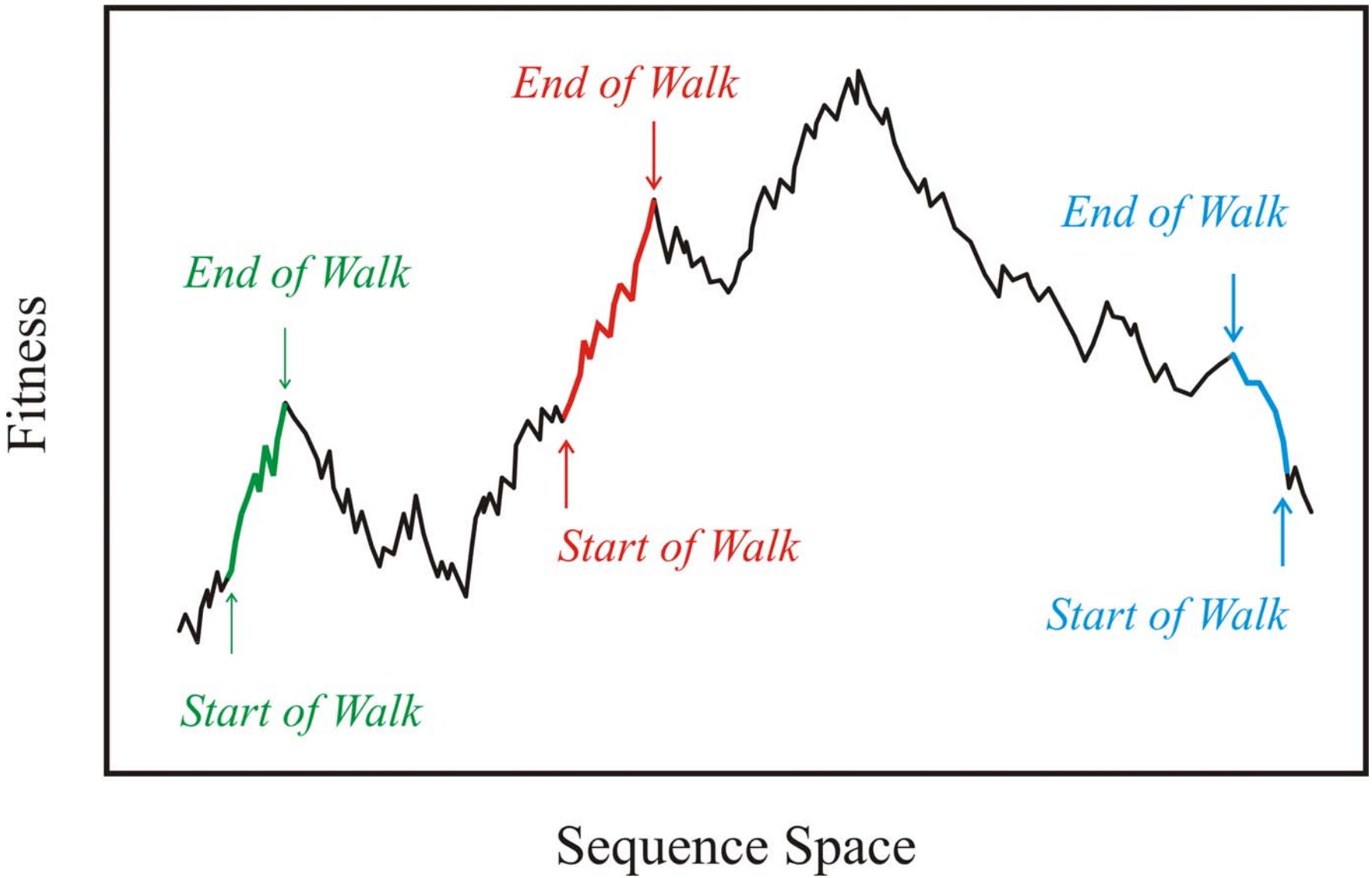
28 neutral point mutations during a long quasi-stationary epoch

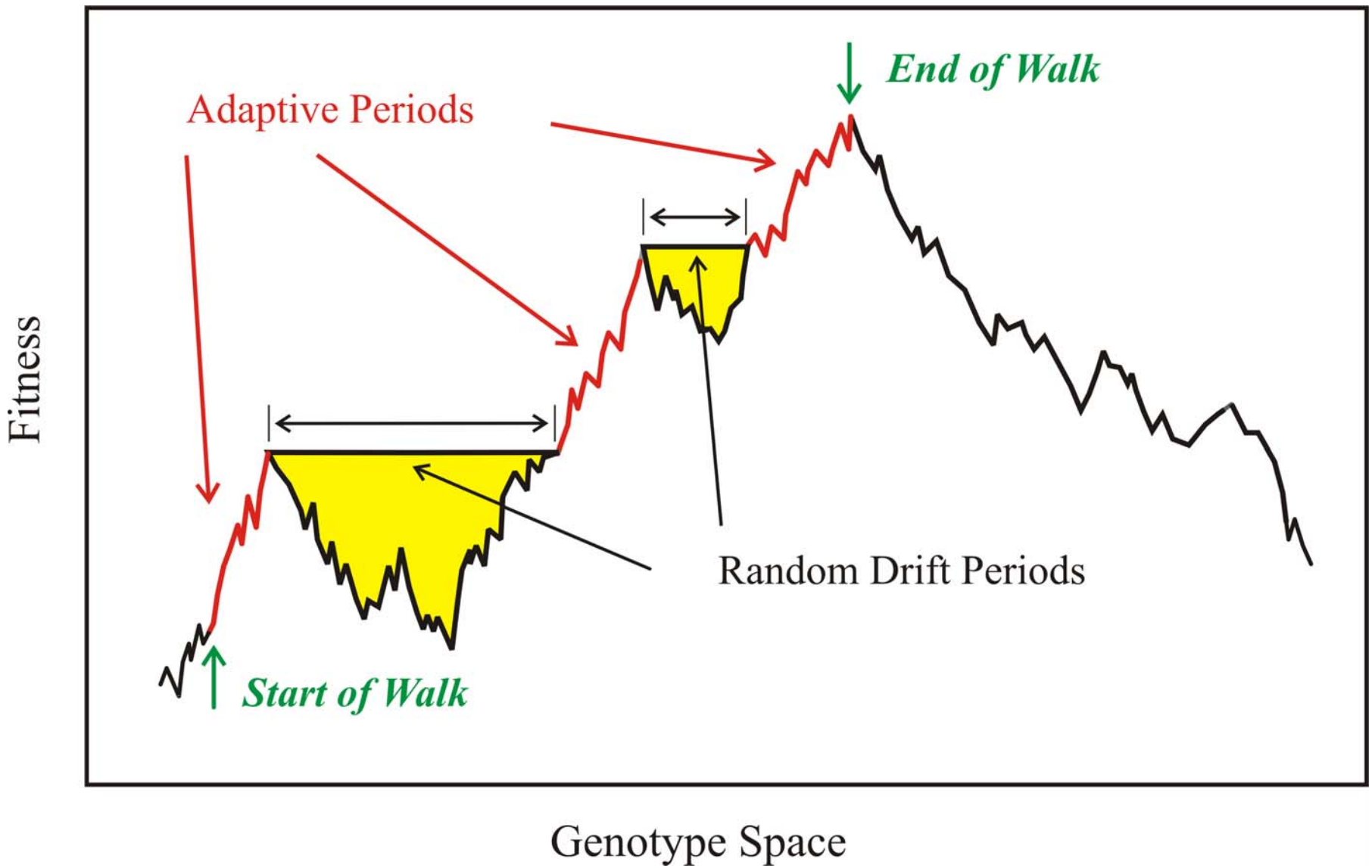



```
entry  GGUAUGGGCGUUGAAUAGUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCGUACAGAA
8      .((((((((((((.........(((....)))........)))))).....((((.........))))))))))....
exit   GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCAUACAGAA
entry  GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUACCAUACAGAA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

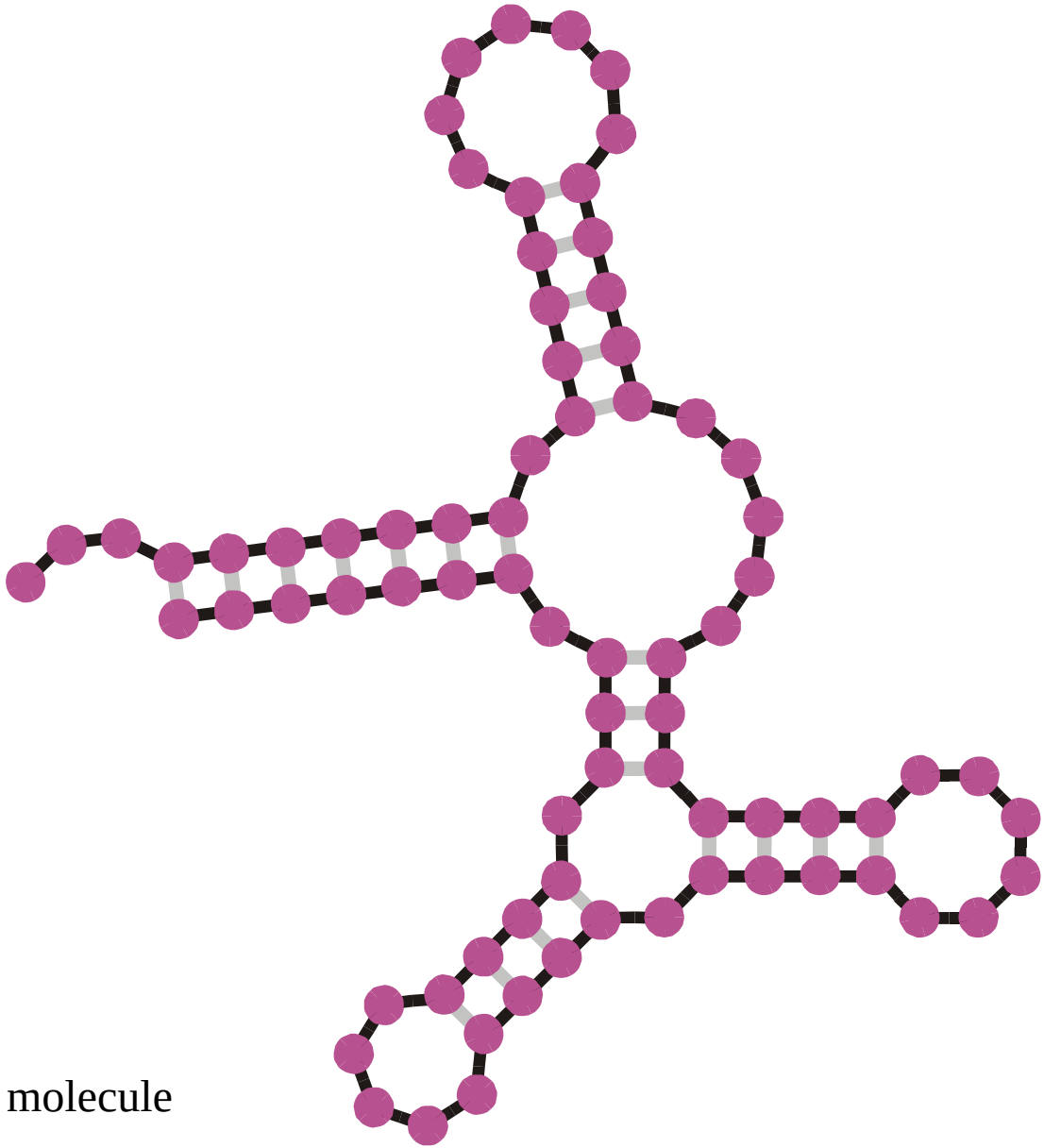
Transition inducing point mutations **Neutral point mutations**

Neutral genotype evolution during phenotypic stasis

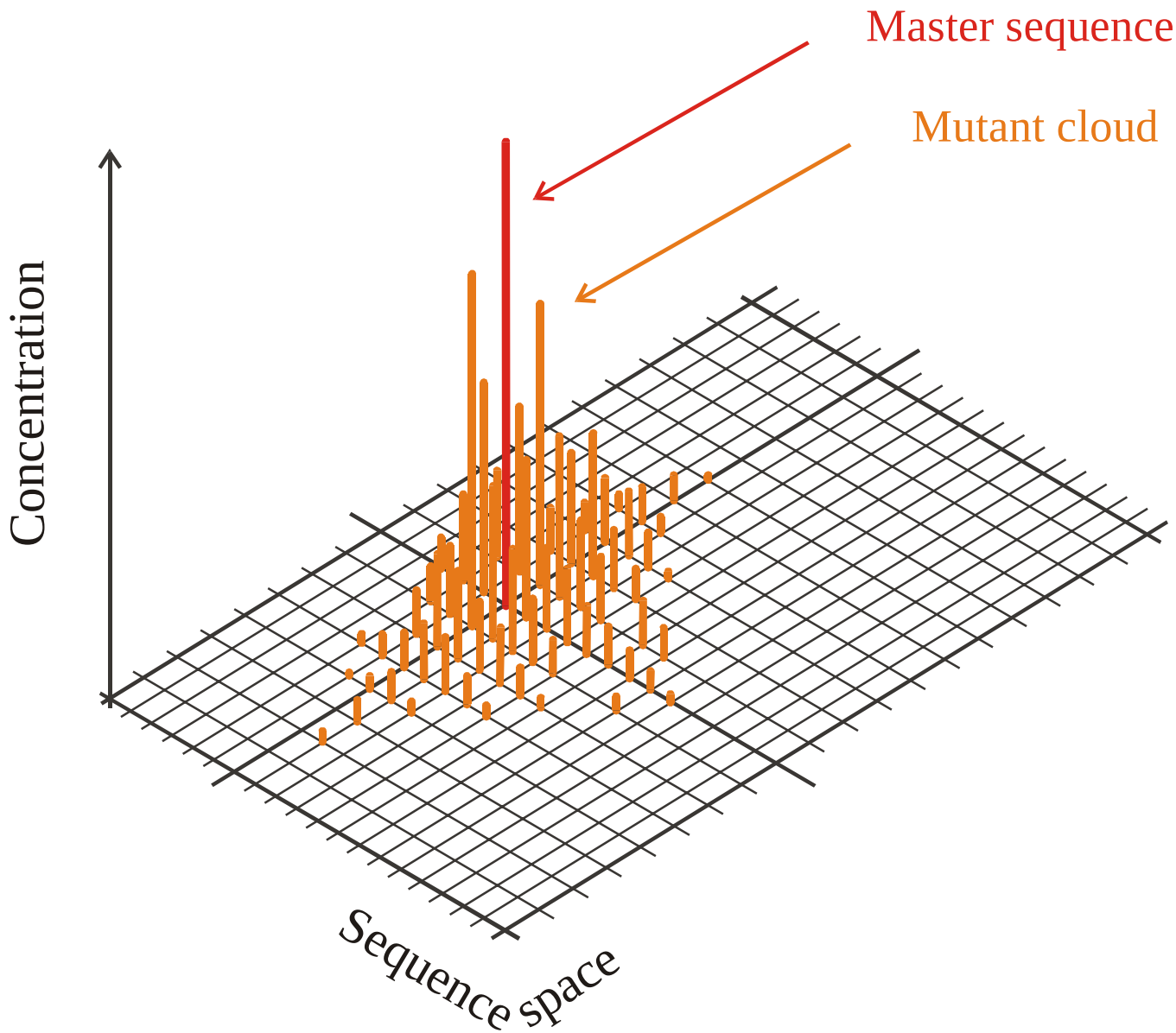

End of Walk
End of Walk
End of Walk
Start of Walk
Start of Walk
Start of Walk
Fitness
Sequence Space


End of Walk
Adaptive Periods
Random Drift Periods
Start of Walk
Fitness
Genotype Space

1. Folding and inverse folding of RNA
2. Neutral networks
3. Darwinian evolution of RNA
4. **Learning by the Darwinian mechanism**
5. Folding kinetics and metastable structures
6. Intersections and conformational switches



Element in example 1: The RNA molecule

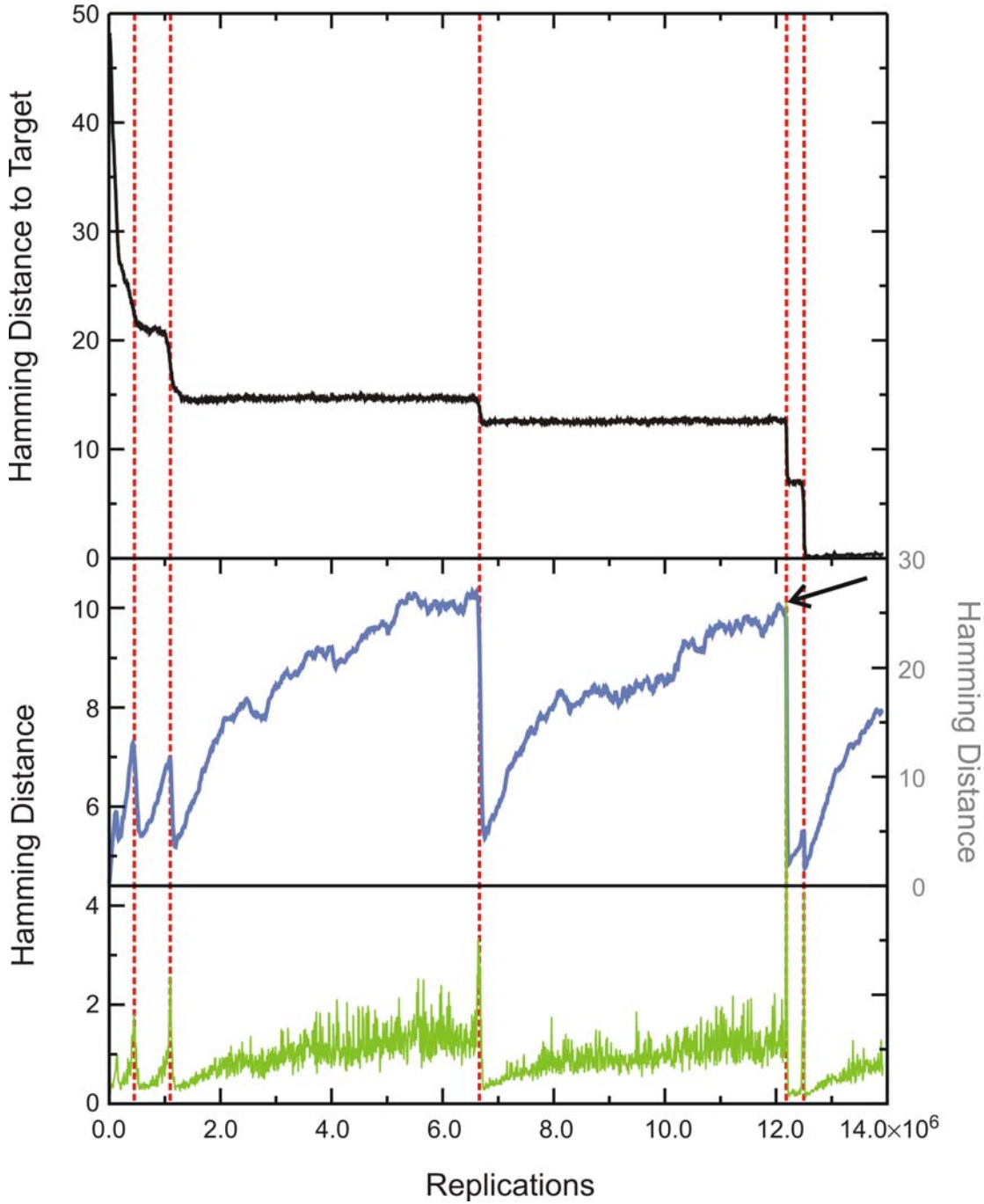


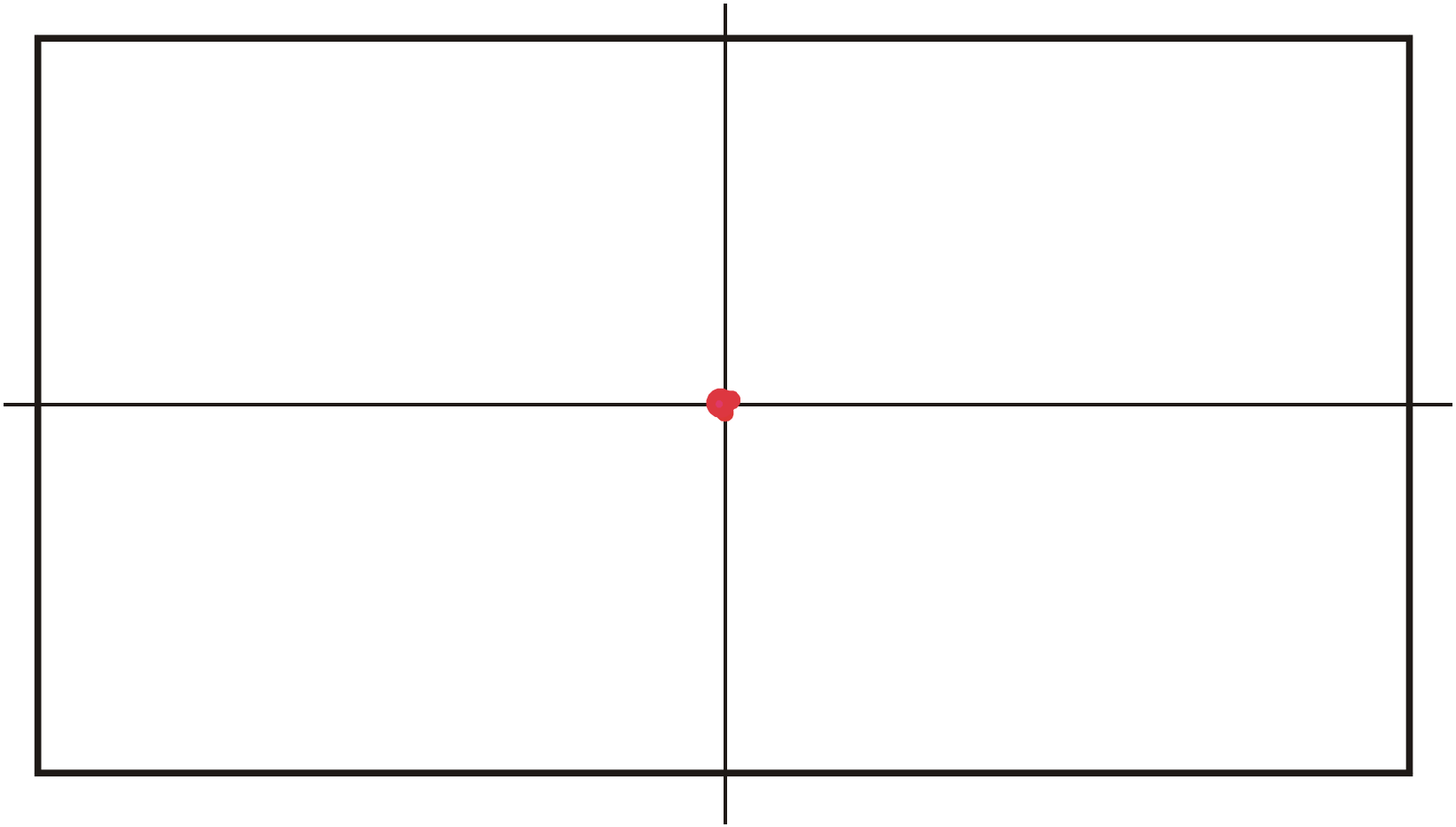
The molecular quasispecies in sequence space

Evolutionary trajectory

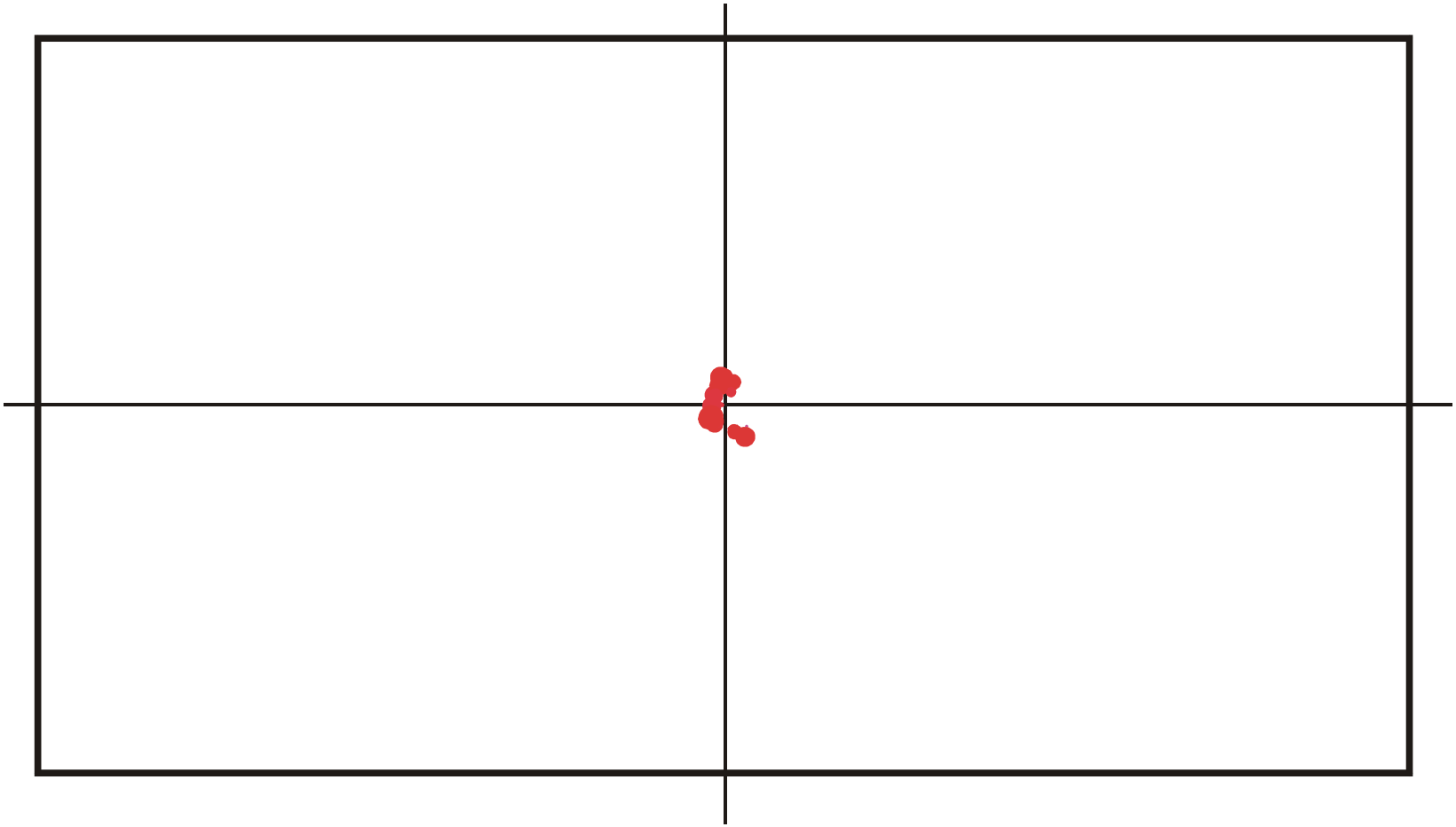
Spreading of the population through diffusion on a neutral network

Drift of the population center in sequence space

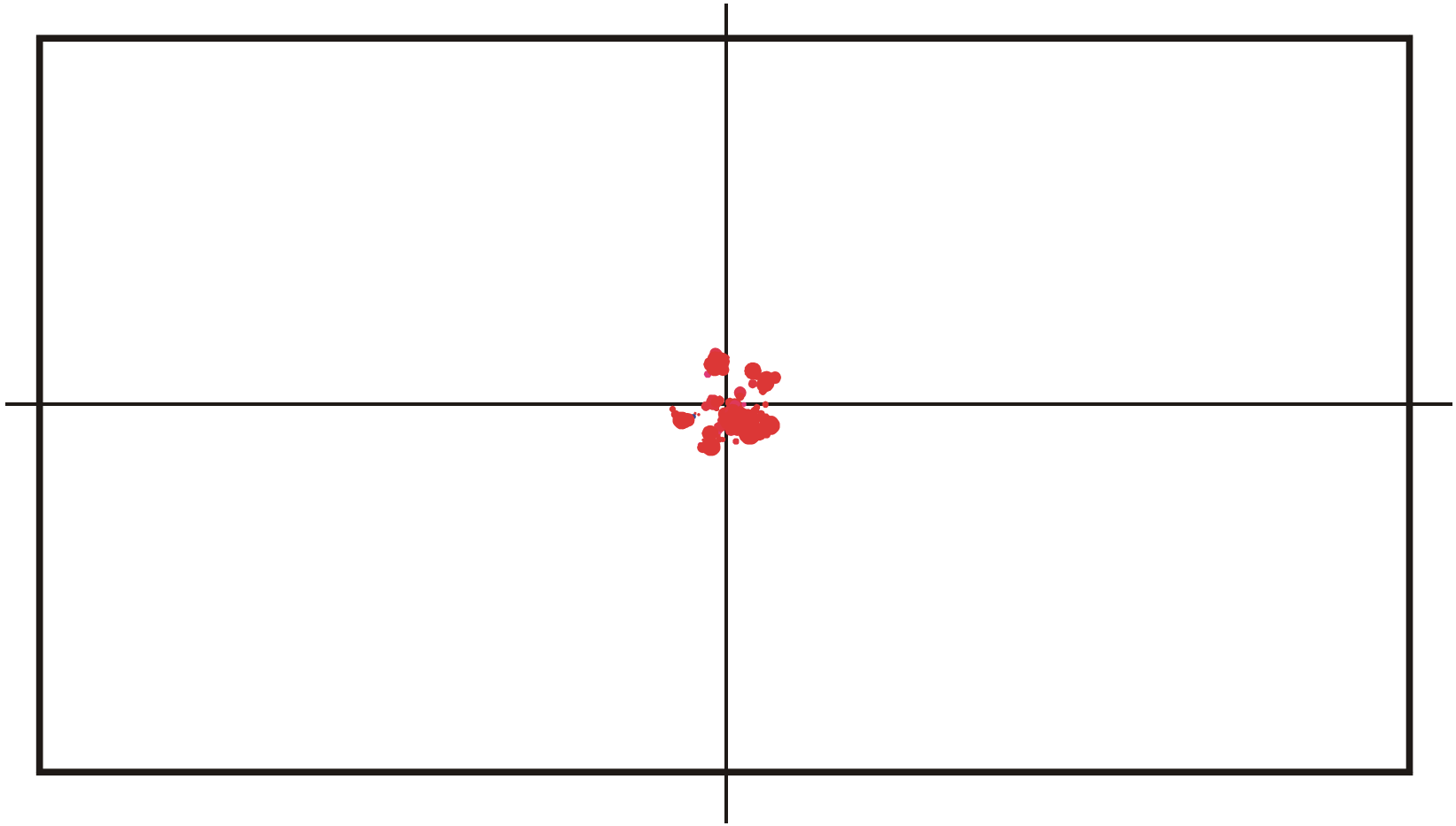




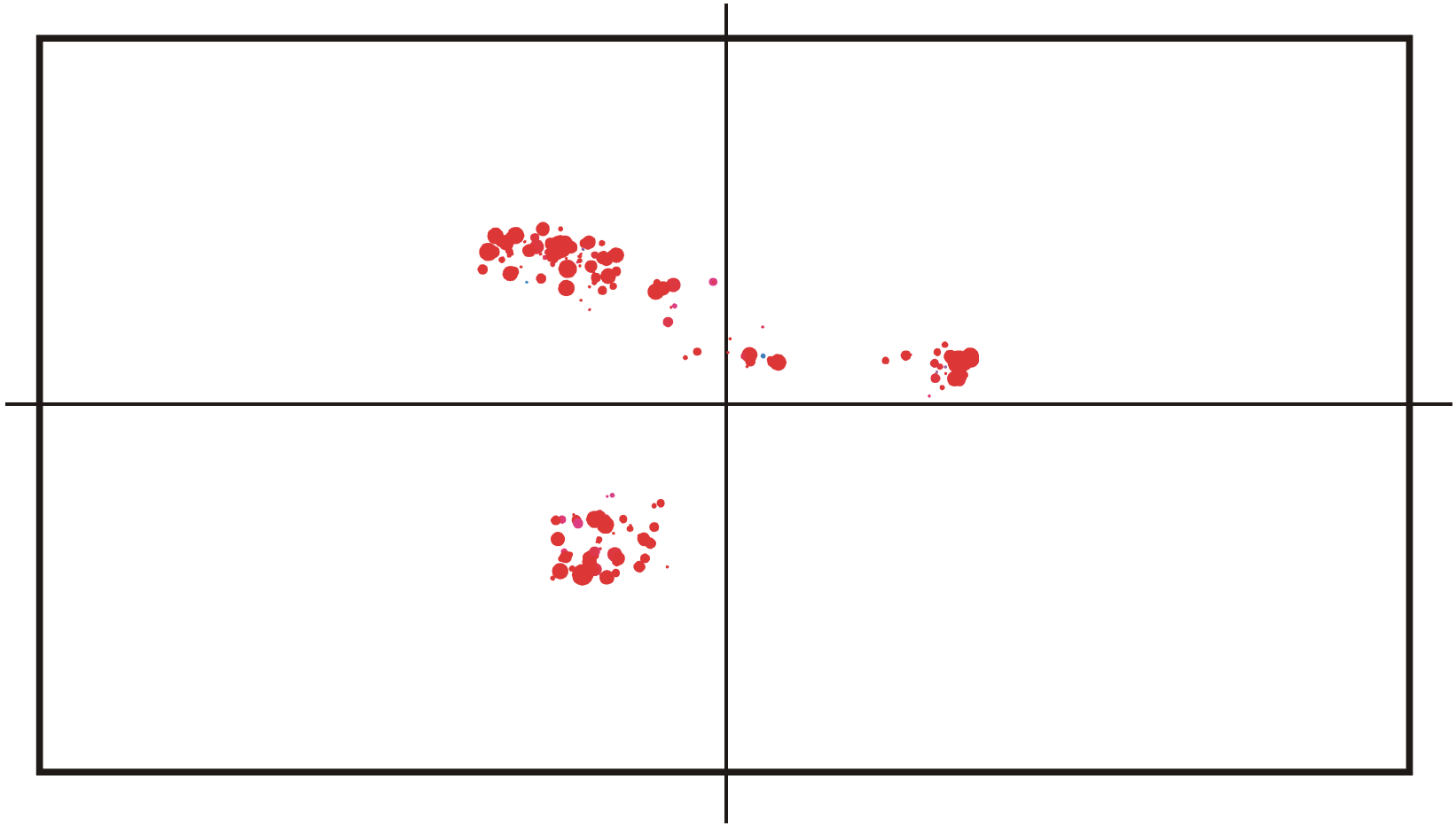
Spread of population in sequence space during a quasistationary epoch: t = 150



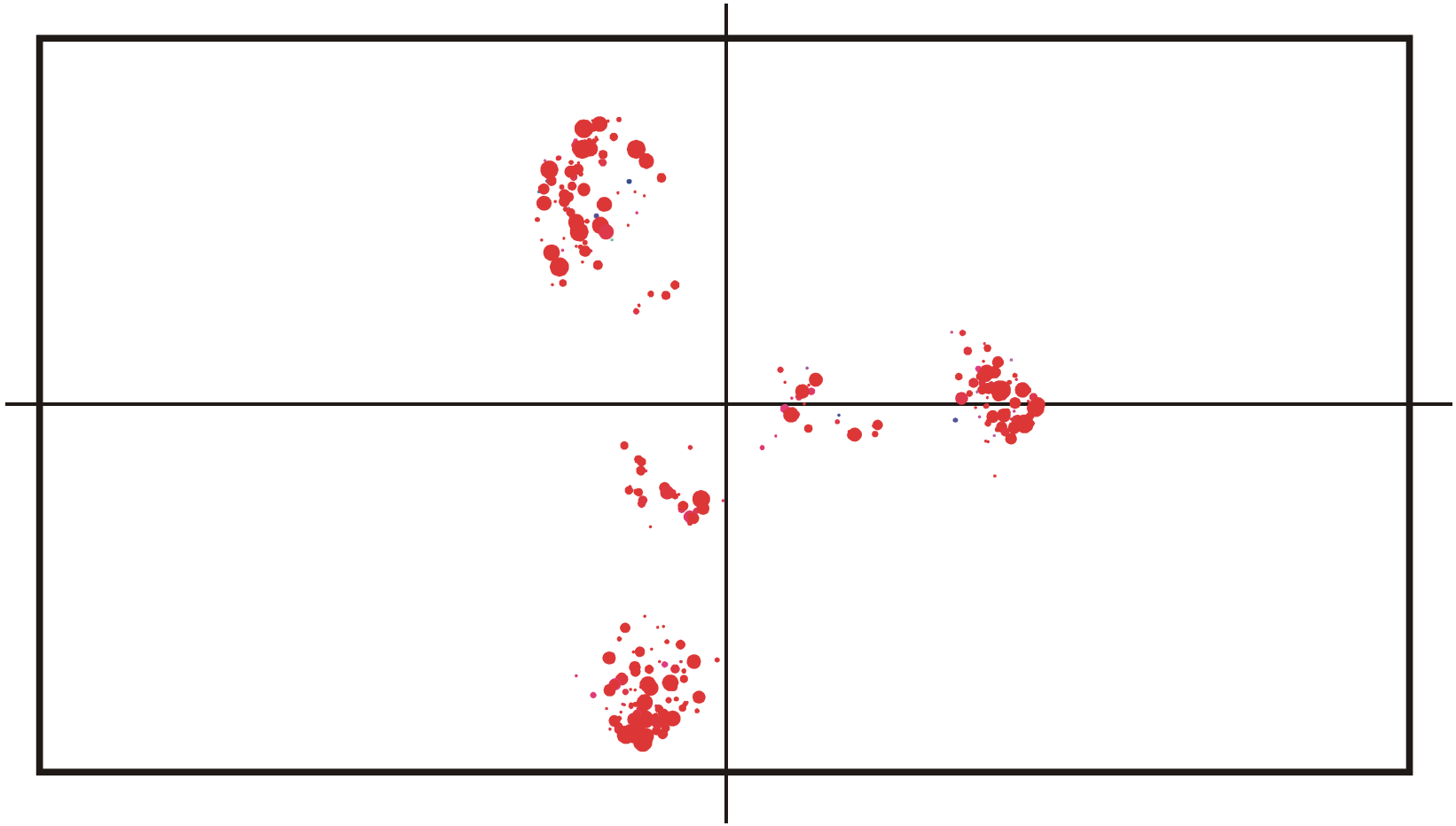
Spread of population in sequence space during a quasistationary epoch: t = 170



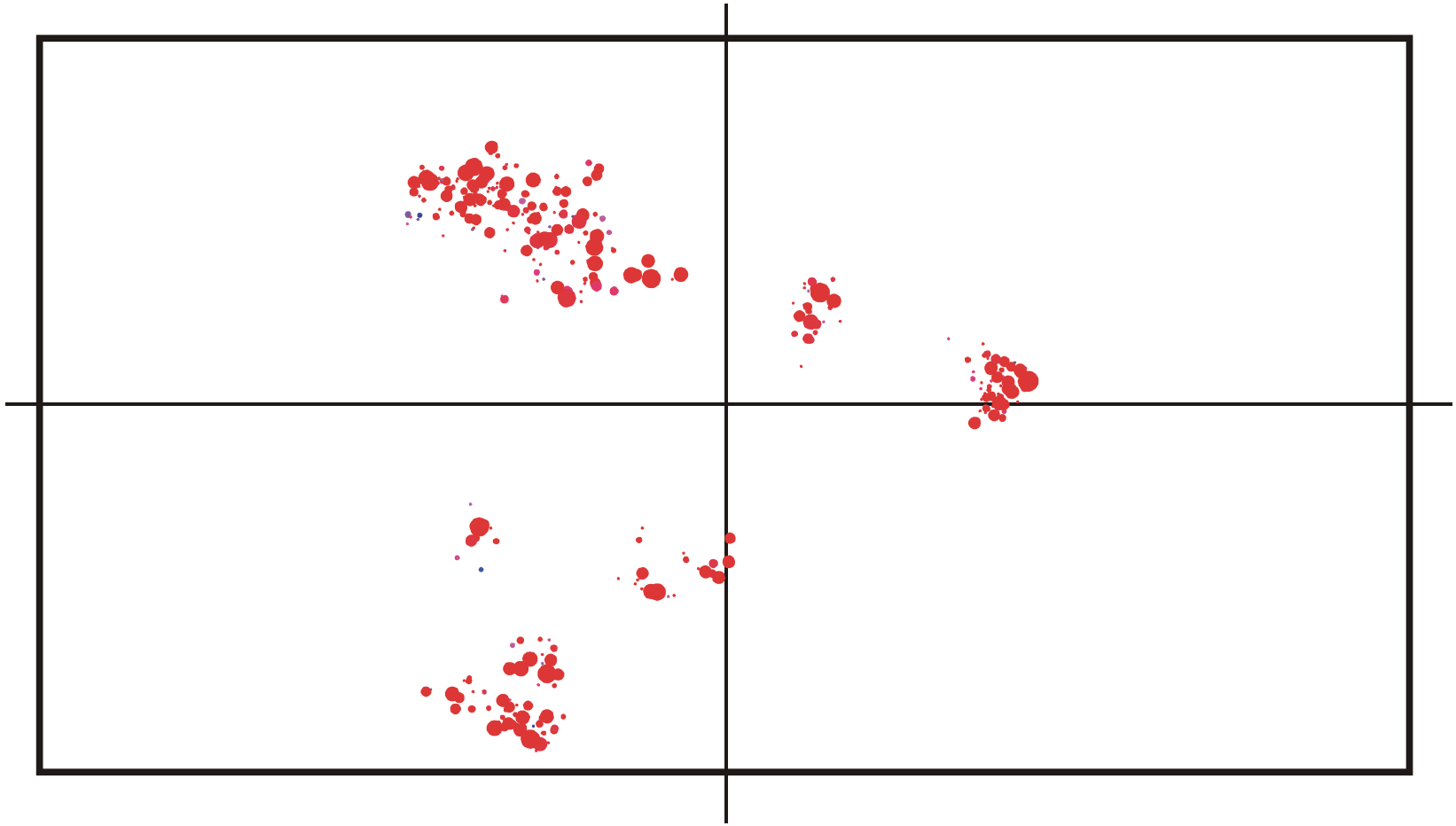
Spread of population in sequence space during a quasistationary epoch: t = 200



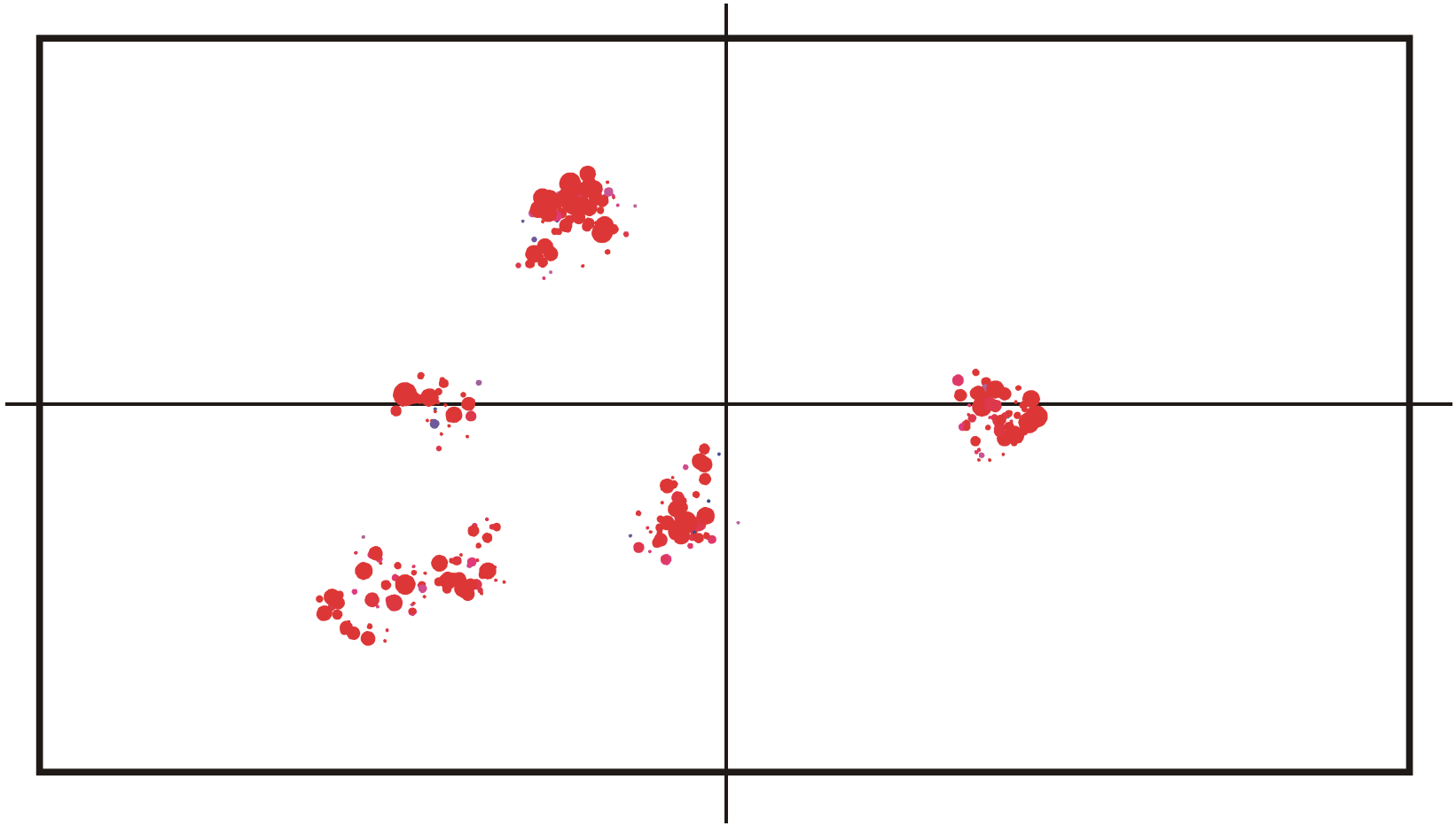
Spread of population in sequence space during a quasistationary epoch: t = 350



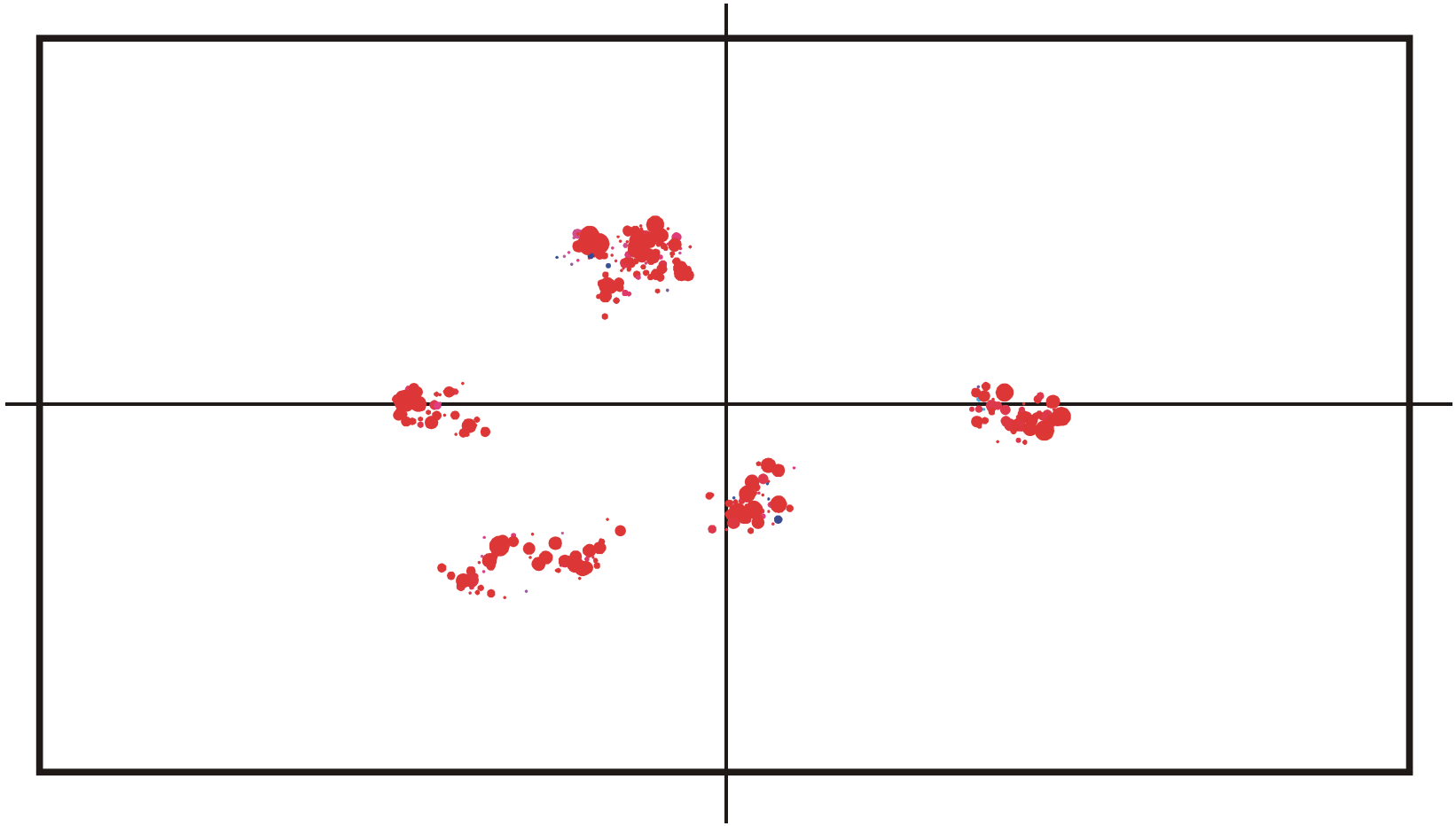
Spread of population in sequence space during a quasistationary epoch: t = 500



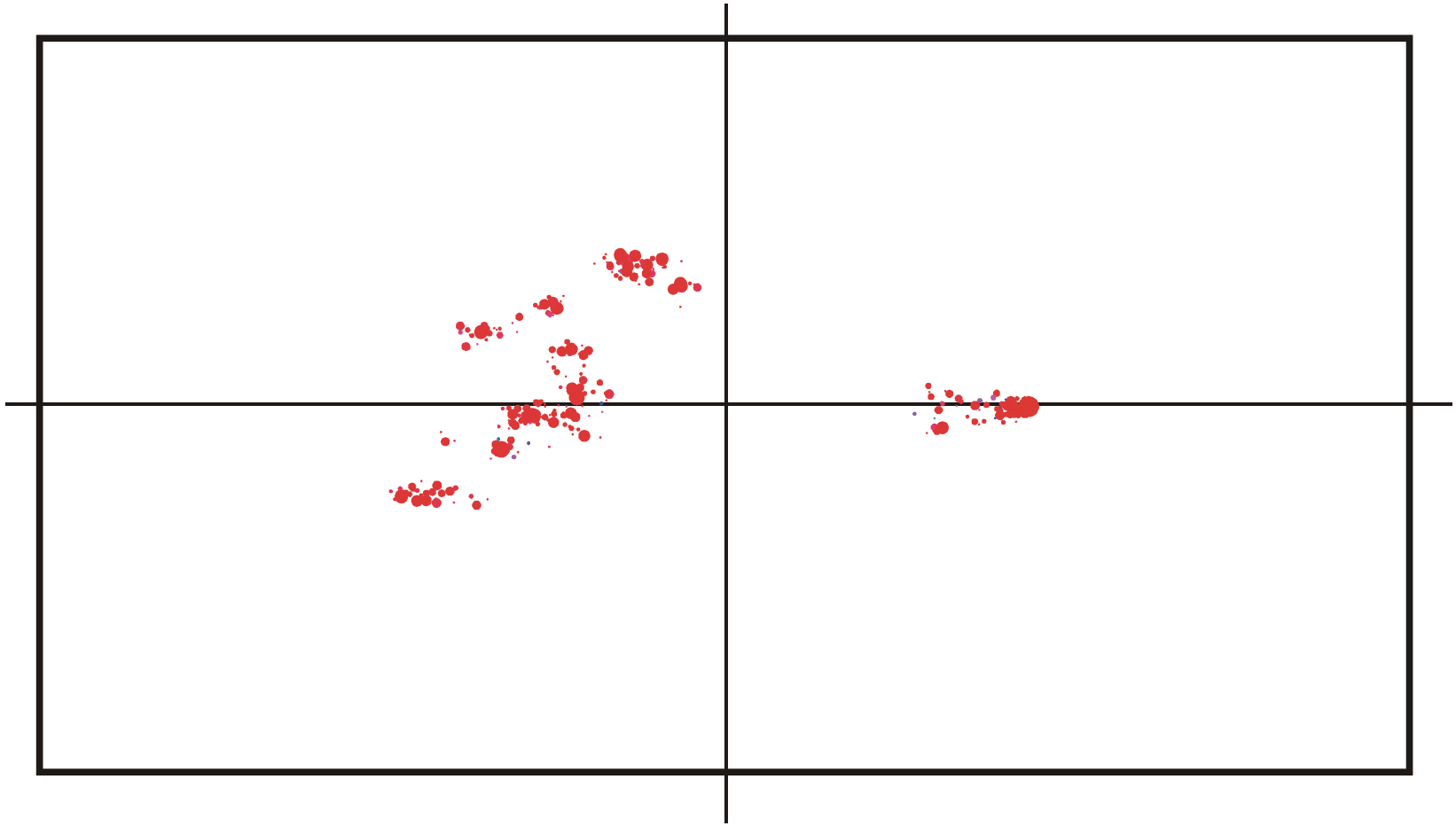
Spread of population in sequence space during a quasistationary epoch: t = 650



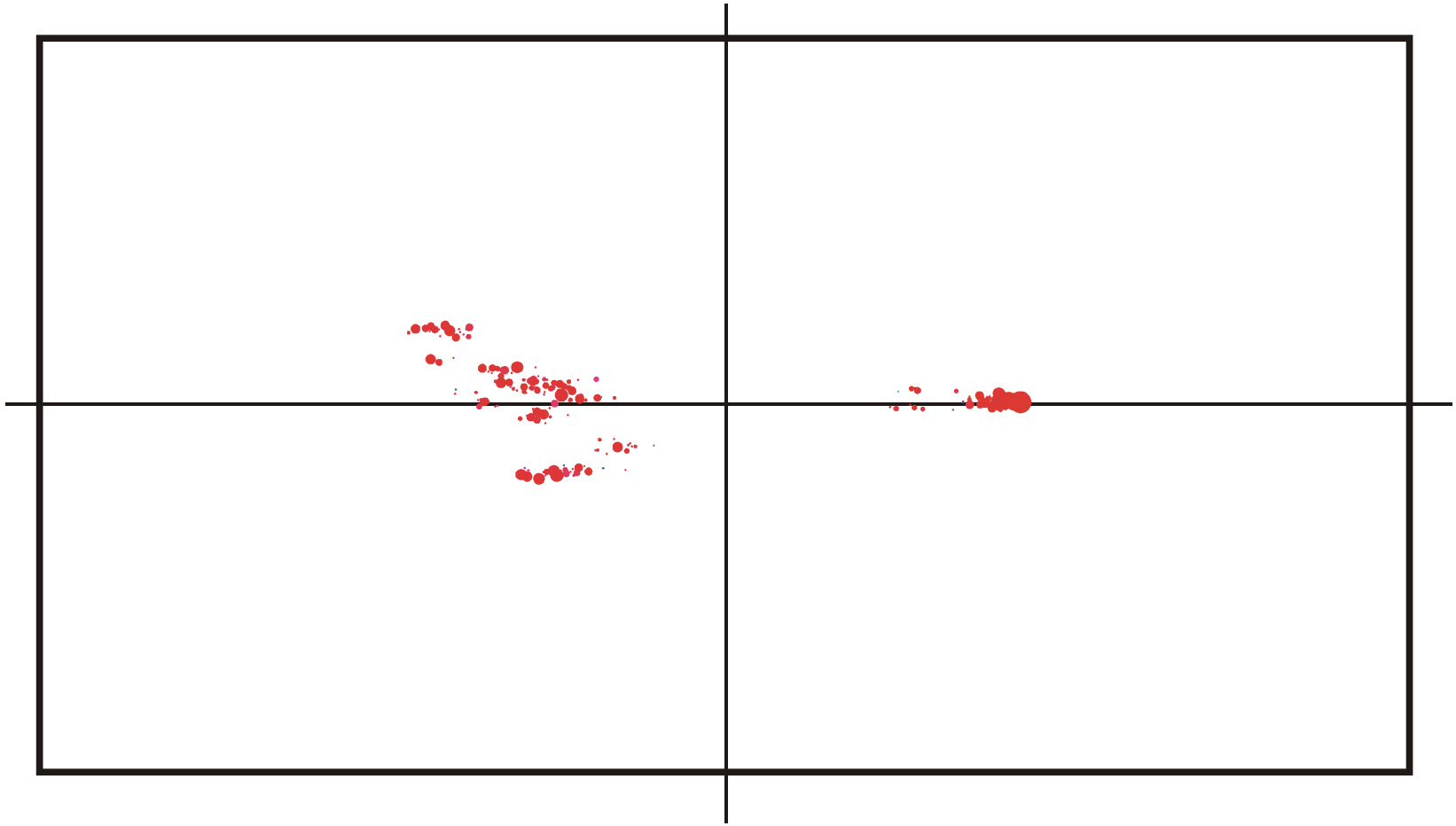
Spread of population in sequence space during a quasistationary epoch: t = 820



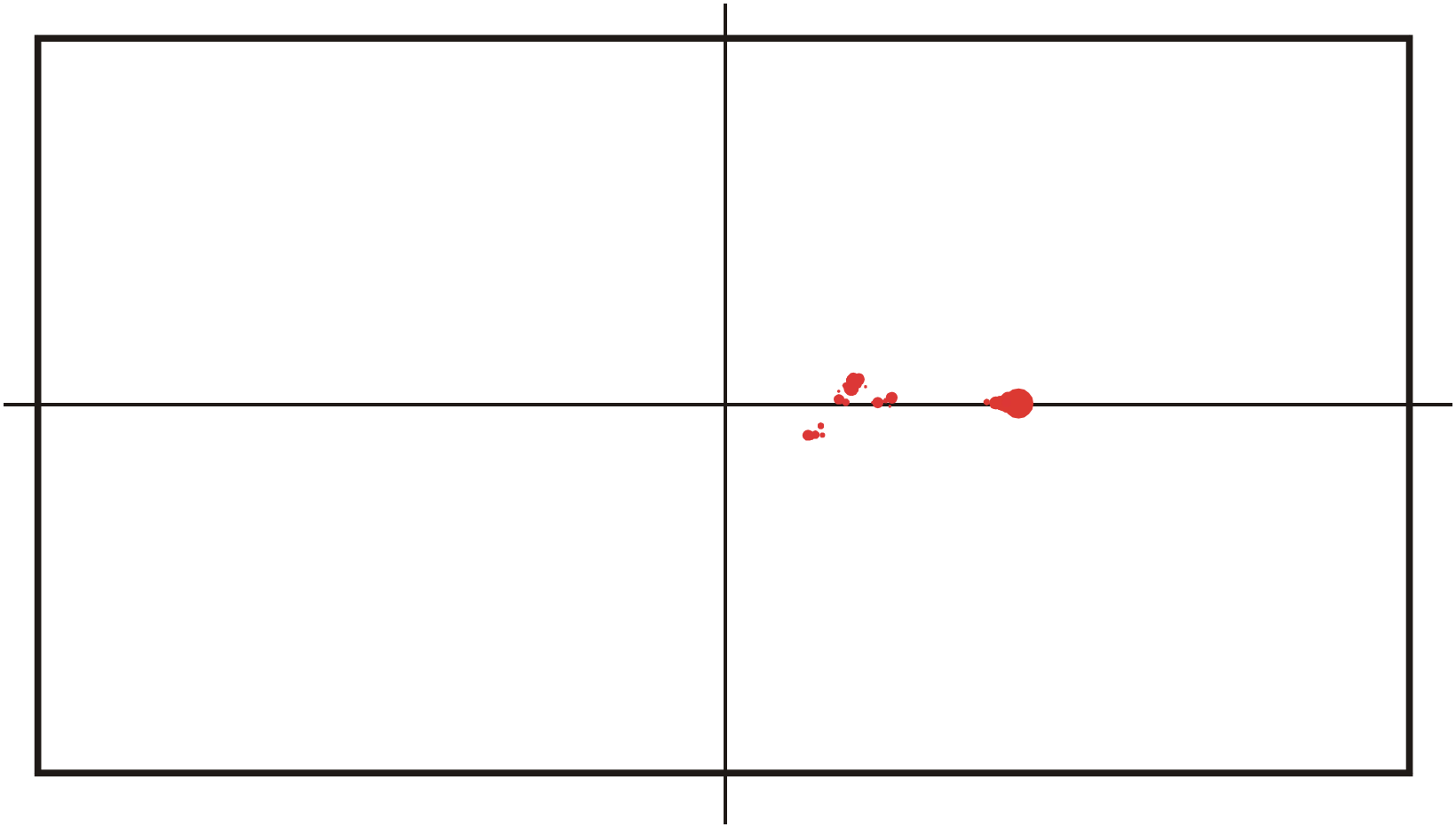
Spread of population in sequence space during a quasistationary epoch: t = 825



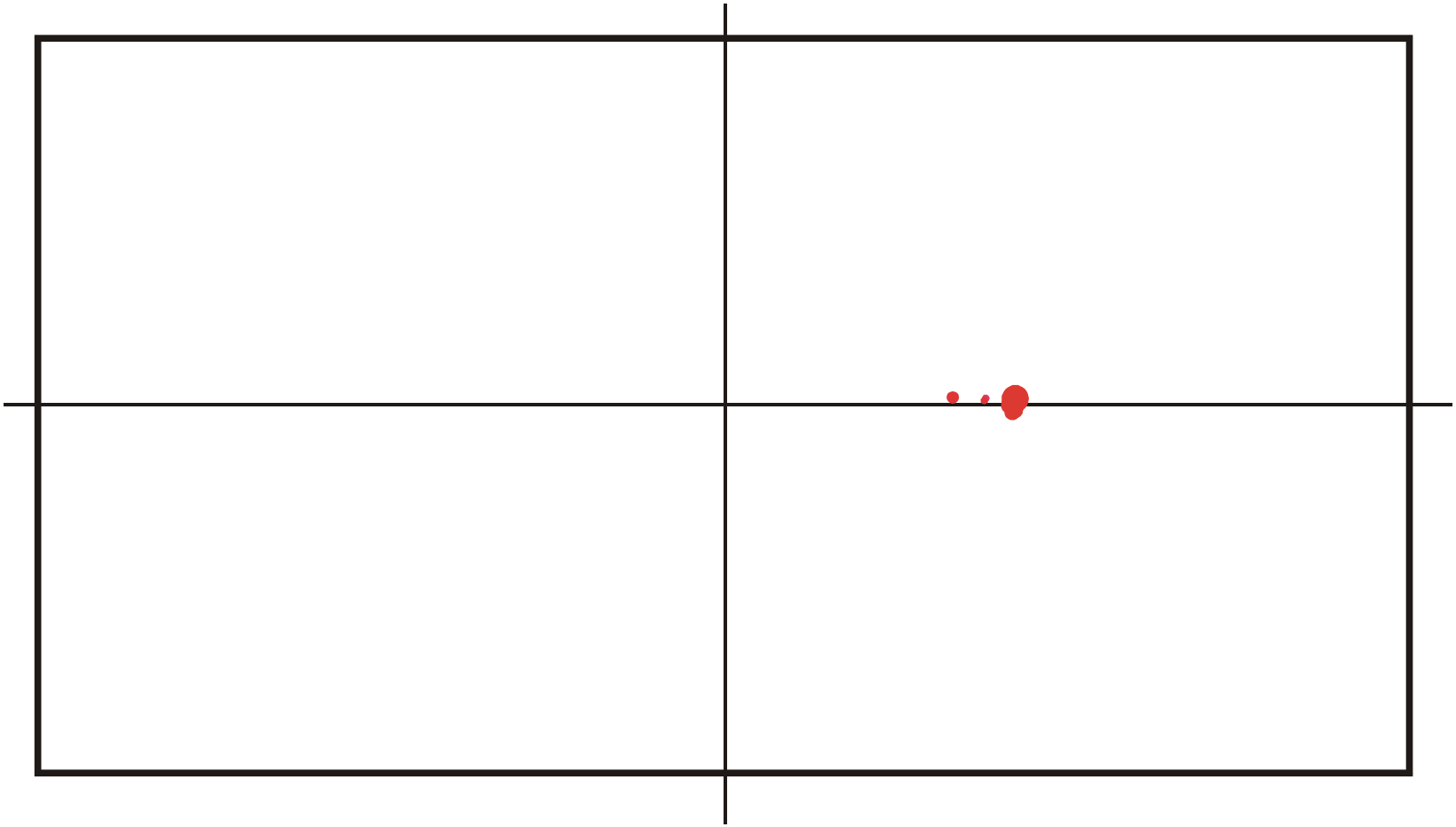
Spread of population in sequence space during a quasistationary epoch: t = 830



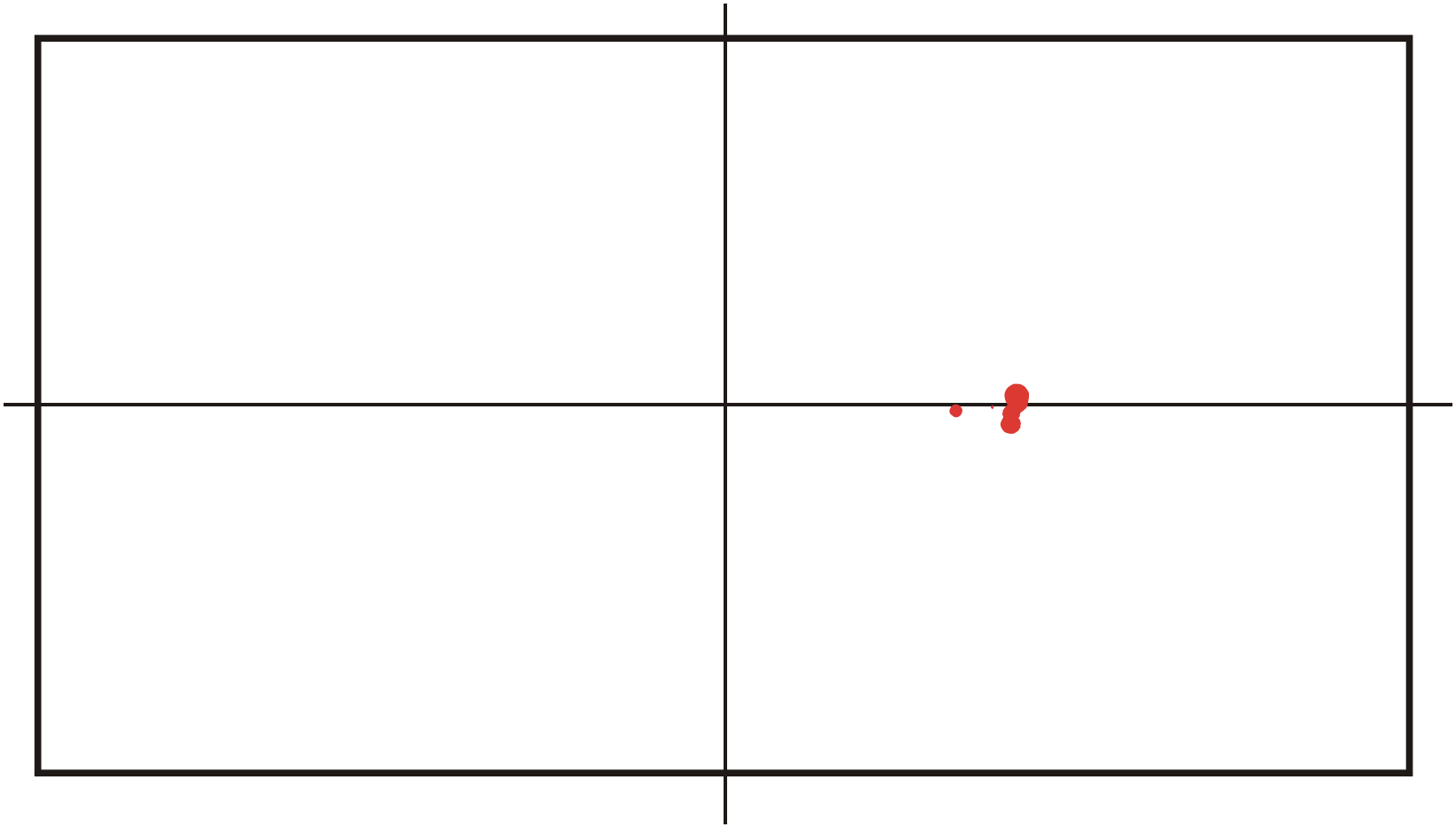
Spread of population in sequence space during a quasistationary epoch: t = 835



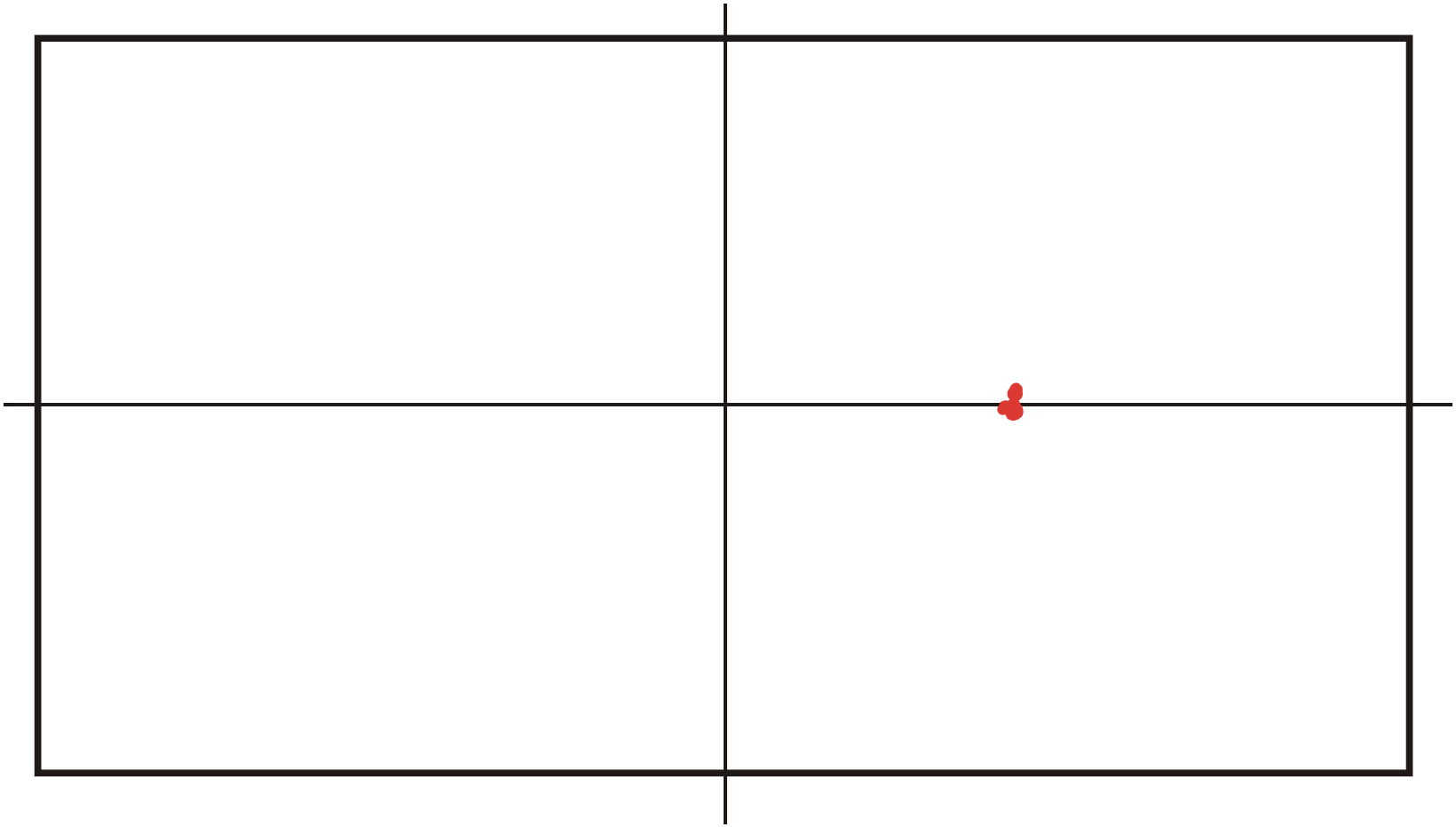
Spread of population in sequence space during a quasistationary epoch: t = 840



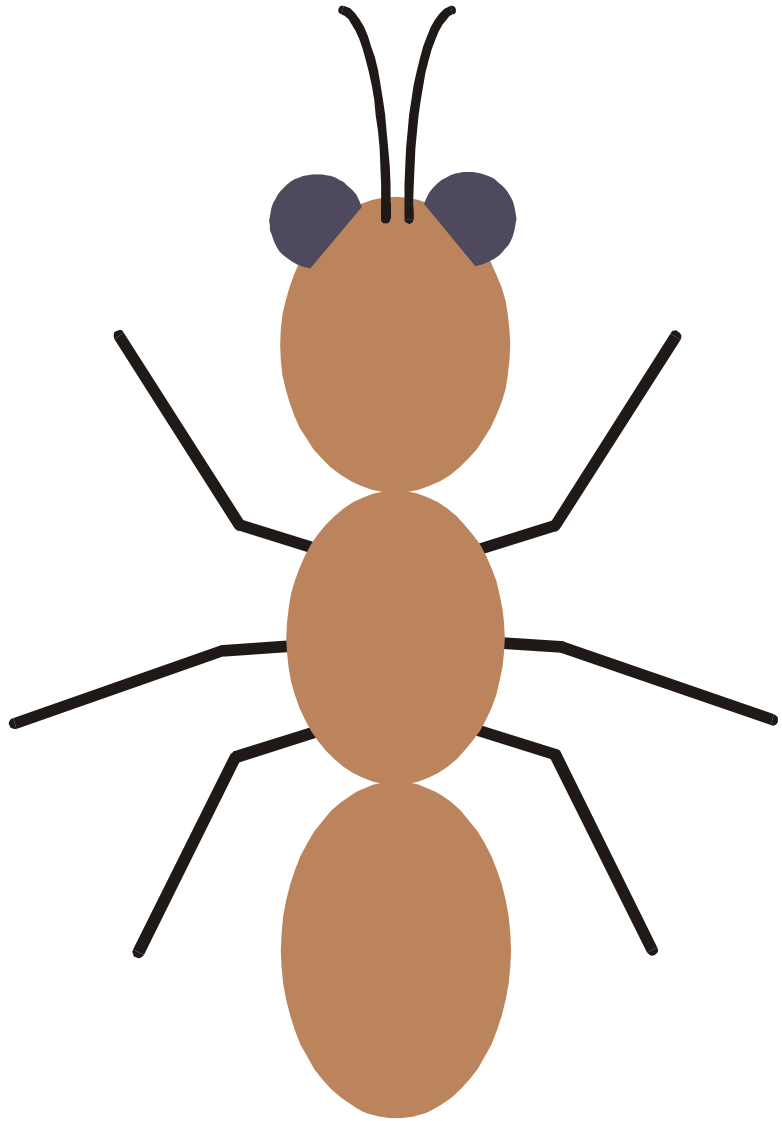
Spread of population in sequence space during a quasistationary epoch: t = 845



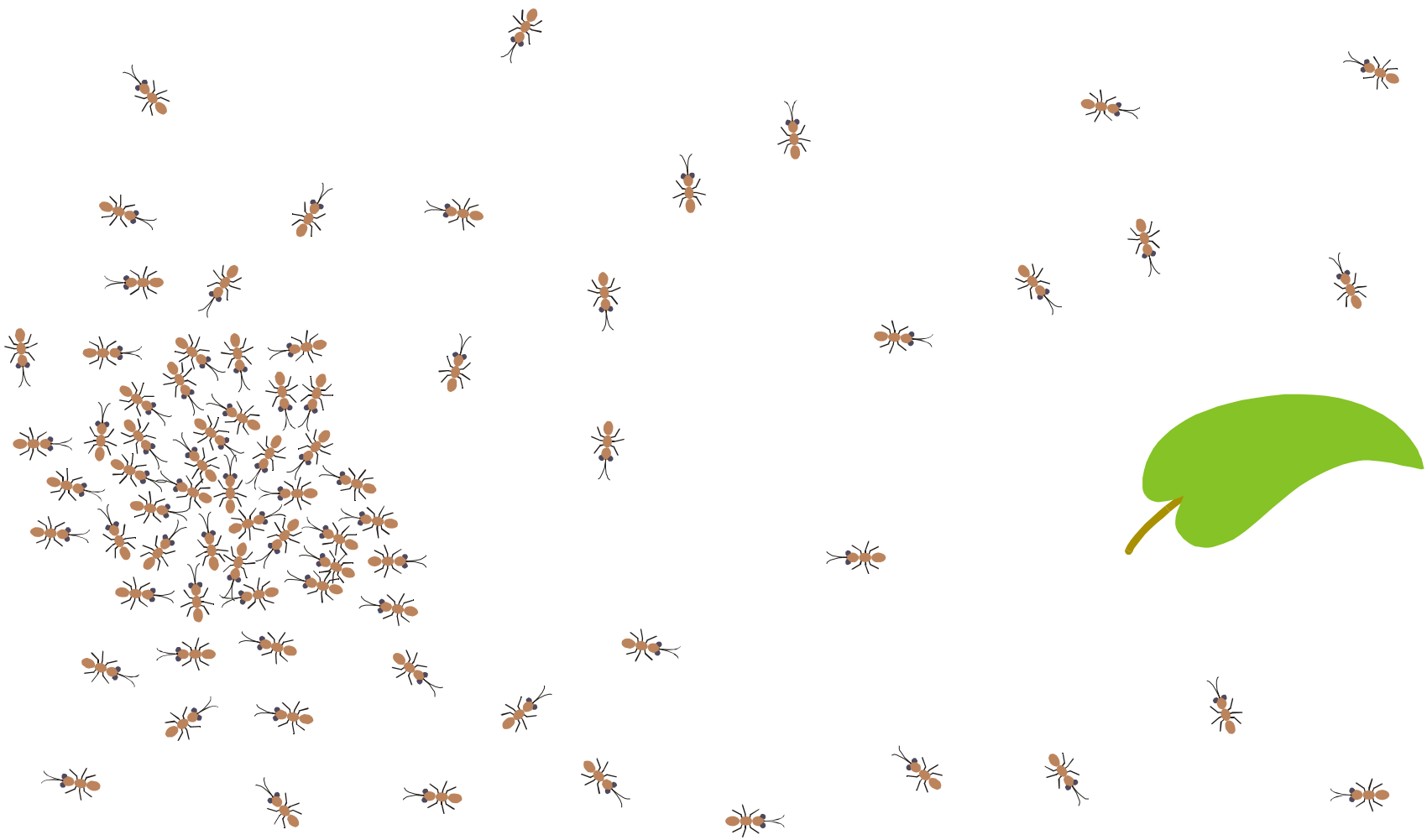
Spread of population in sequence space during a quasistationary epoch: t = 850



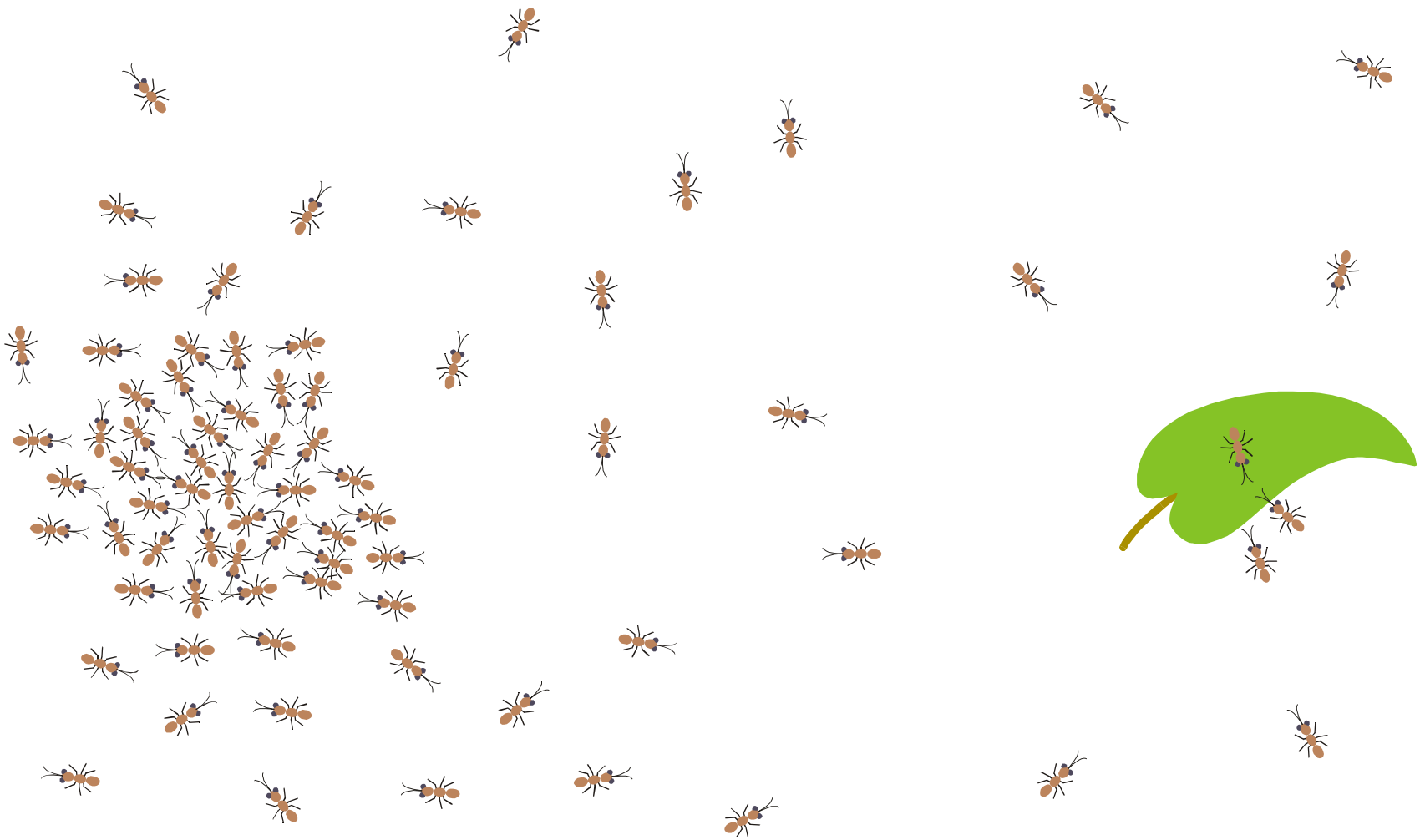
Spread of population in sequence space during a quasistationary epoch: t = 855



Element in example 2: The ant worker



Foraging behavior of ant colonies

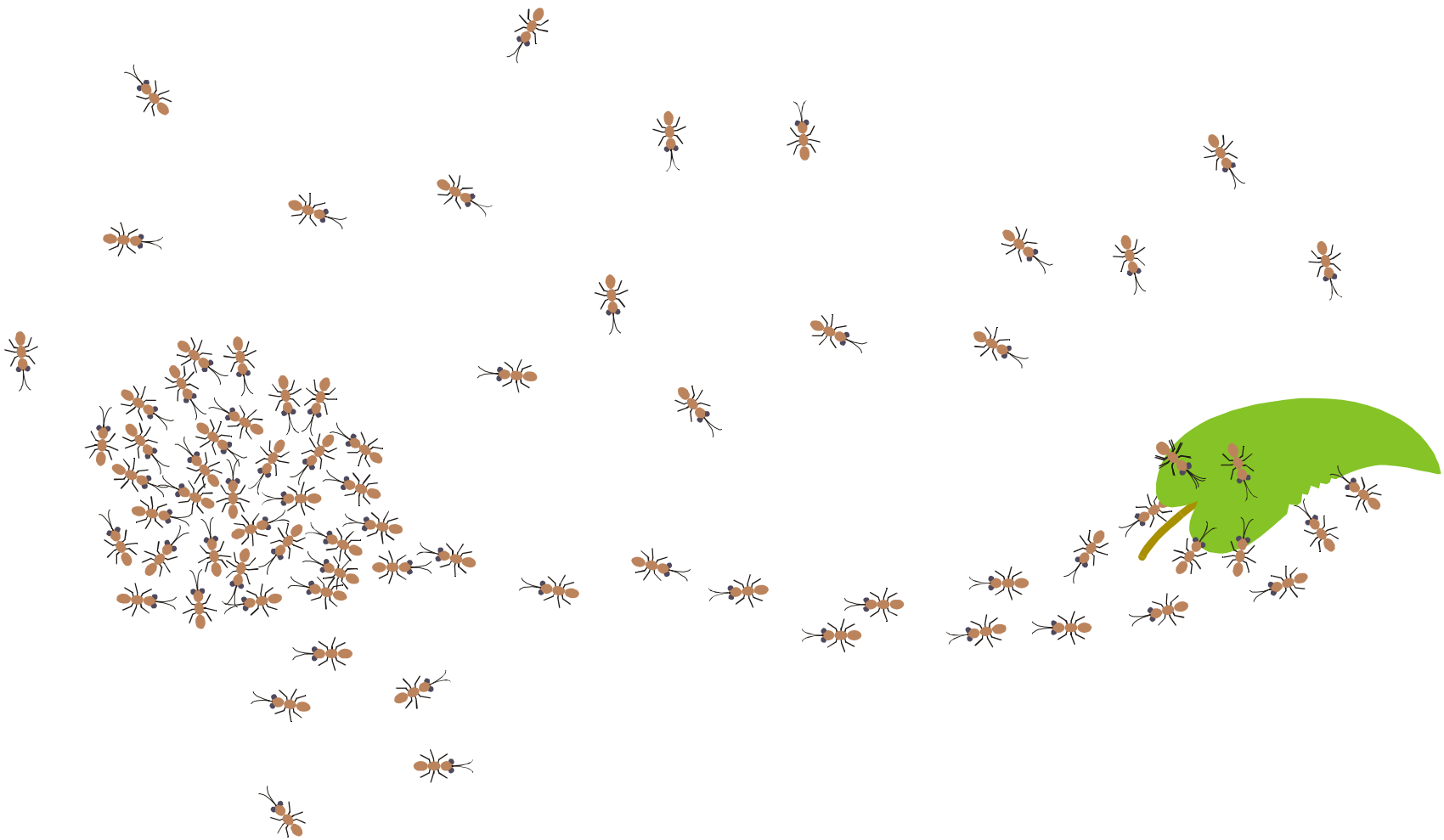


Ant colony

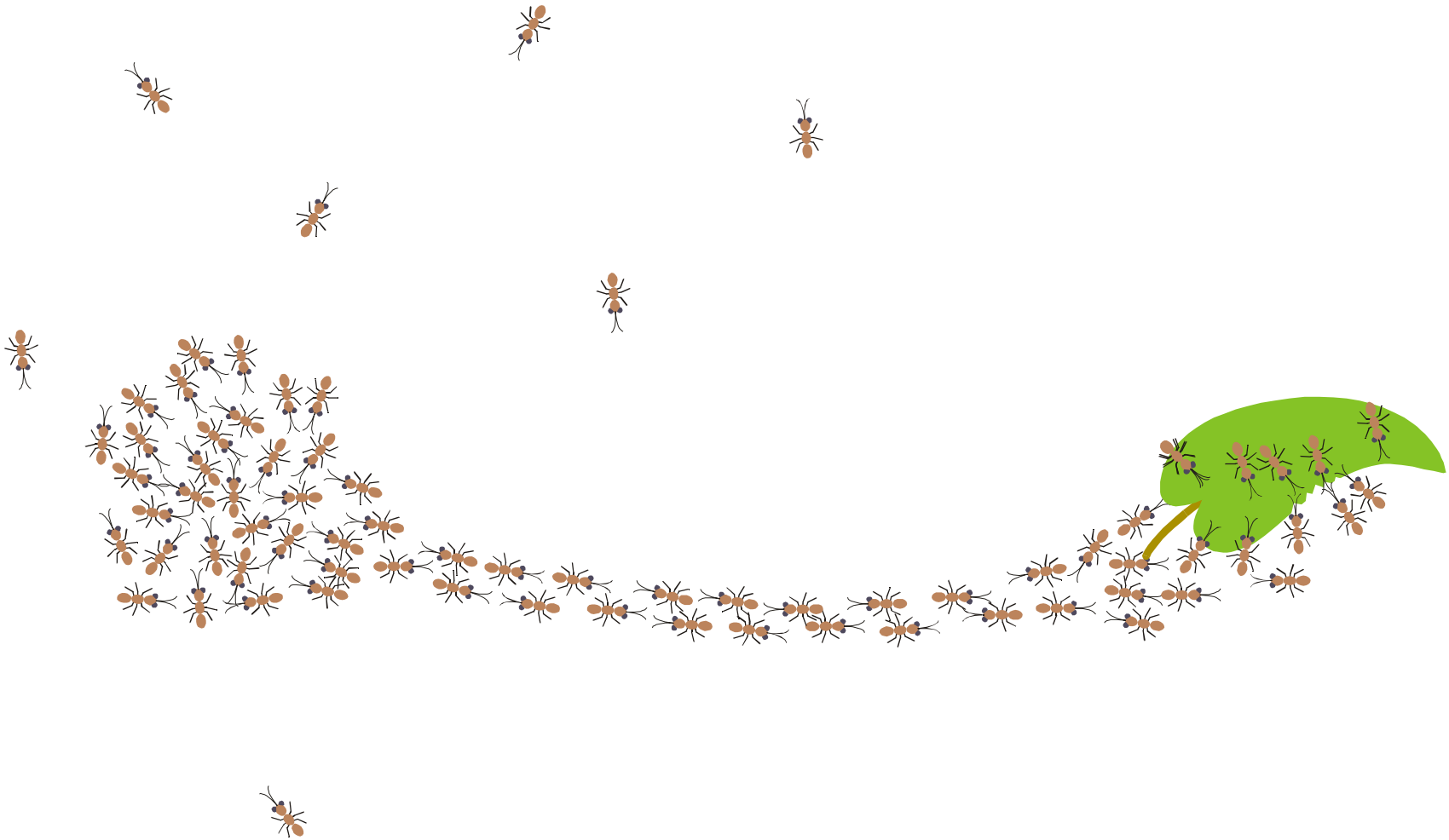
Food source detected

Food source

Foraging behavior of ant colonies



Foraging behavior of ant colonies



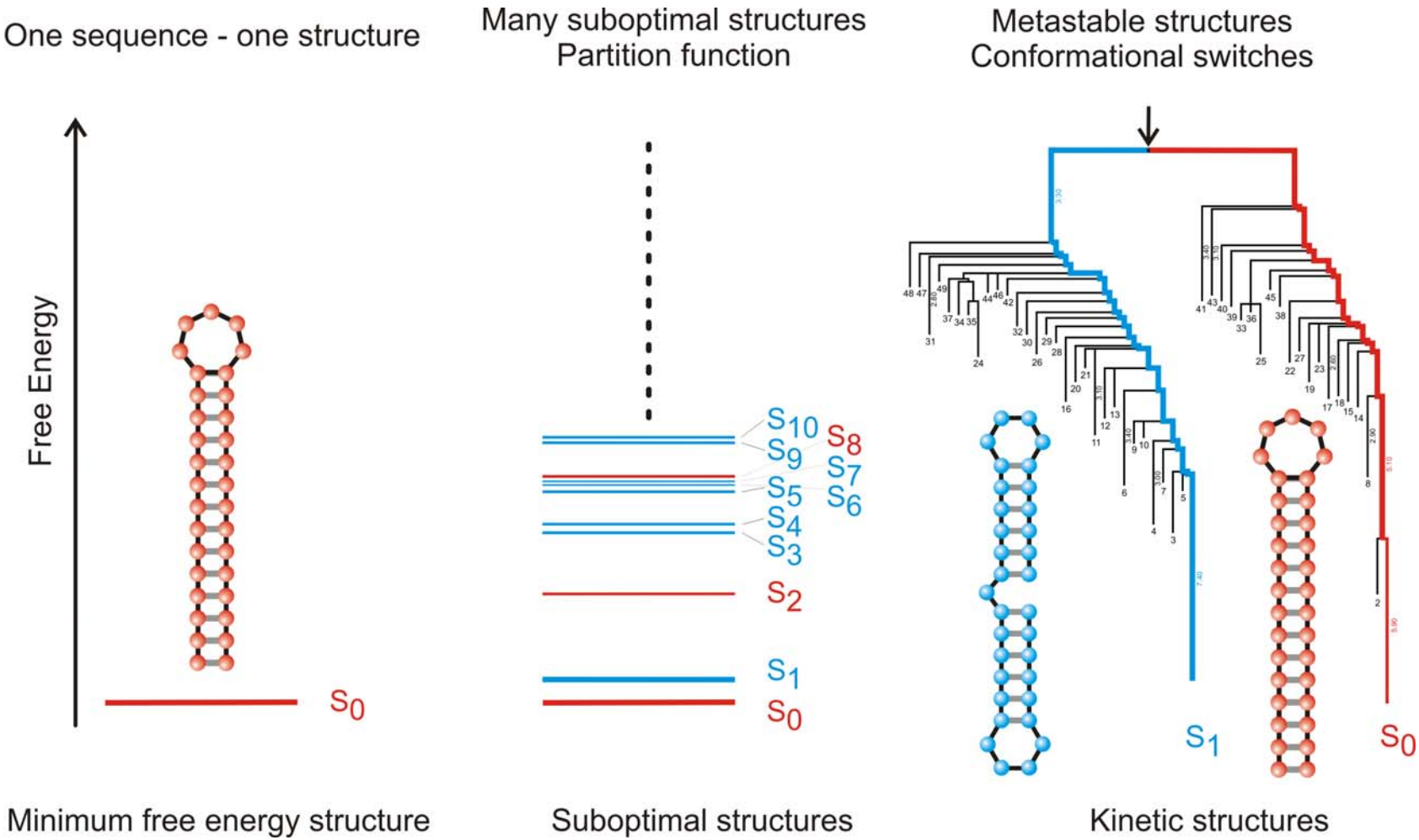
Foraging behavior of ant colonies

	Evolution of RNA	Foraging ants
Element	RNA nucleotide	Individual worker ant
Genotype	RNA sequence	Worker ant collective
Phenotype	RNA structure	Foraging path
Learning entity	Population of molecules	Ant colony
Relation between elements	Mutation	Reorientation of path segment
Search process	Optimization of structure	Optimization of path
Search space	Sequence space	Three-dimensional space
Random step	Mutation	Segment of ant walk
Self-enhancing process	Replication	Secretion of pheromone
Measure of activity	Mean replication rate	Mean pheromone concentration
Goal of the search	Target structure	Richest food source
Temporary memory	Sequence distribution	Pheromone trail

Learning at population or colony level **by trial and error**

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

1. Folding and inverse folding of RNA
2. Neutral networks
3. Darwinian evolution of RNA
4. Learning by the Darwinian mechanism
5. **Folding kinetics and metastable structures**
6. Intersections and conformational switches



RNA secondary structures derived from a single sequence

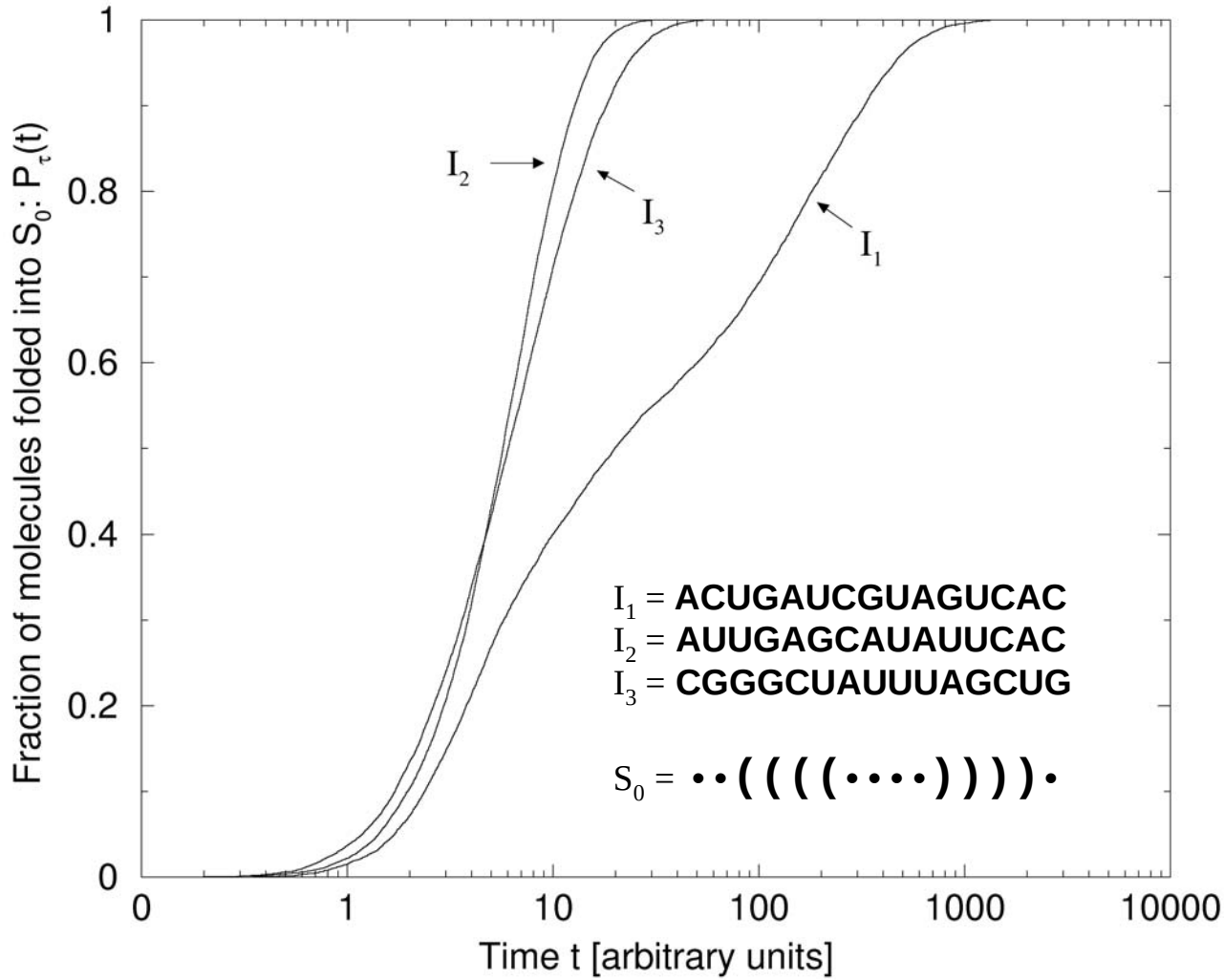
Kinetic Folding of RNA Secondary Structures

Christoph Flamm, Walter Fontana, Ivo L. Hofacker, Peter Schuster. *RNA folding kinetics at elementary step resolution.* RNA **6**:325-338, 2000

Christoph Flamm, Ivo L. Hofacker, Sebastian Maurer-Stroh, Peter F. Stadler, Martin Zehl. *Design of multistable RNA molecules.* RNA **7**:325-338, 2001

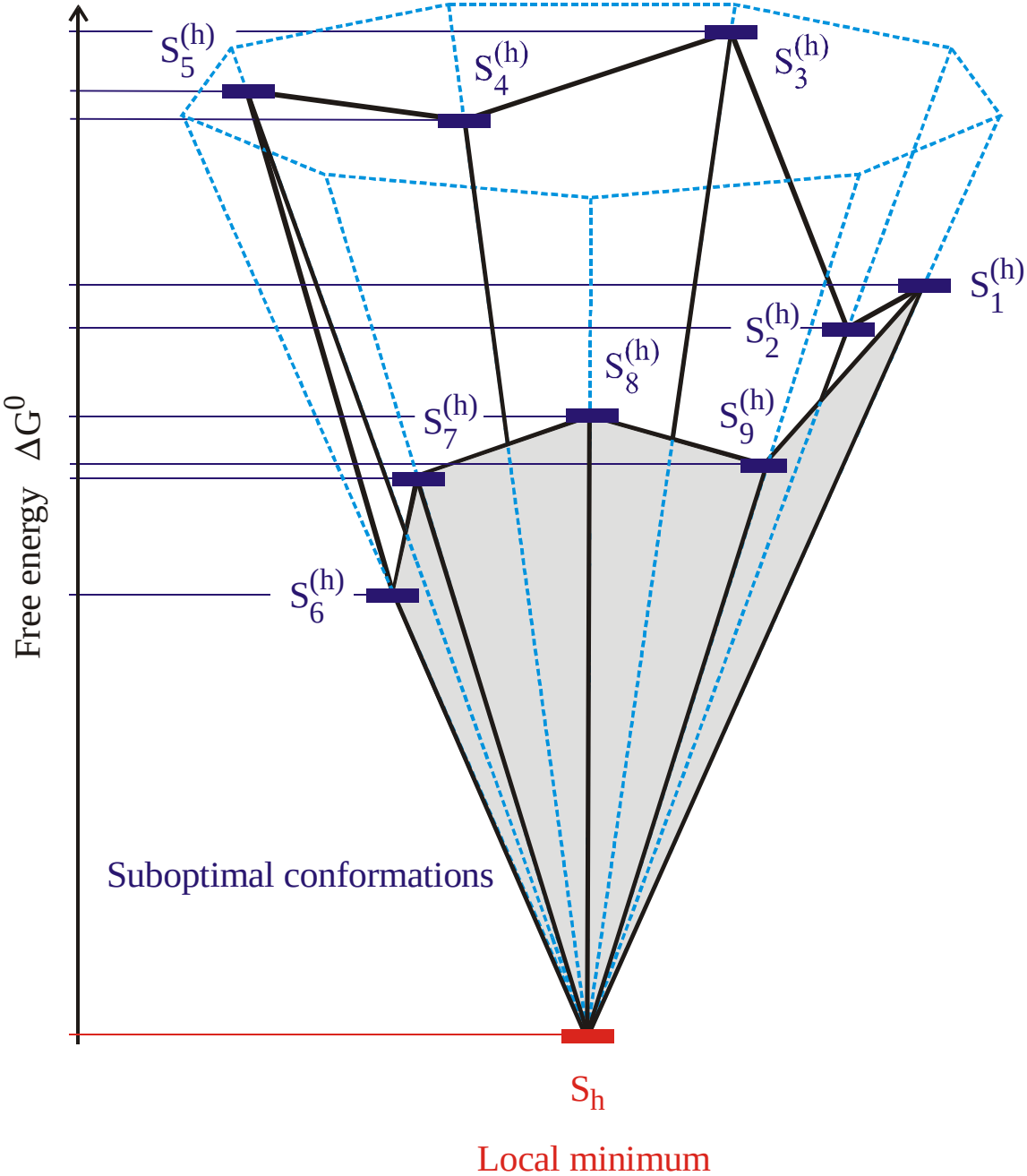
Christoph Flamm, Ivo L. Hofacker, Peter F. Stadler, Michael T. Wolfinger. *Barrier trees of degenerate landscapes.* Z.Phys.Chem. **216**:155-173, 2002

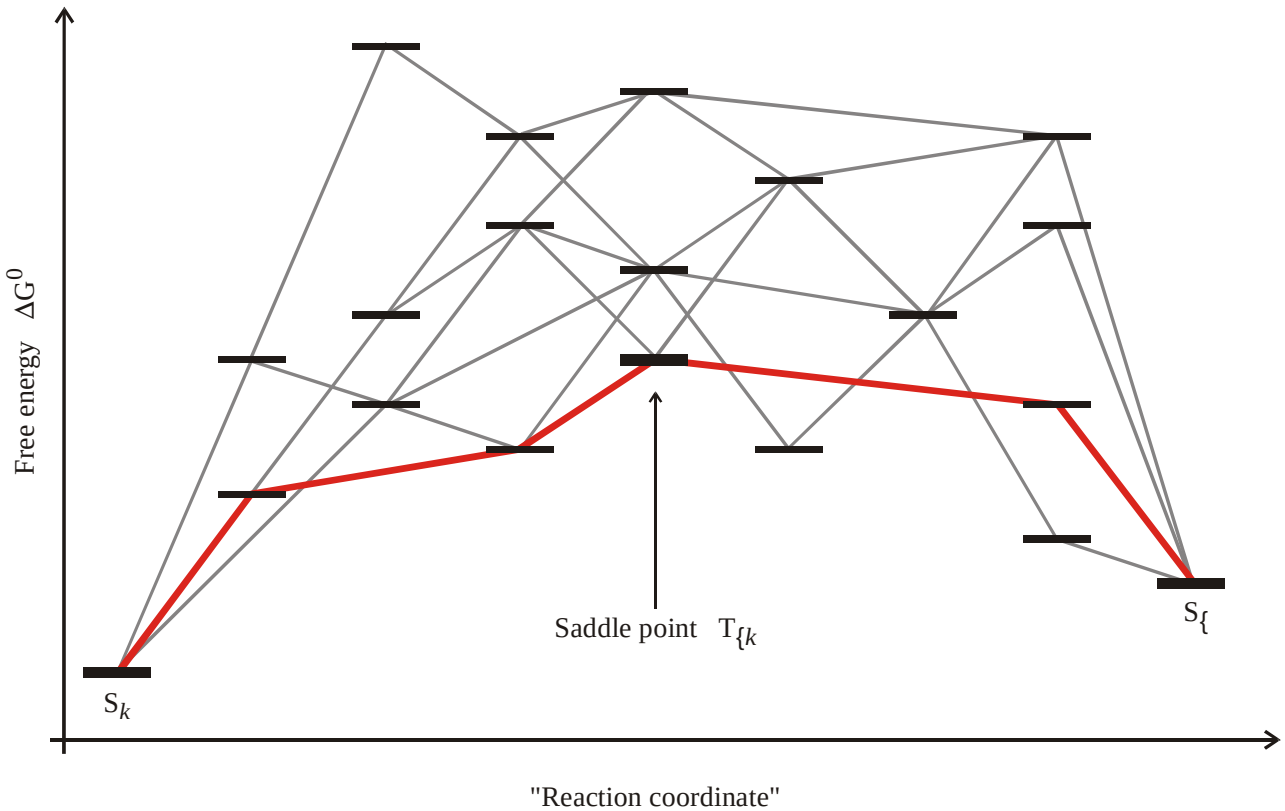
Michael T. Wolfinger, W. Andreas Svrcek-Seiler, Christoph Flamm, Ivo L. Hofacker, Peter F. Stadler. *Efficient computation of RNA folding dynamics.* J.Phys.A: Math.Gen. **37**:4731-4741, 2004

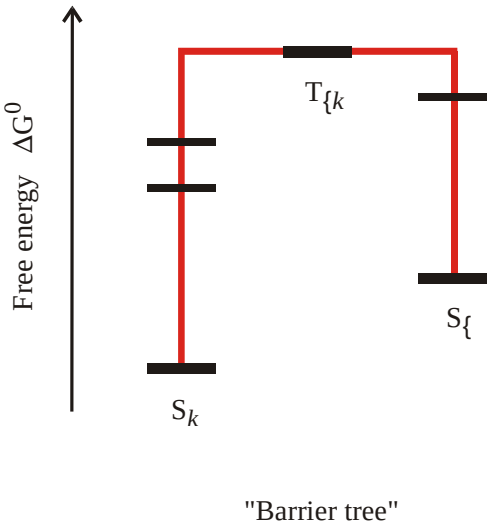


Mean folding curves for three small RNA molecules with different folding behavior

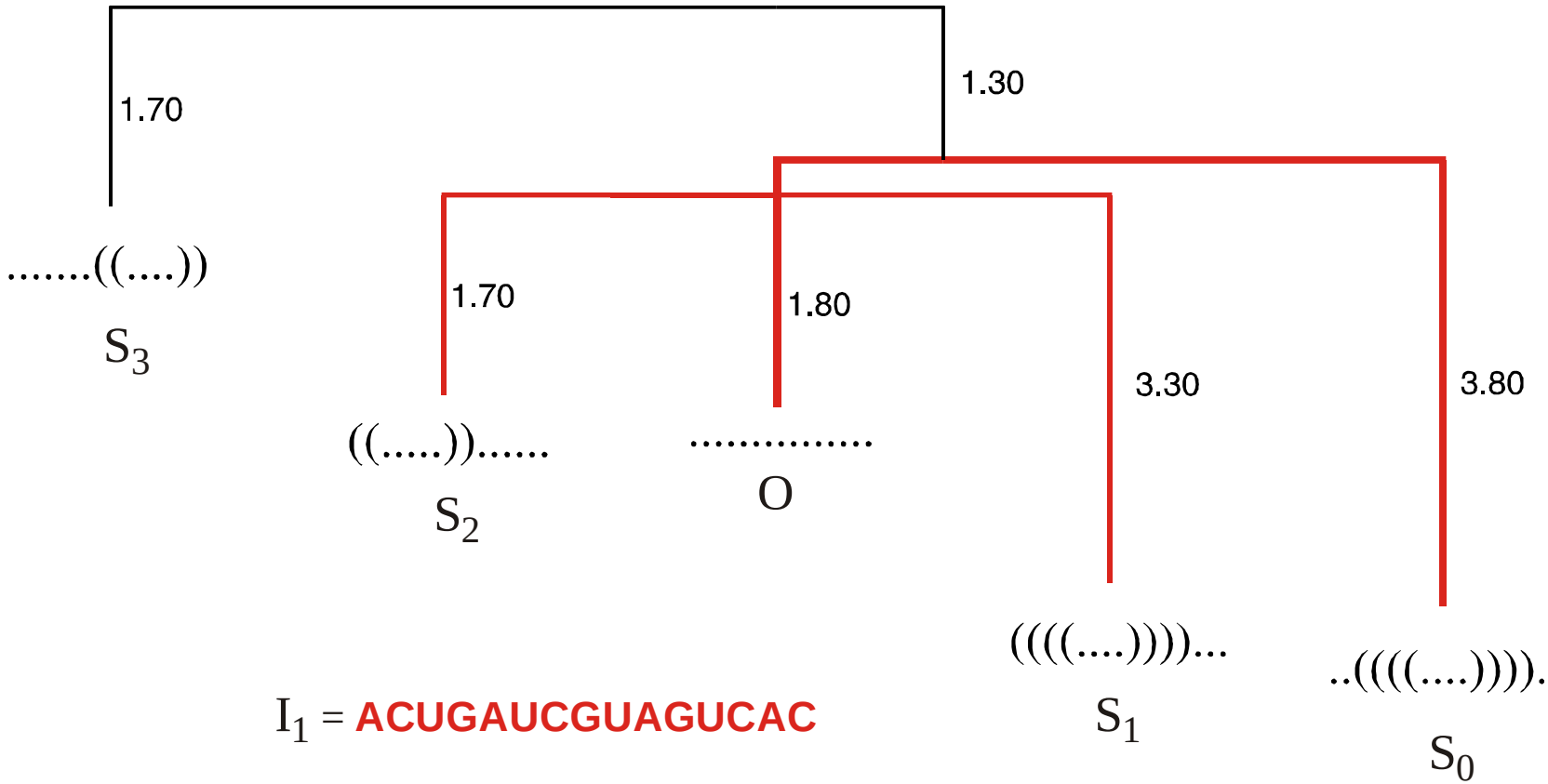
Search for local minima in conformation space



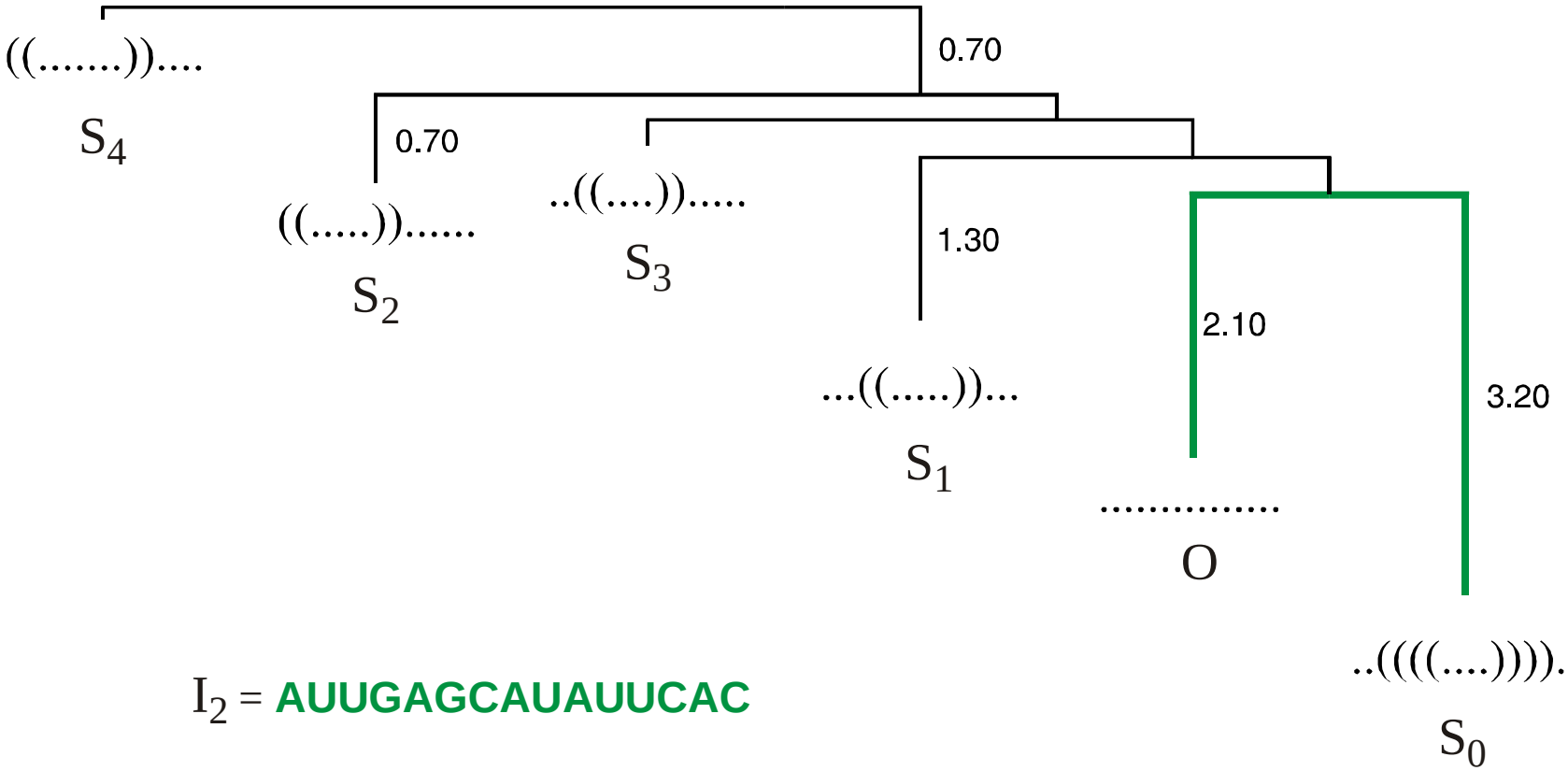




Definition of a ‚barrier tree‘

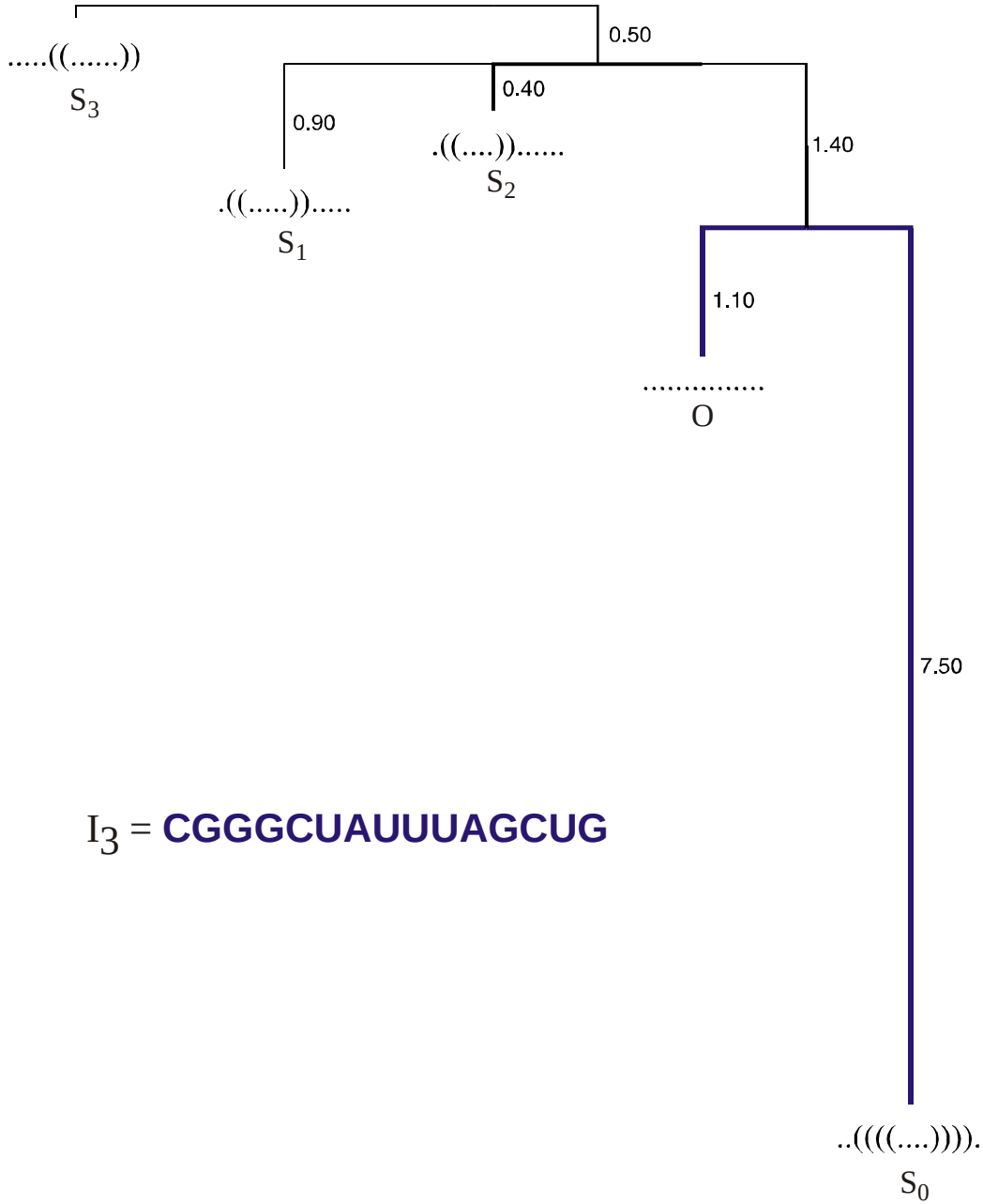


Example of an unefficiently folding small RNA molecule with n = 15



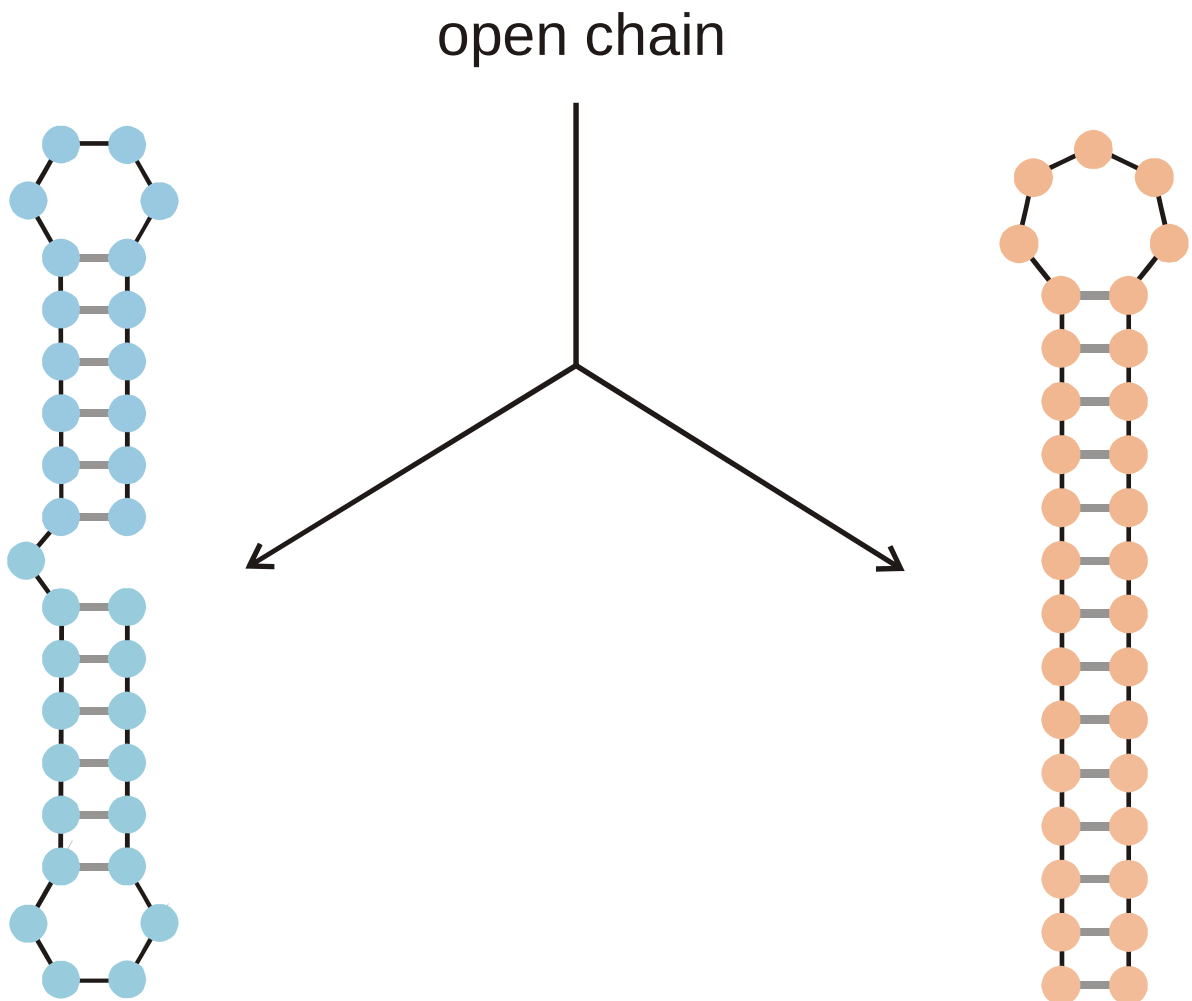
I_2 = **AUUGAGCAUAUUCAC**

Example of an easily folding small RNA molecule with n = 15

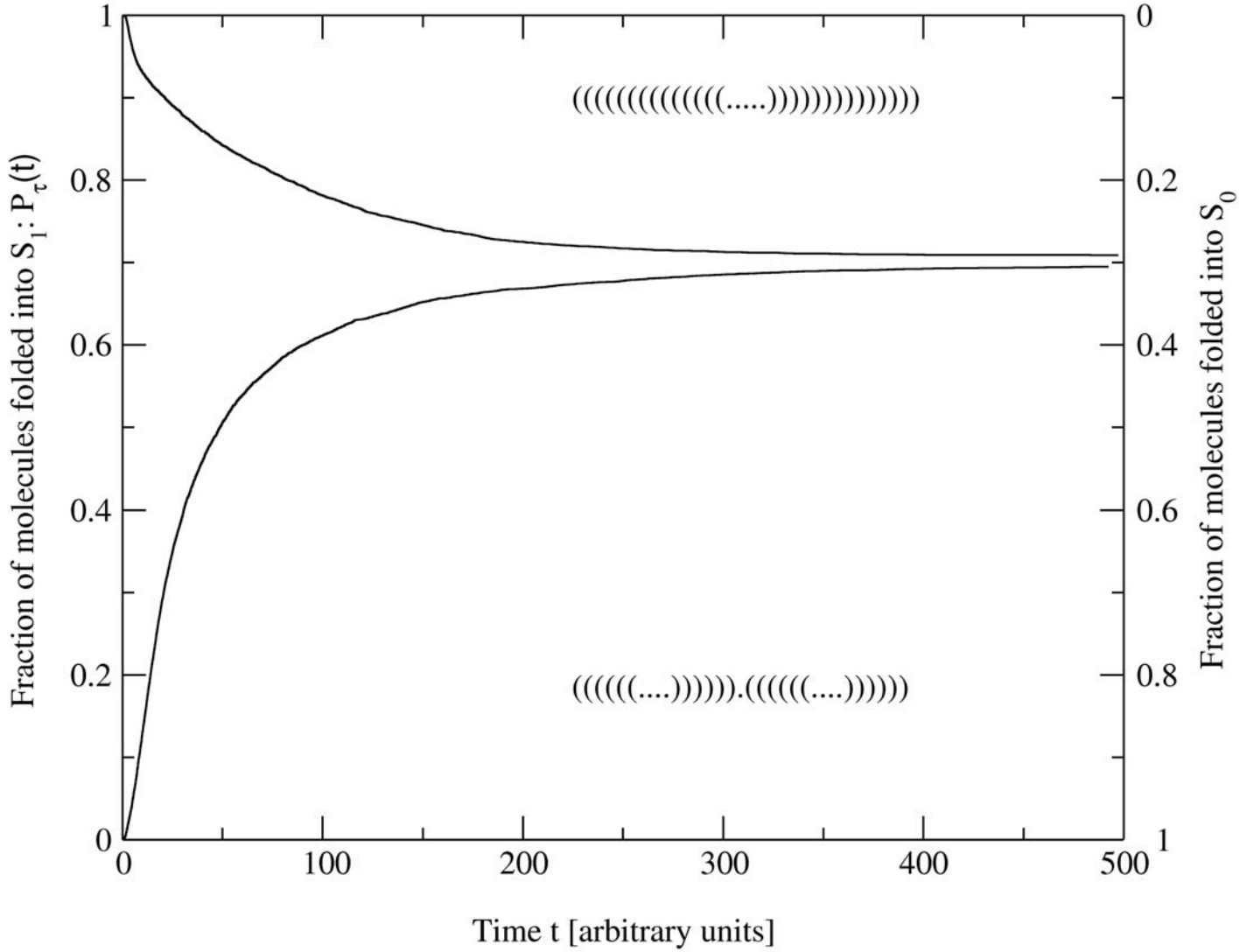


I_3 = **CGGGCUAUUUAGCUG**

Example of an easily folding and especially stable small RNA molecule with n = 15

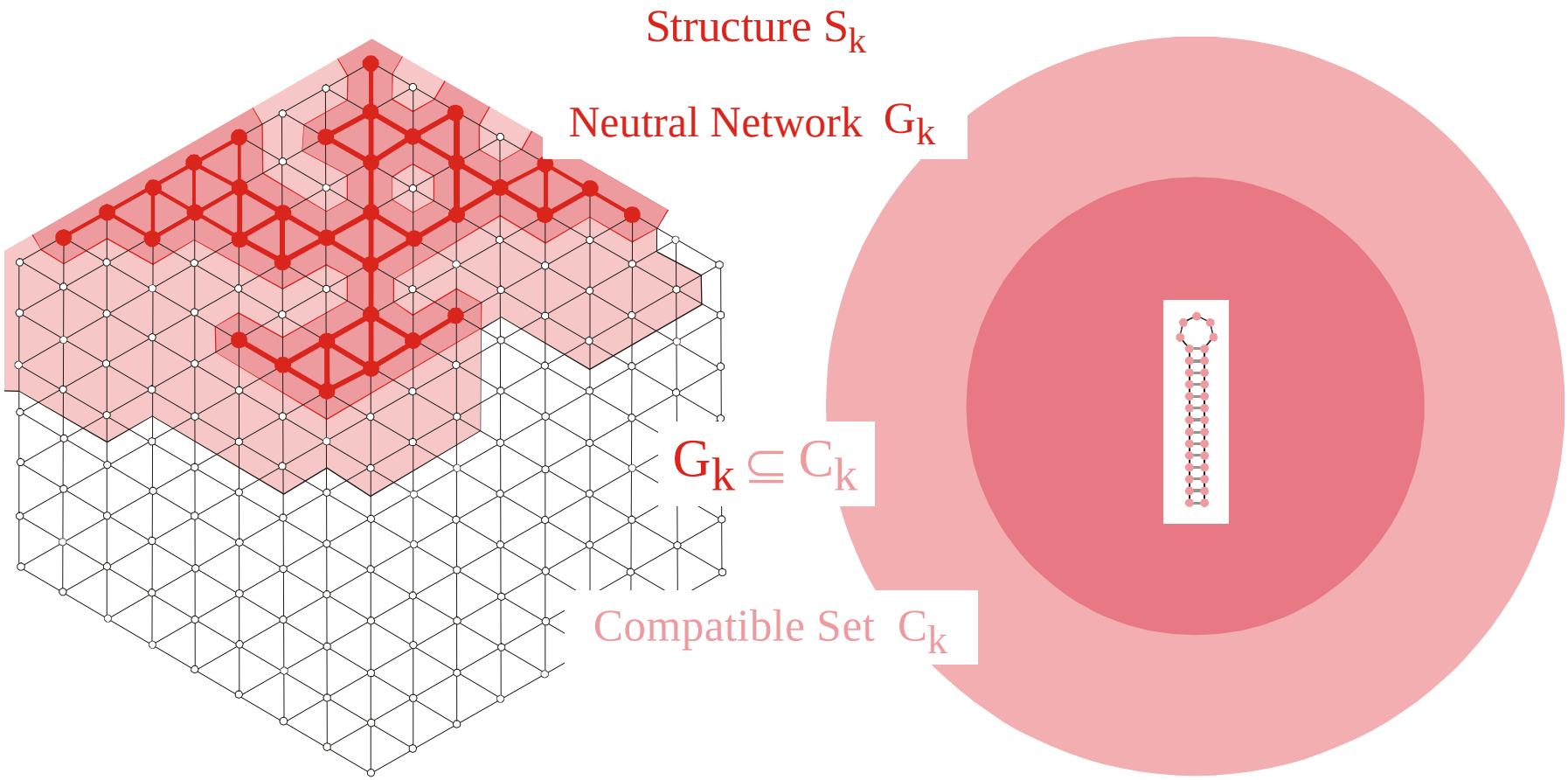


A nucleic acid molecule folding in two dominant conformations

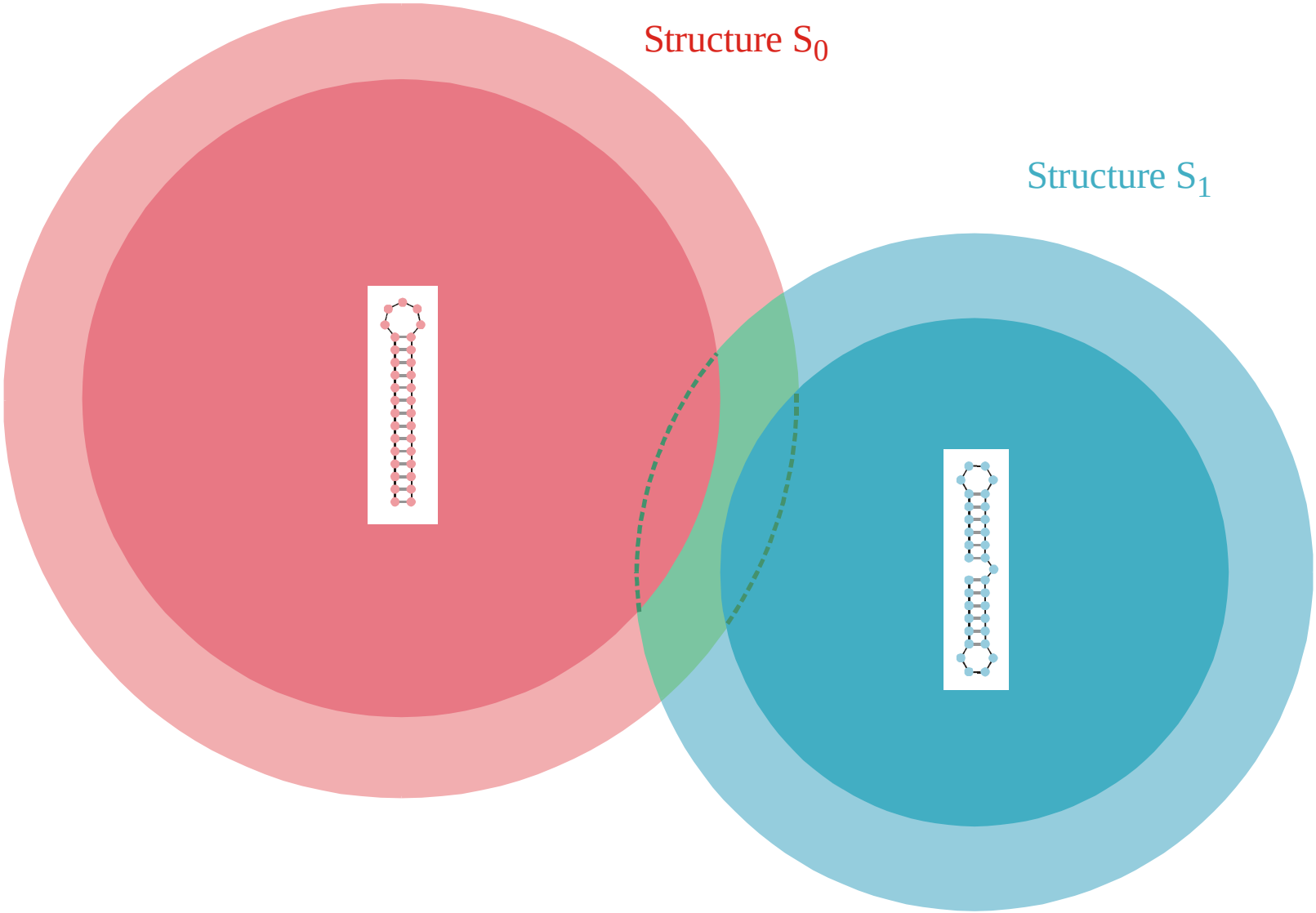


Folding dynamics of the sequence **GGCCCCUUUGGGGGCCAGACCCCUAAAAAGGGUC**

1. Folding and inverse folding of RNA
2. Neutral networks
3. Darwinian evolution of RNA
4. Learning by the Darwinian mechanism
5. Folding kinetics and metastable structures
6. **Intersections and conformational switches**



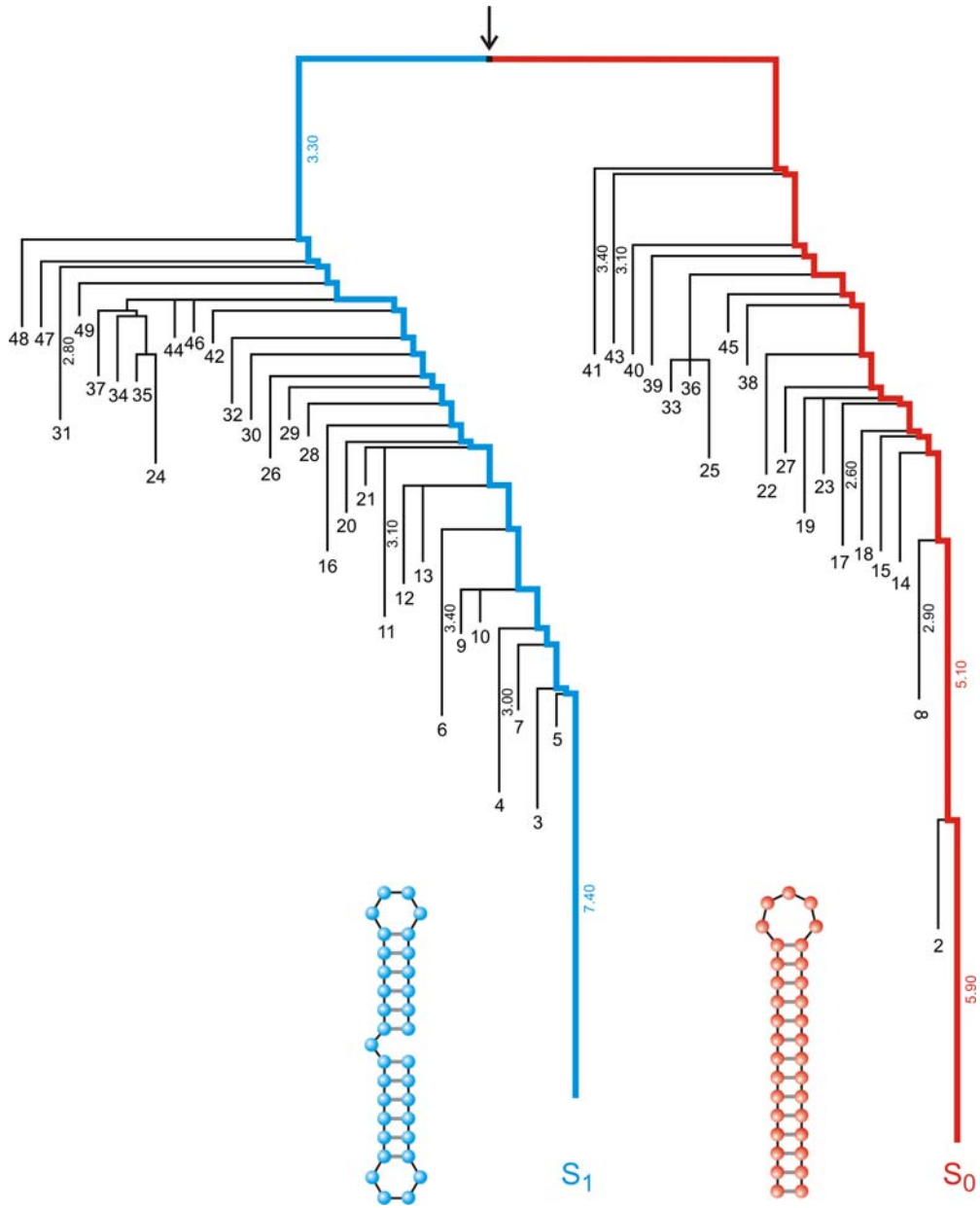
The **compatible set C_k** of a structure **S_k** consists of all sequences which form **S_k** as its minimum free energy structure (the neutral network **G_k**) or one of its suboptimal structures.



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

The intersection of two compatible sets is always non empty: $C_0 \cap C_1 \neq \varnothing$

The barrier tree connecting S_1 and S_0



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*).

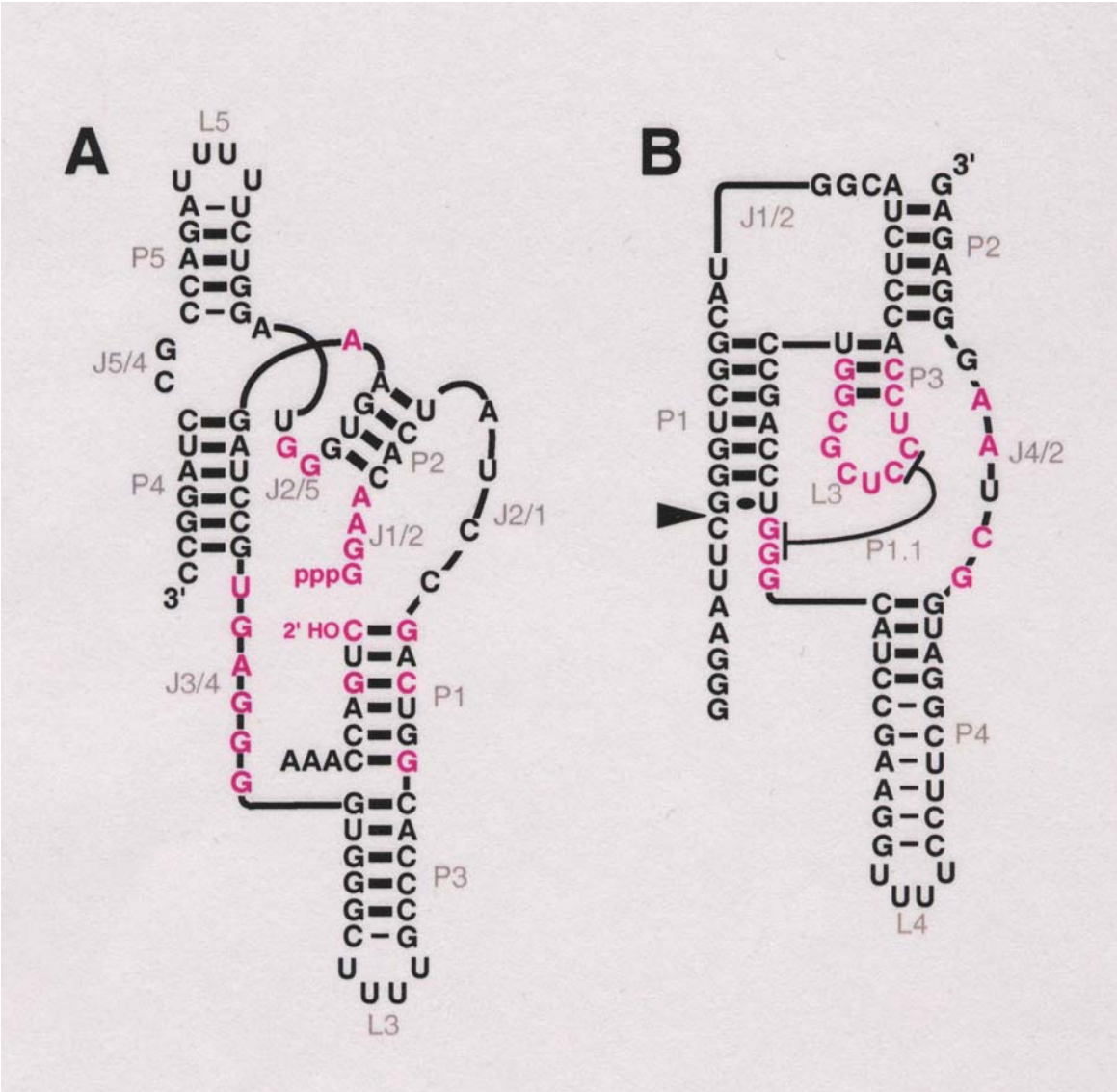
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5*–*8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect. 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.

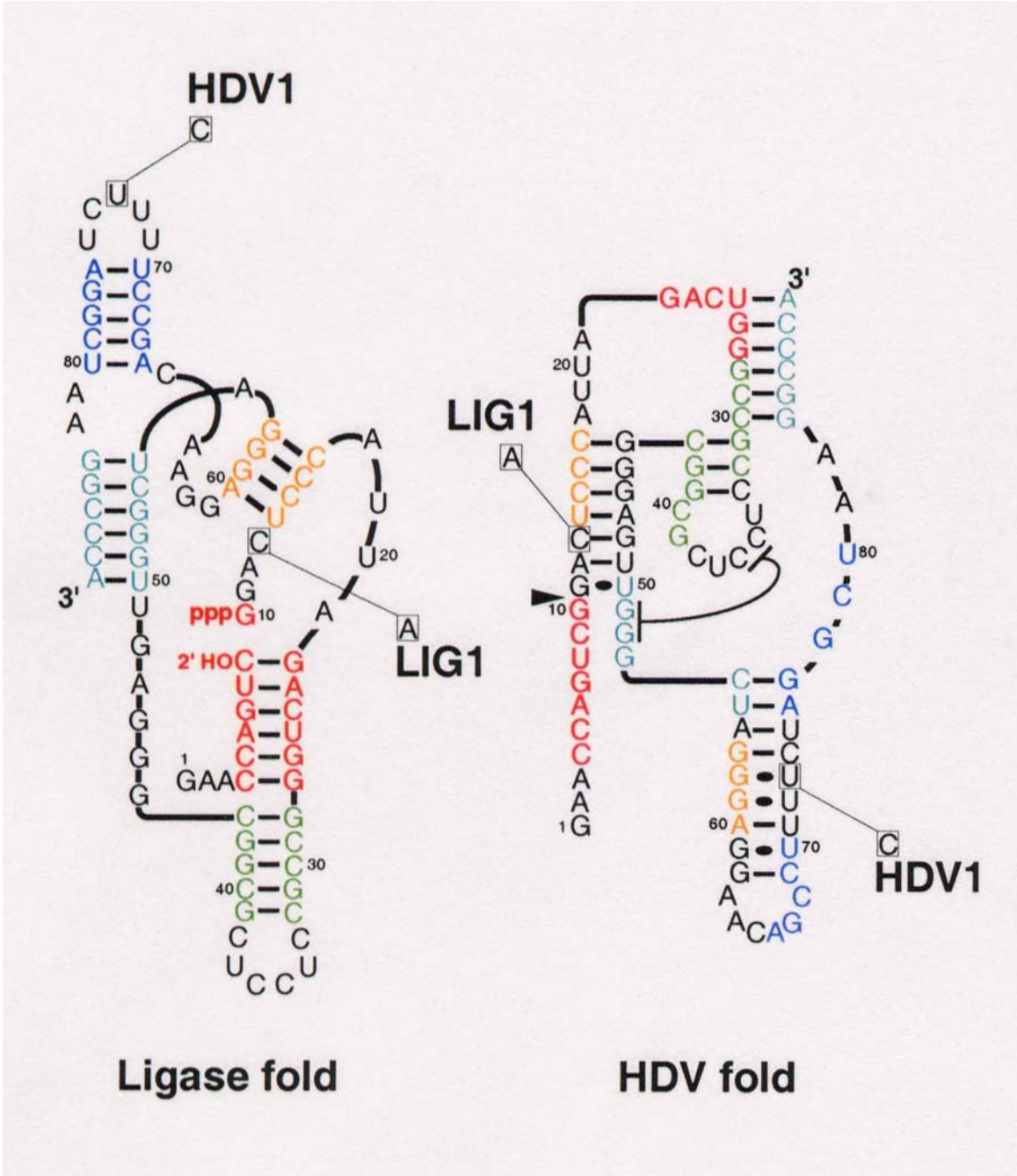
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

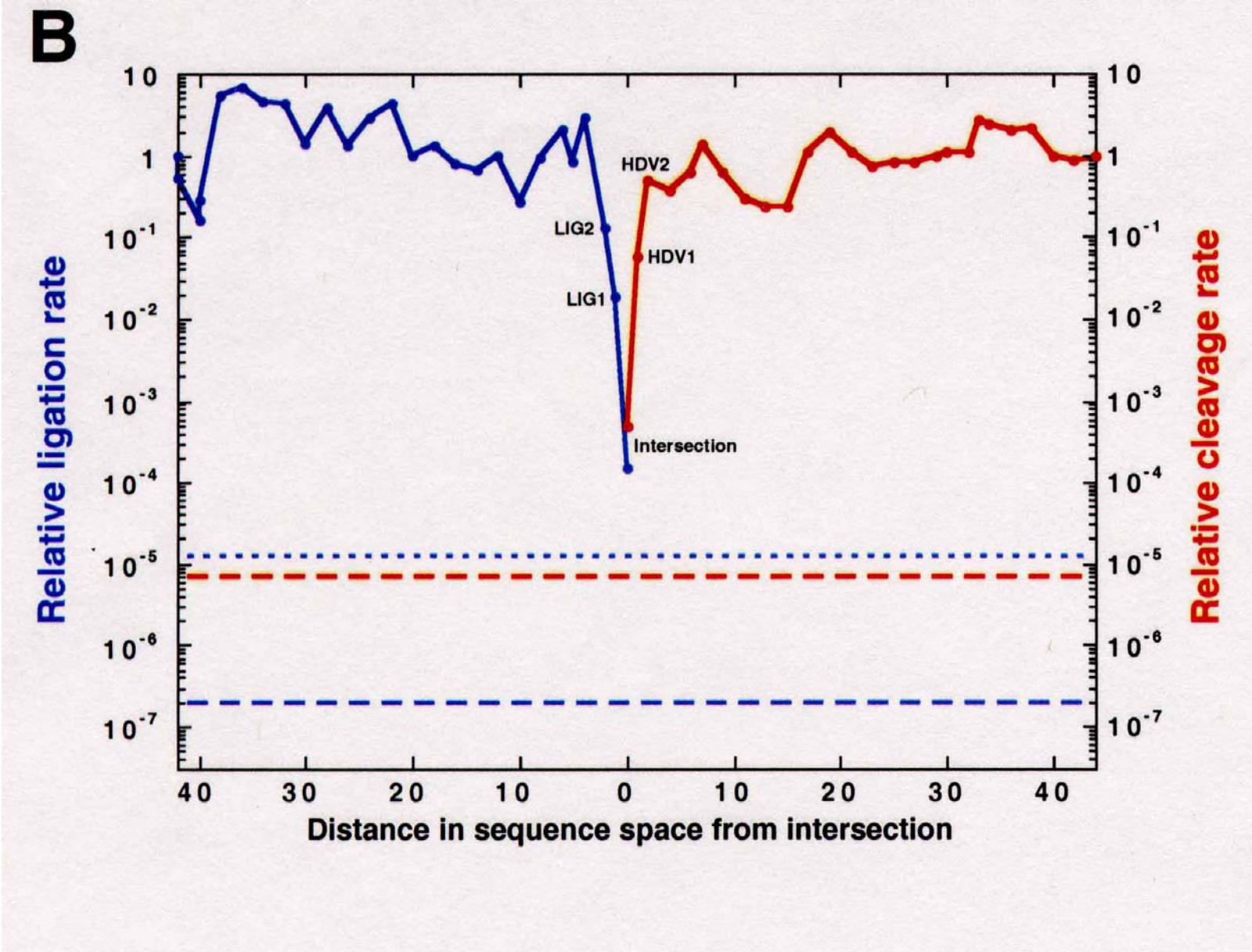


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

The sequence at the ***intersection***:

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

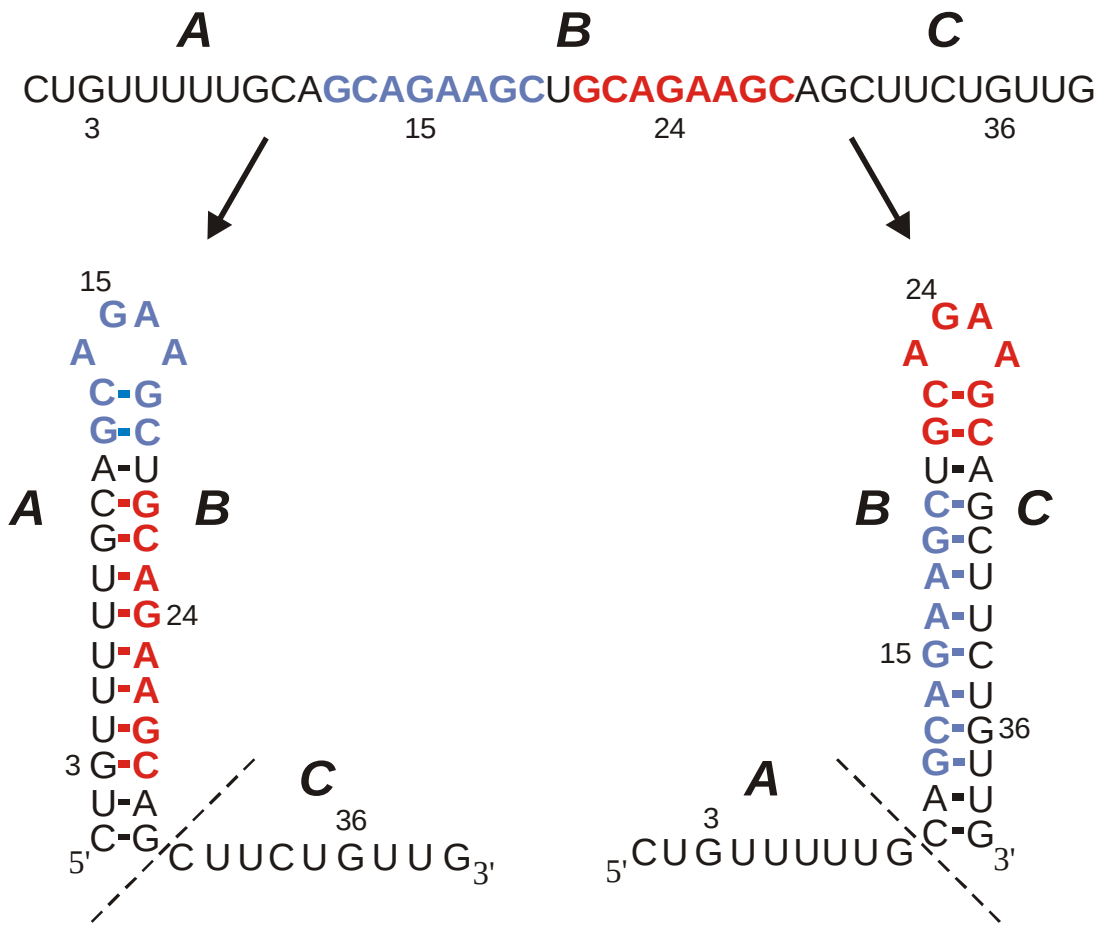




Two neutral walks through sequence space with conservation of structure and catalytic activity

J. H. A. Nagel, C. Flamm, I. L. Hofacker, K. Franke, M. H. de Smit, P. Schuster, and C. W. A. Pleij. ***Structural parameters affecting the kinetic competition of RNA hairpin formation***, in press 2005.

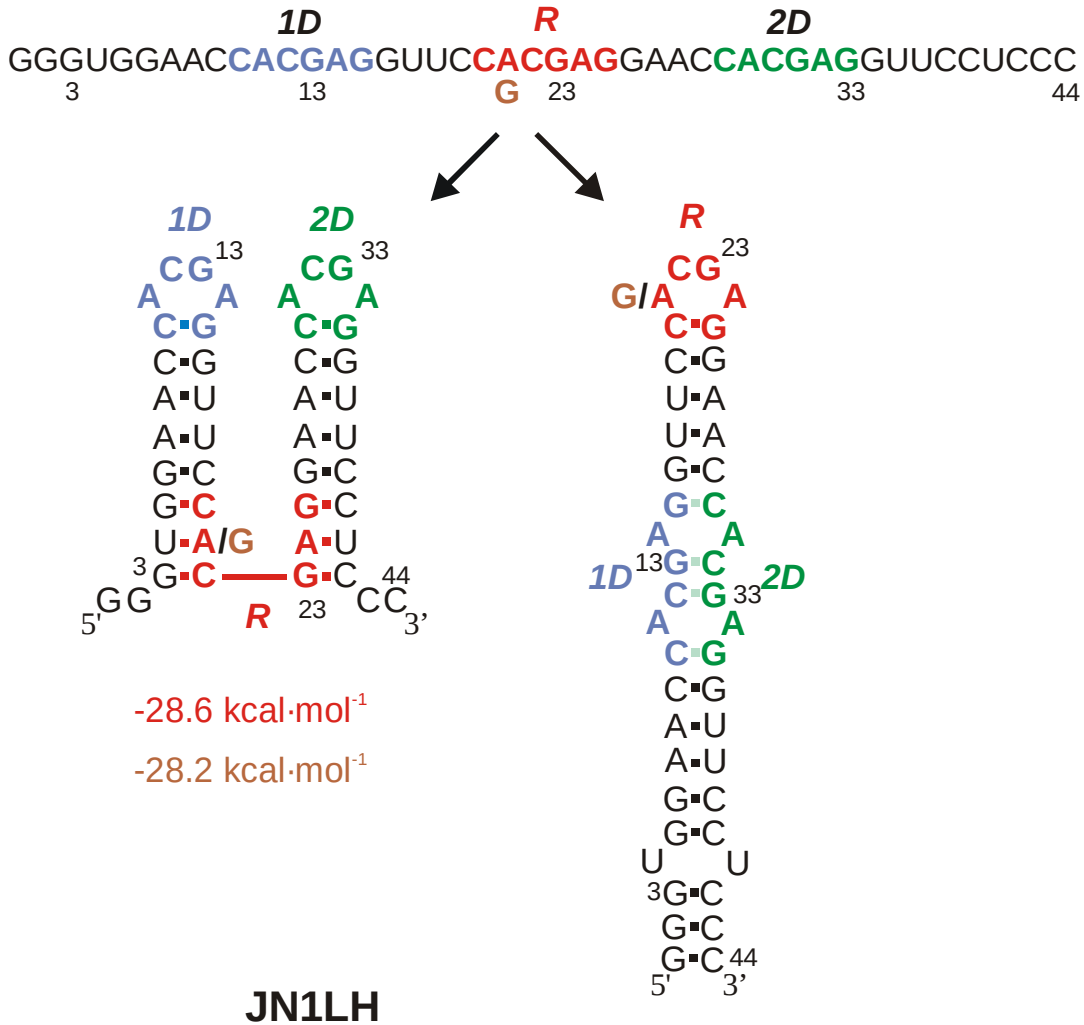
J. H. A. Nagel, J. Møller-Jensen, C. Flamm, K. J. Öistämö, J. Besnard, I. L. Hofacker, A. P. Gultyaev, M. H. de Smit, P. Schuster, K. Gerdes and C. W. A. Pleij. ***The refolding mechanism of the metastable structure in the 5'-end of the*** hok ***mRNA of plasmid*** R1, submitted 2005.



JN2C

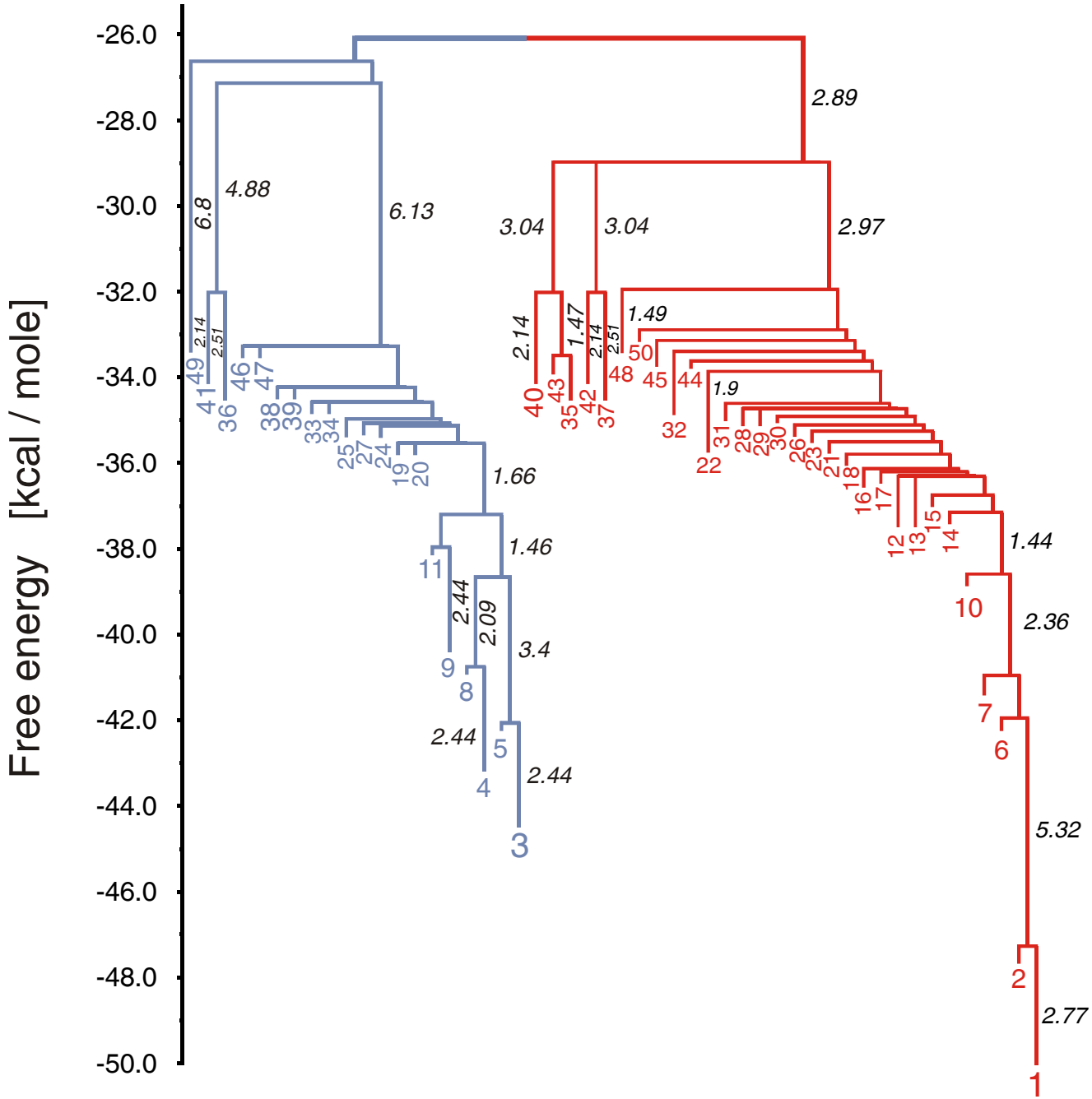
J.H.A. Nagel, C. Flamm, I.L. Hofacker, K. Franke, M.H. de Smit, P. Schuster, and C.W.A. Pleij.

Structural parameters affecting the kinetic competition of RNA hairpin formation, in press 2004.



J.H.A. Nagel, C. Flamm, I.L. Hofacker, K. Franke, M.H. de Smit, P. Schuster, and C.W.A. Pleij.

Structural parameters affecting the kinetic competition of RNA hairpin formation, in press 2004.



J1LH barrier tree

Conclusions

I. The Darwinian mechanism of optimization through variation and selection operates equally well on simple and complex repoducing elements because only the number of fertile offspring counts.

II. Darwinian learning through trial and error takes place on the level of populations. It does not require sophisticated elements and occurs even with self-replicating molecules.

III. Even simple molecules have the capacity for a rich repertoire of properties and interactions. For example, they can have multiple structures and functions.

Acknowledgement of support

Universität Wien

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

Austrian Genome Research Program – GEN-AU

Siemens AG, Austria

Universität Wien and the Santa Fe Institute

Coworkers

Universität Wien

Walter Fontana, Harvard Medical School, MA

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

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

Jord Nagel, **Kees Pleij**, Universiteit Leiden, NL

Ivo L.Hofacker, **Christoph Flamm**, **Andreas Svrček-Seiler**
Universität Wien, AT

Stefan Bernhart, Ulrike Langhammer, **Ulrike Mückstein**,
Universität Wien, AT

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

Andreas Wernitznig, **Michael Kospach**, **Kurt Grünberger**, **Stefan Wuchty**

Web-Page for further information:

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