Modeling in Molecular Biology

Peter Schuster

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

Third GEN-AU Summer School: Ultra-Sensitive Proteomics and Genomics

Litschau, 29.– 31.08.2005

Web-Page for further information:

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

Mathematical models

Discrete methods

Enumeration, combinatorics
Graph theory, network theory
String matching (sequence comparison)

Continuous methods

Differentiation, Integration
Optimization

Dynamical systems

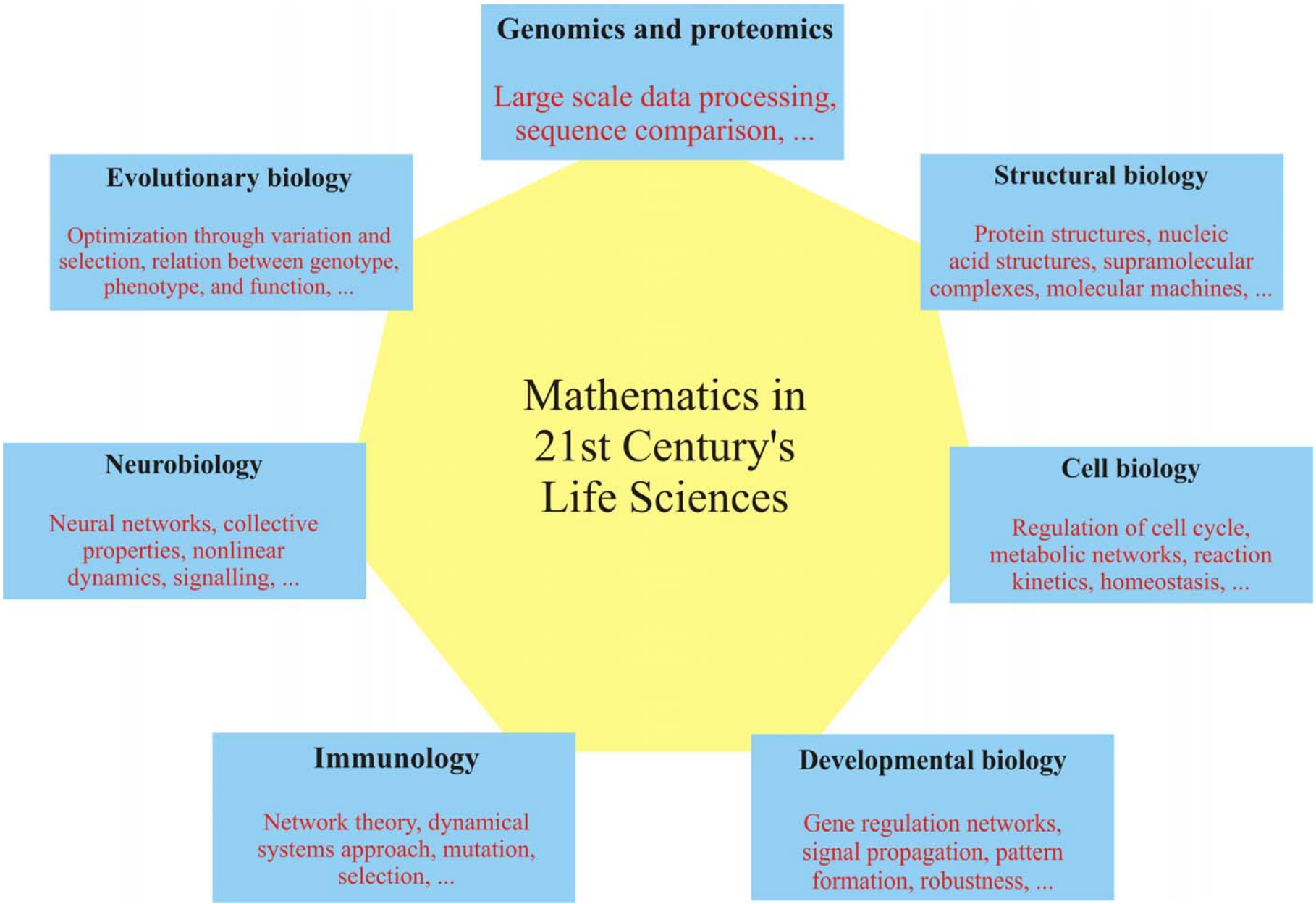
Stochastic difference equs.		Difference equs.		Stochastic differential equs.		Differential equs.
Δn , Δt	$\leftrightarrow$	dn , Δt	$\leftrightarrow$	Δn , dt	$\leftrightarrow$	dn , dt … ODE
Δn , Δx , Δt	$\leftrightarrow$	dn , Δx , Δt	$\leftrightarrow$	Δn , dx , dt	$\leftrightarrow$	dn , dx , dt … PDE

Simulation methods

Cellular automata Genetic algorithms Neural networks Simulated annealing Differential equs.

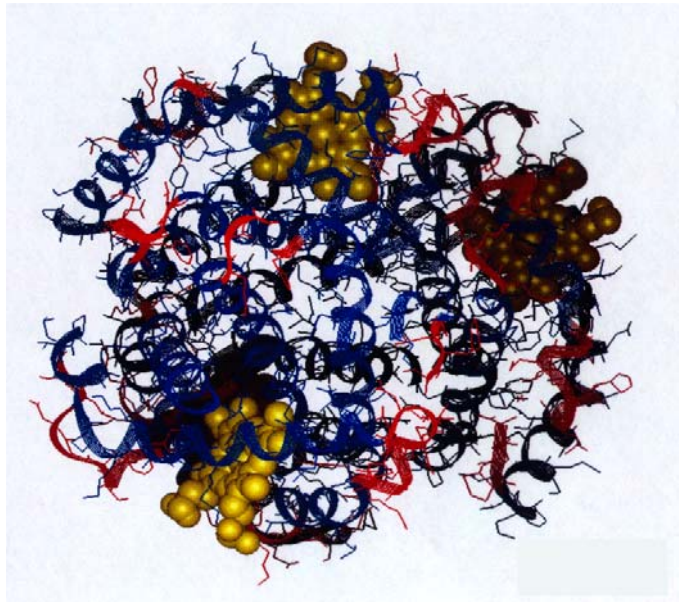
Structure prediction and optimization

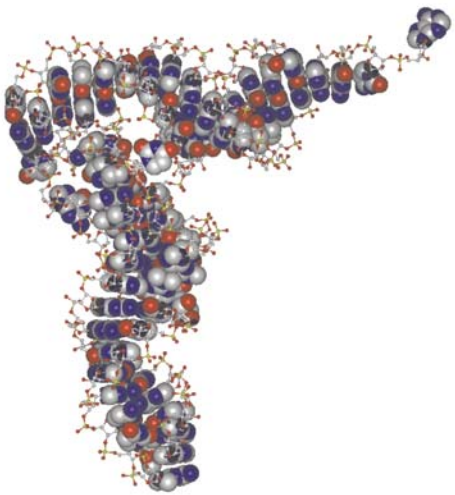
Discrete states		Continuous states
Dynamic programming	Simplex methods	Gradient techniques
RNA secondary structures, lattice proteins		3D-Structures of (bio)molecules

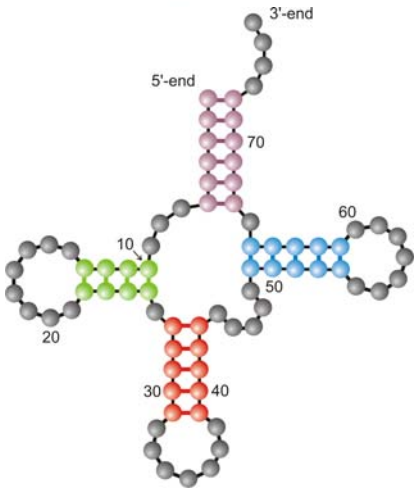

Genomics and proteomics
Large scale data processing, sequence comparison, ...
Evolutionary biology
Optimization through variation and selection, relation between genotype, phenotype, and function, ...
Structural biology
Protein structures, nucleic acid structures, supramolecular complexes, molecular machines, ...
Mathematics in 21st Century's Life Sciences
Neurobiology
Neural networks, collective properties, nonlinear dynamics, signalling, ...
Cell biology
Regulation of cell cycle, metabolic networks, reaction kinetics, homeostasis, ...
Immunology
Network theory, dynamical systems approach, mutation, selection, ...
Developmental biology
Gene regulation networks, signal propagation, pattern formation, robustness, ...

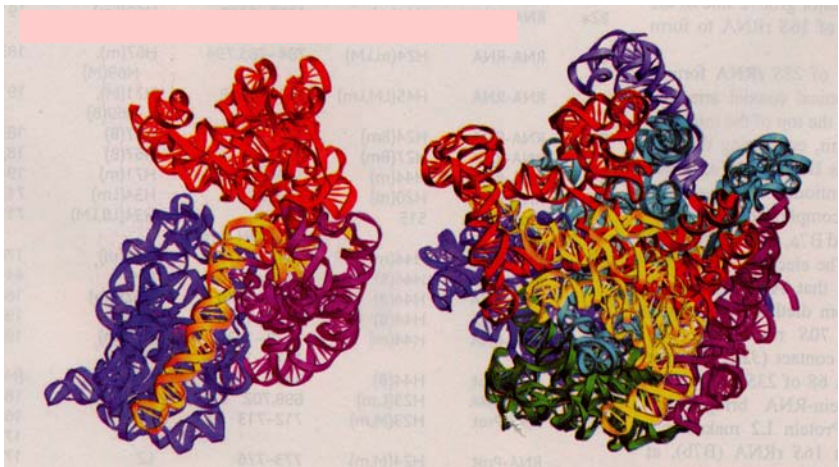
Structural biology

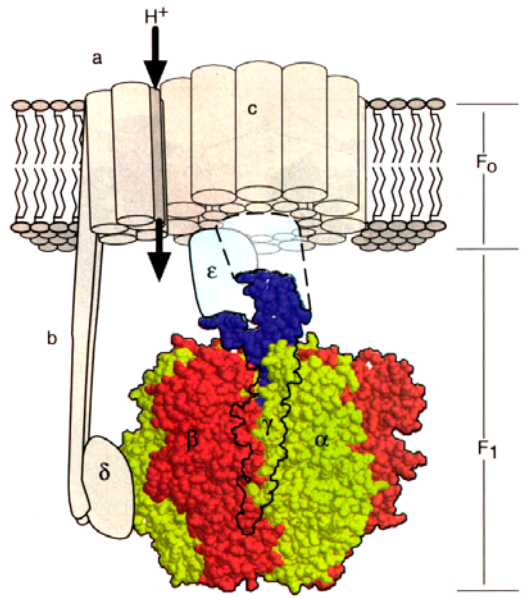
Protein structures, nucleic acid structures, supramolecular complexes, molecular machines, ...

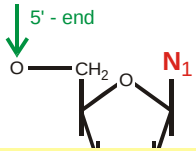




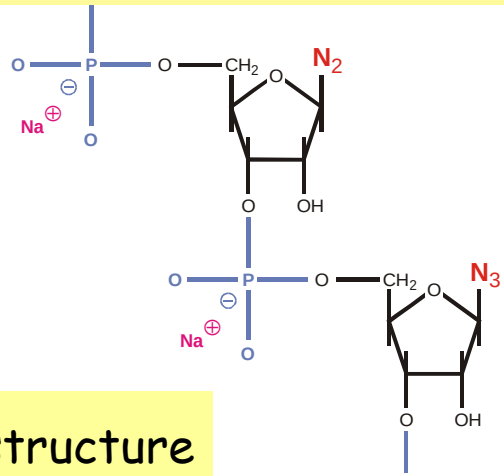


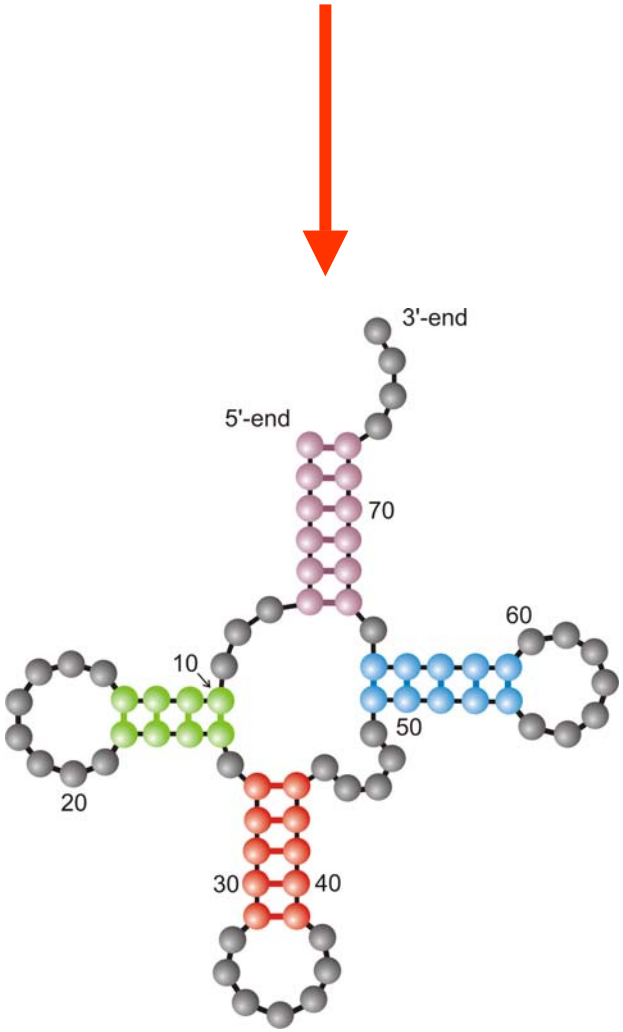




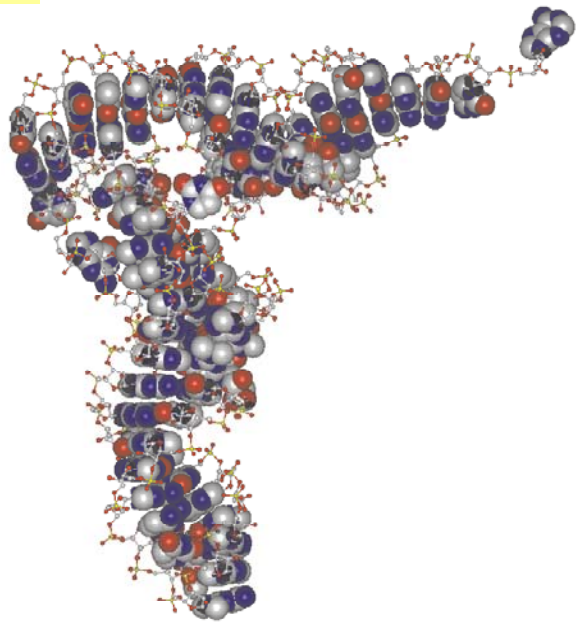


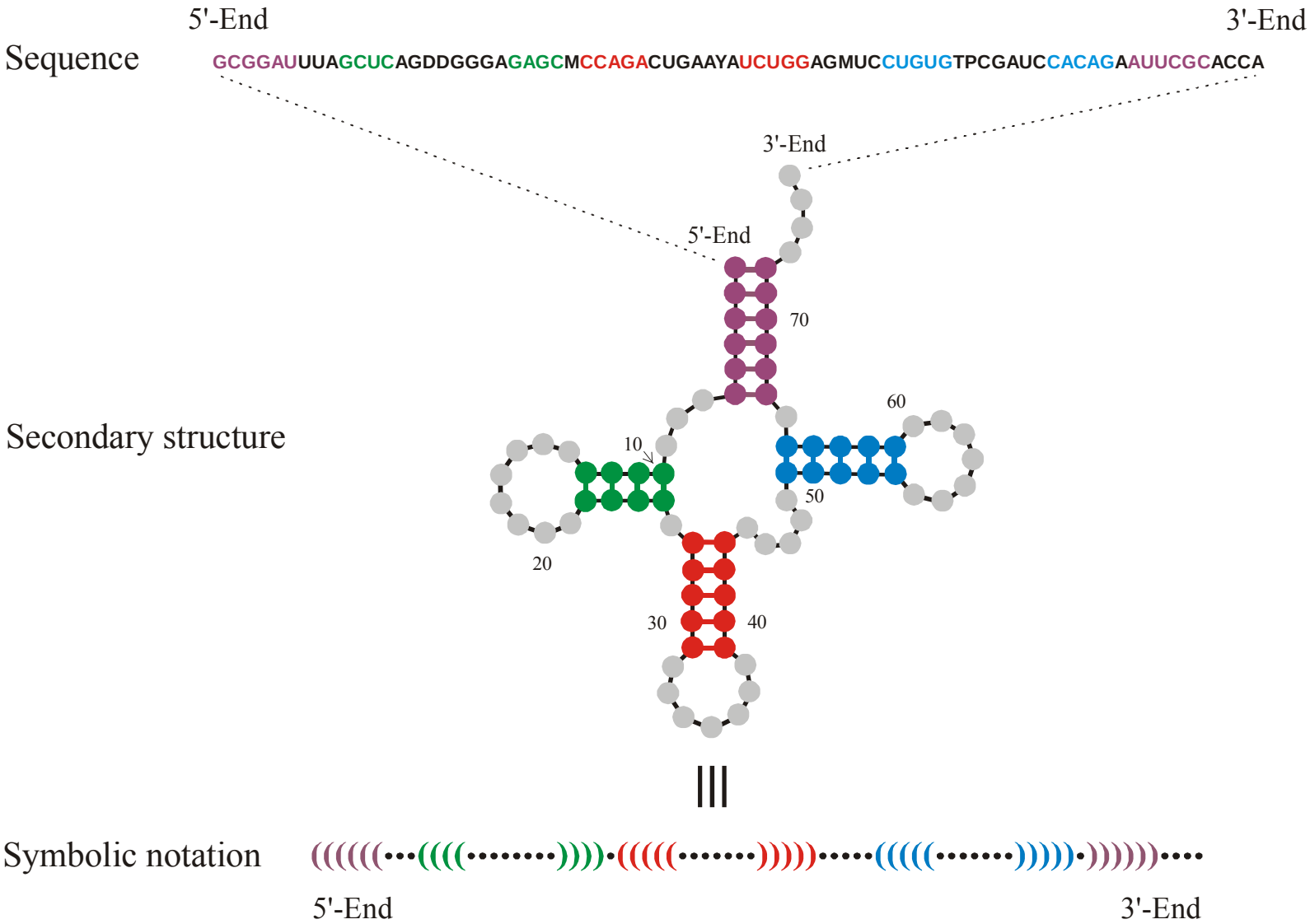
5'-end GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'-end





Definition of RNA structure





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

RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

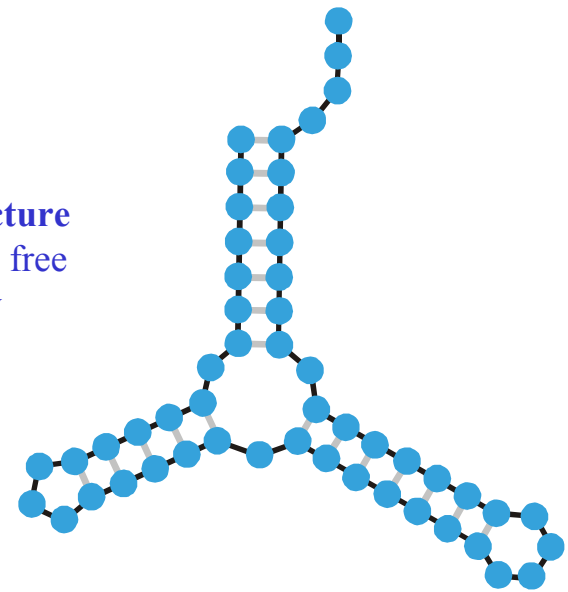
RNA folding:

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

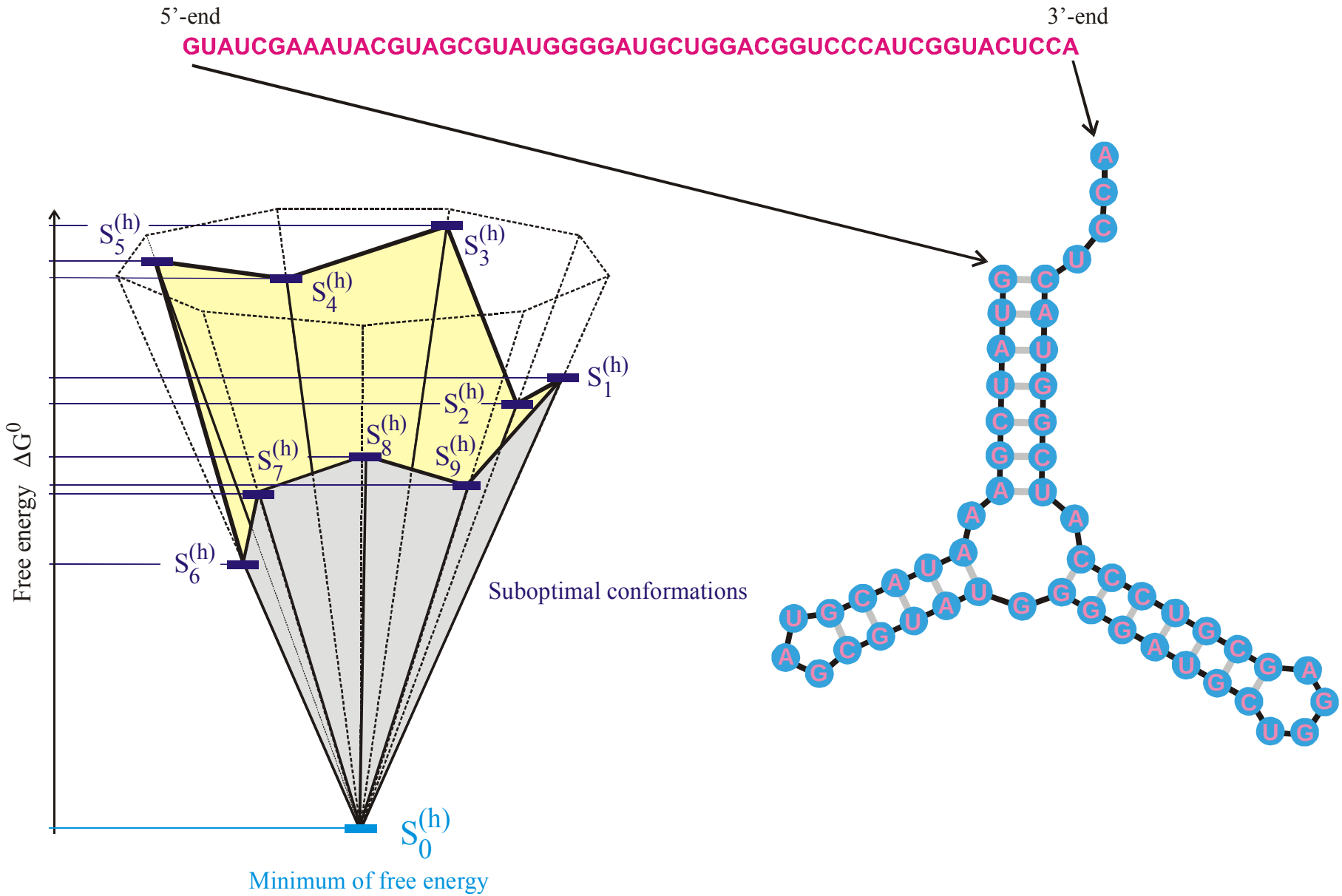
Biophysical chemistry: thermodynamics and kinetics

Empirical parameters

RNA structure of minimal free energy



Sequence, structure, and design



The minimum free energy structures on a discrete space of conformations

RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

RNA folding:

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

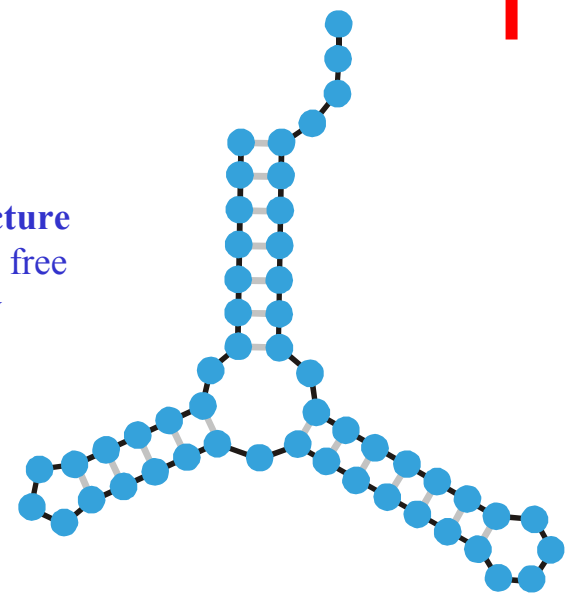
Iterative determination of a sequence for the given secondary structure

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

The ***Vienna RNA-Package***:

A library of routines for folding, ***inverse folding***, sequence and structure alignment, ***kinetic folding***, ***cofolding***, …

Monatshefte für Chemie 125, 167–188 (1994)

Monatshefte für Chemie
Chemical Monthly

Printed in Austria

Fast Folding and Comparison of RNA Secondary Structures

I. L. Hofacker[1,*], W. Fontana[3], P. F. Stadler[1,3], L. S. Bonhoeffer[4], M. Tacker[1] and P. Schuster[1,2,3]

[1] Institut für Theoretische Chemie, Universität Wien, A-1090 Wien, Austria
[2] Institut für Molekulare Biotechnologie, D-07745 Jena, Federal Republic of Germany
[3] Santa Fe Institute, Santa Fe, NM 87501, U.S.A.
[4] Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

Summary. Computer codes for computation and comparison of RNA secondary structures, the Vienna RNA package, are presented, that are based on dynamic programming algorithms and aim at predictions of structures with minimum free energies as well as at computations of the equilibrium partition functions and base pairing probabilities.

An efficient heuristic for the inverse folding problem of RNA is introduced. In addition we present compact and efficient programs for the comparison of RNA secondary structures based on tree editing and alignment.

All computer codes are written in ANSI C. They include implementations of modified algorithms on parallel computers with distributed memory. Performance analysis carried out on an Intel Hypercube shows that parallel computing becomes gradually more and more efficient the longer the sequences are.

Keywords. Inverse folding; parallel computing; public domain software; RNA folding; RNA secondary structures; tree editing.

Schnelle Faltung und Vergleich von Sekundärstrukturen von RNA

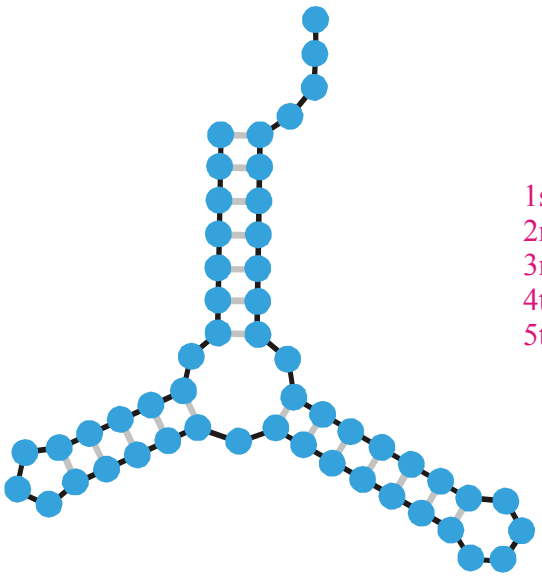
Zusammenfassung. Die im Vienna RNA package enthaltenen Computer Programme für die Berechnung und den Vergleich von RNA Sekundärstrukturen werden präsentiert. Ihren Kern bilden Algorithmen zur Vorhersage von Strukturen minimaler Energie sowie zur Berechnung von Zustandssumme und Basenpaarungswahrscheinlichkeiten mittels dynamischer Programmierung.

Ein effizienter heuristischer Algorithmus für das inverse Faltungsproblem wird vorgestellt. Darüberhinaus präsentieren wir kompakte und effiziente Programme zum Vergleich von RNA Sekundärstrukturen durch Baum-Editierung und Alignierung.

Alle Programme sind in ANSI C geschrieben, darunter auch eine Implementation des Faltungsalgorithmus für Parallelrechner mit verteiltem Speicher. Wie Tests auf einem Intel Hypercube zeigen, wird das Parallelrechnen umso effizienter je länger die Sequenzen sind.

1. Introduction

Recent interest in RNA structures and functions was caused by their catalytic capacities [1, 2] as well as by the success of selection methods in producing RNA



Minimum free energy criterion

1st
2nd
3rd trial
4th
5th

GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

UGGUUACGCGUUGGGGUAACGAAGAUUCCGAGAGGAGUUUAGUGACUAGAGG

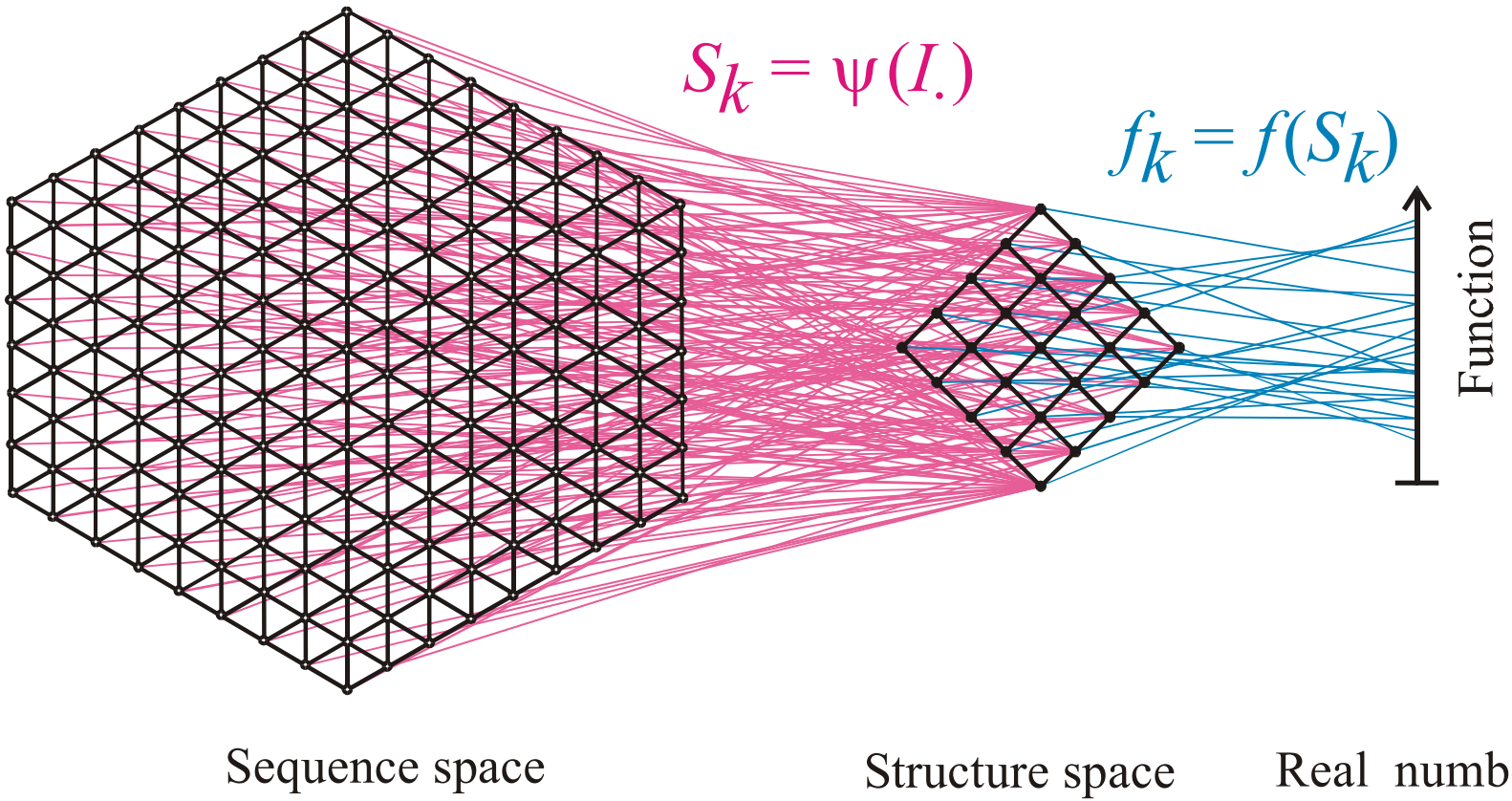
CUUCUUGAGCUAGUACCUAGUCGGAUAGGAUUUCCUAUCUCCAGGGAGGAUG

CUUUUCUUCACGUUAGAUGUGUAAUGGACAUGUGUUUAUUUAGGAAAGGCGC

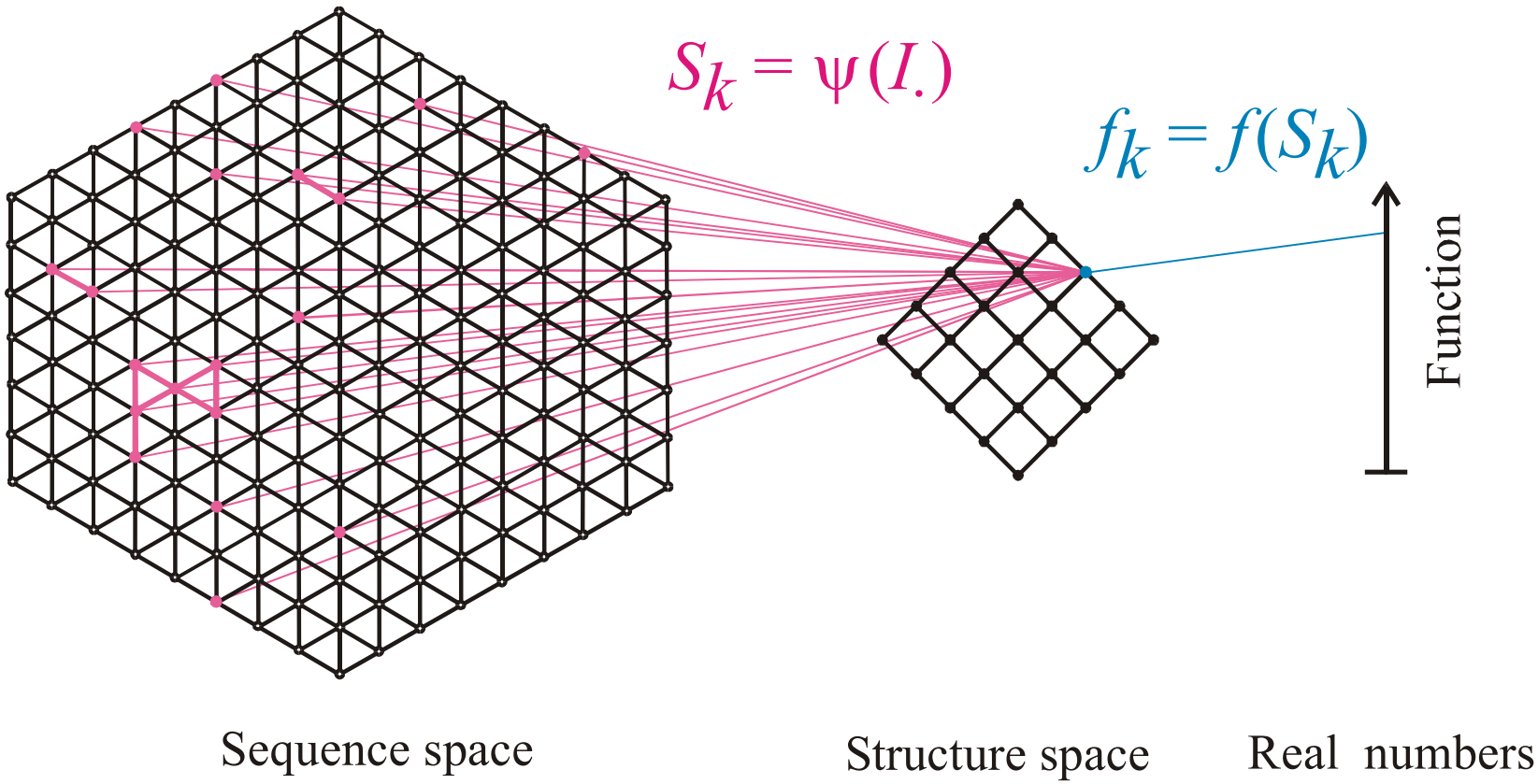
AUAACGUGAGUGUCUAAUACUGAUCGCUCCGGAGGGUGGUGGCGUUGUUAAU

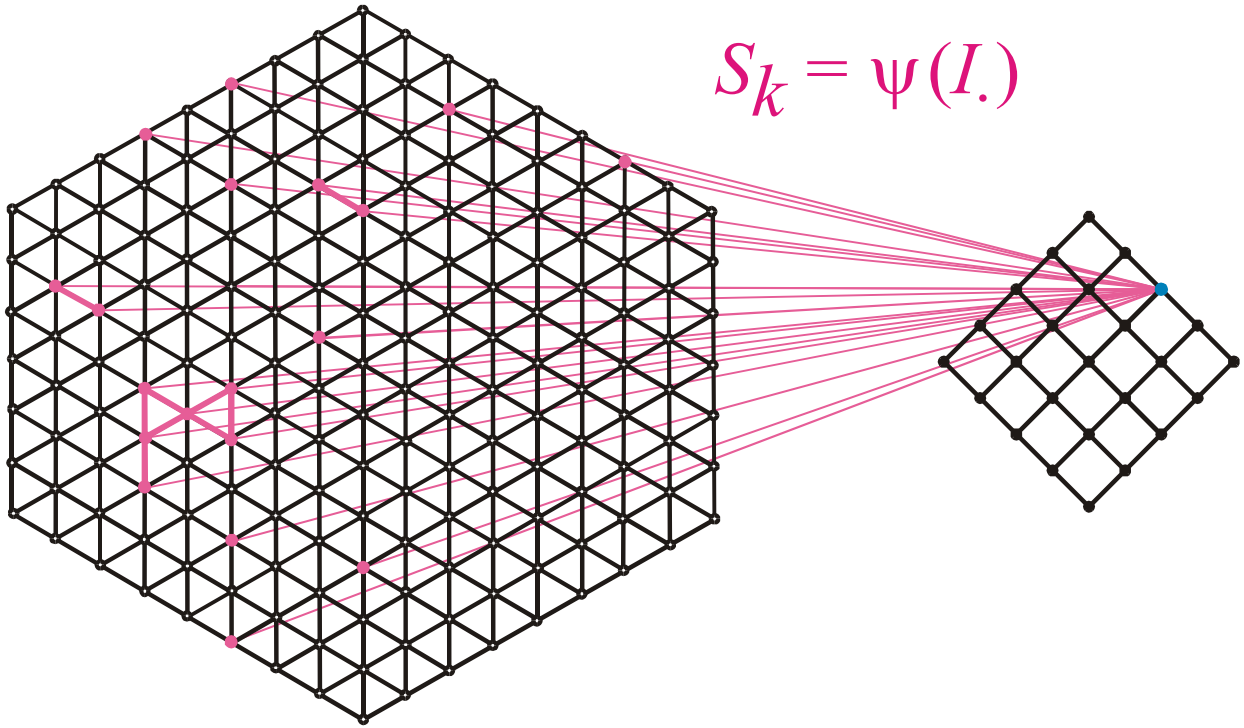
Inverse folding of RNA secondary structures

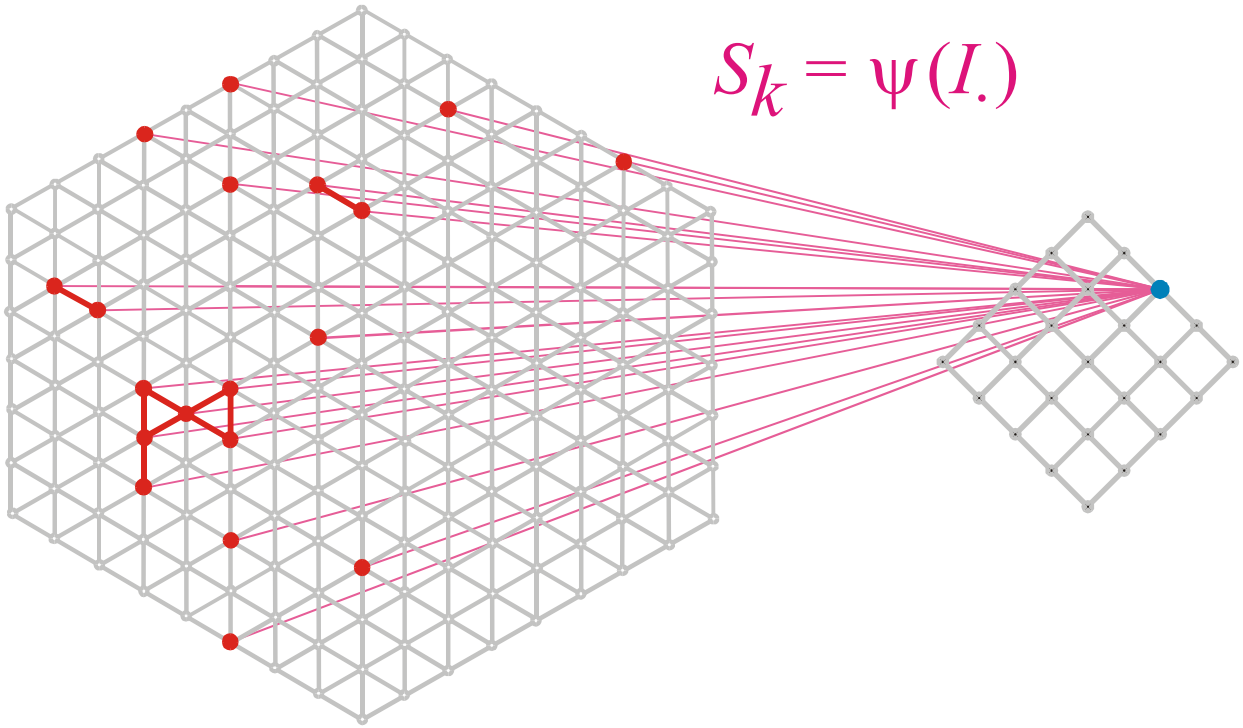
The inverse folding algorithm searches for sequences that form a given RNA secondary structure under the minimum free energy criterion.



Mapping from sequence space into structure space and into function

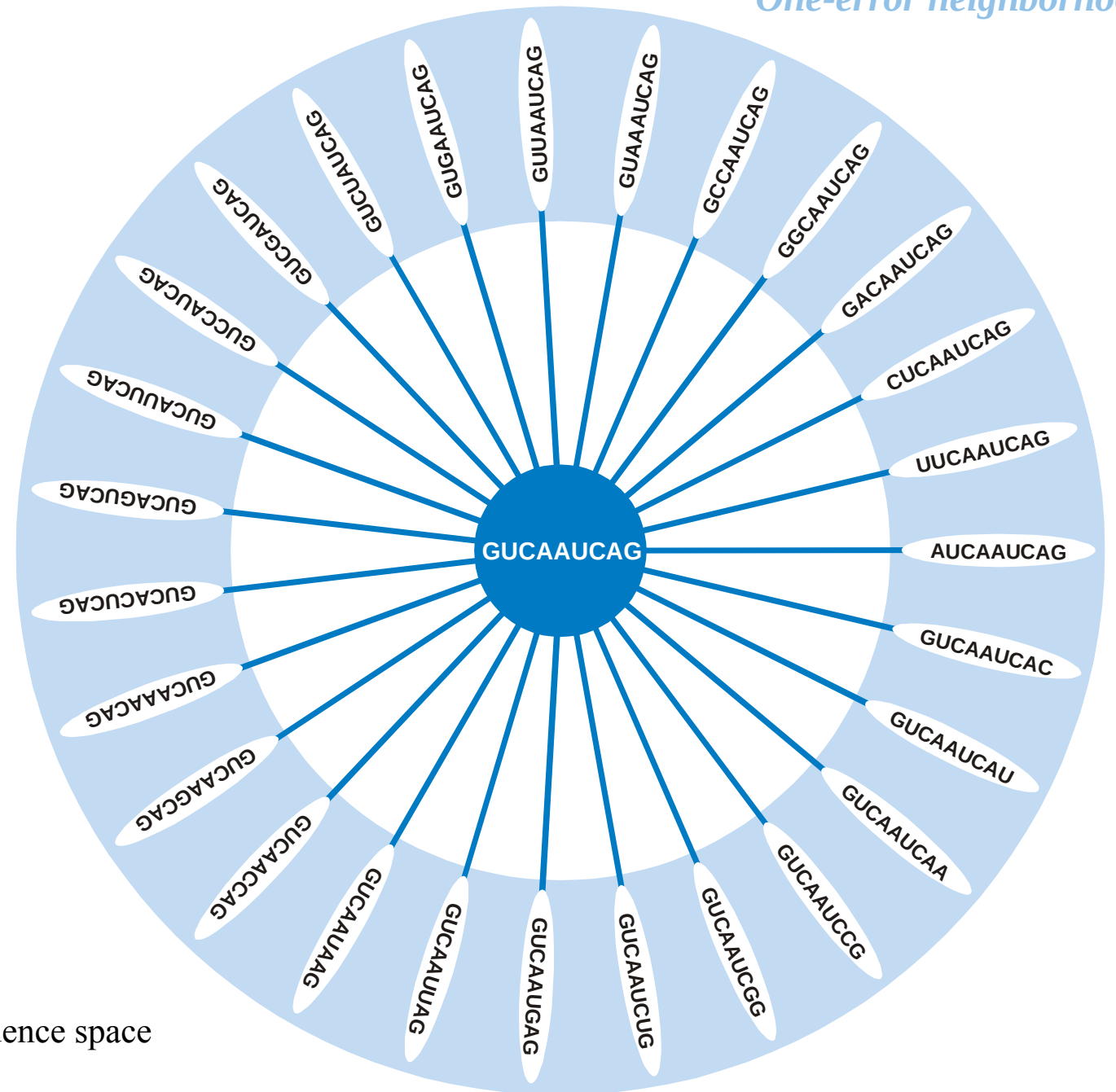

$S_k = \psi(I_.)$
$f_k = f(S_k)$
Function
Sequence space
Structure space
Real numbers


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

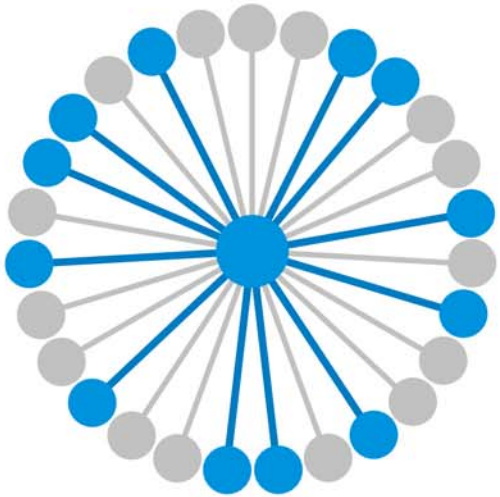


The pre-image of the structure S_k in sequence space is the **neutral network G_k**

One-error neighborhood



The surrounding of
GUCAAUCAG in sequence space



$$G_k = \psi^{-1}(S_k) \doteq \{ I_j \mid \psi(I_j) = S_k \}$$

$$\bar{\lambda}_k = \frac{\sum_{j \in |G_k|} \lambda_j(k)}{|G_k|}$$

$\lambda_j = 12 / 27 = 0.444$

Alphabet size κ :

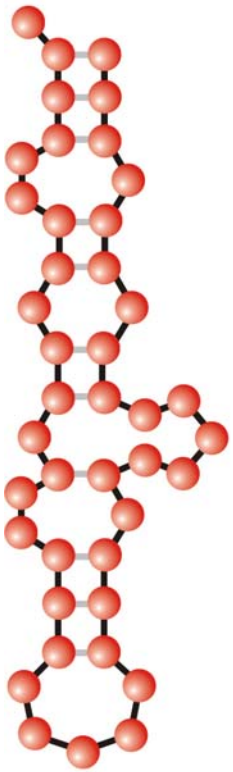
κ	λ_{cr}	
2	0.5	**AU,GC,DU**
3	0.423	**AUG , UGC**
4	0.370	**AUGC**

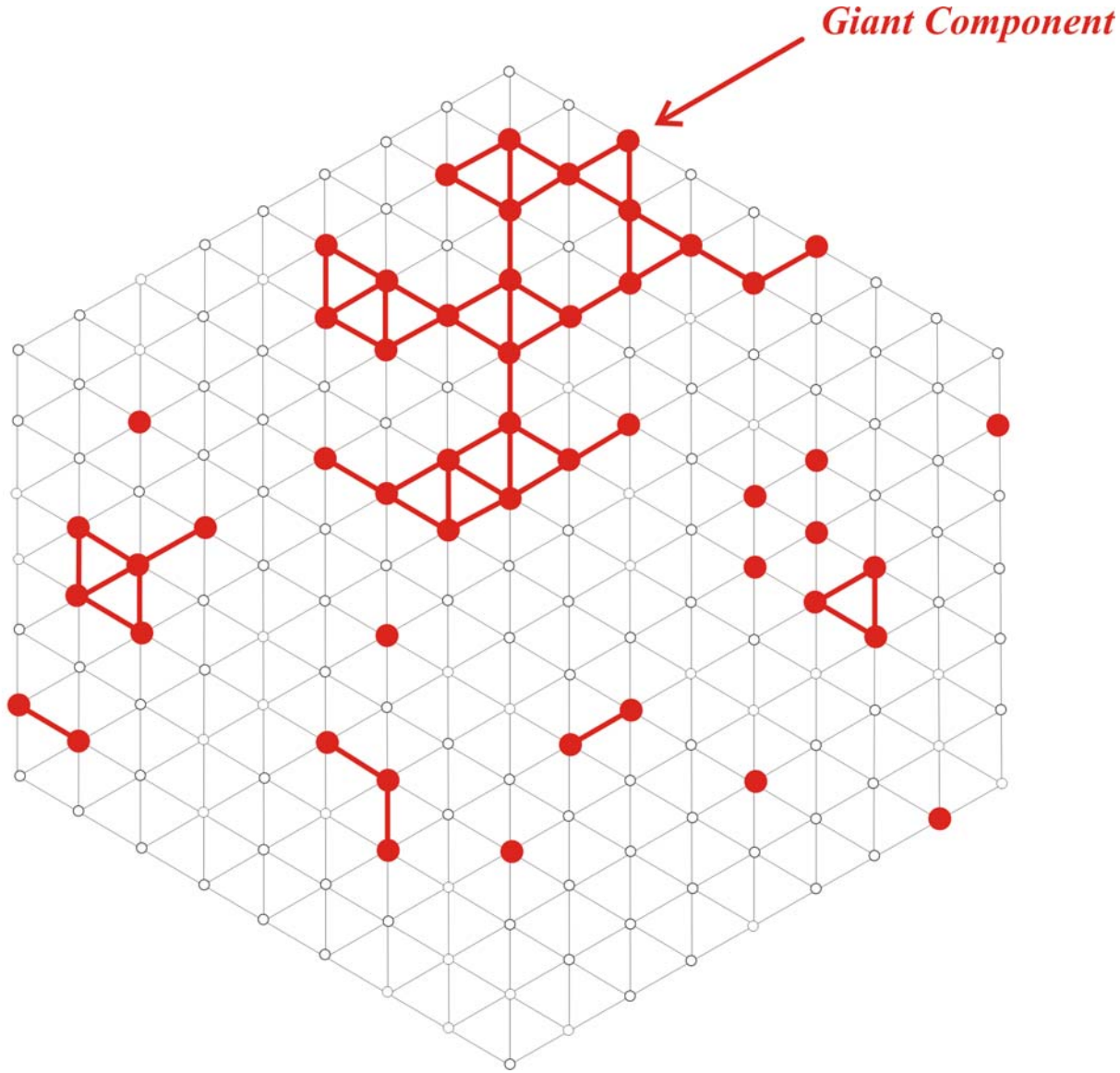
$\bar{\lambda}_k > \lambda_{cr}$ .... network G_k is connected

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

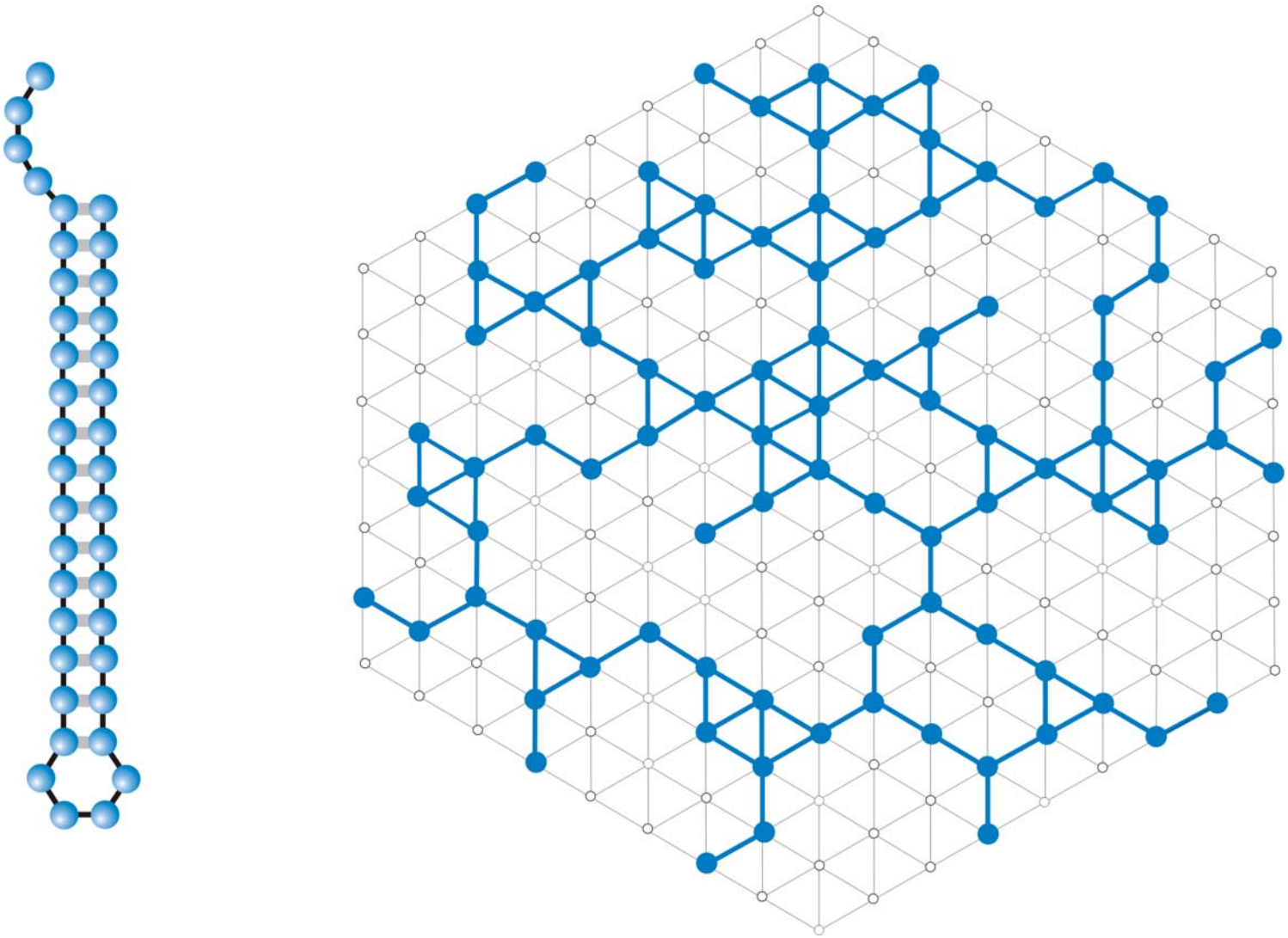
Connectivity threshold: $\lambda_{cr} = 1 - \kappa^{-1/(\kappa-1)}$

Degree of neutrality of neutral networks and the connectivity threshold





A multi-component neutral network formed by a rare structure: $\lambda < \lambda_{cr}$



A connected neutral network formed by a common structure: $\lambda > \lambda_{cr}$

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.

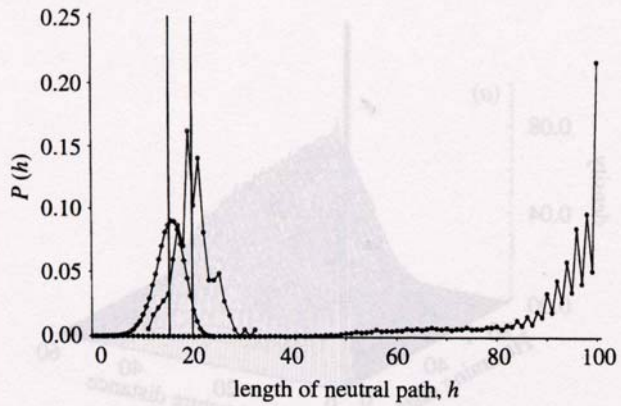


Figure 4. Neutral paths. A neutral path is defined by a series of nearest neighbour sequences that fold into identical structures. Two classes of nearest neighbours are admitted: neighbours of Hamming distance 1, which are obtained by single base exchanges in unpaired stretches of the structure, and 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).

Proc. R. Soc. Lond. B (1994) **255**, 279–284
Printed in Great Britain

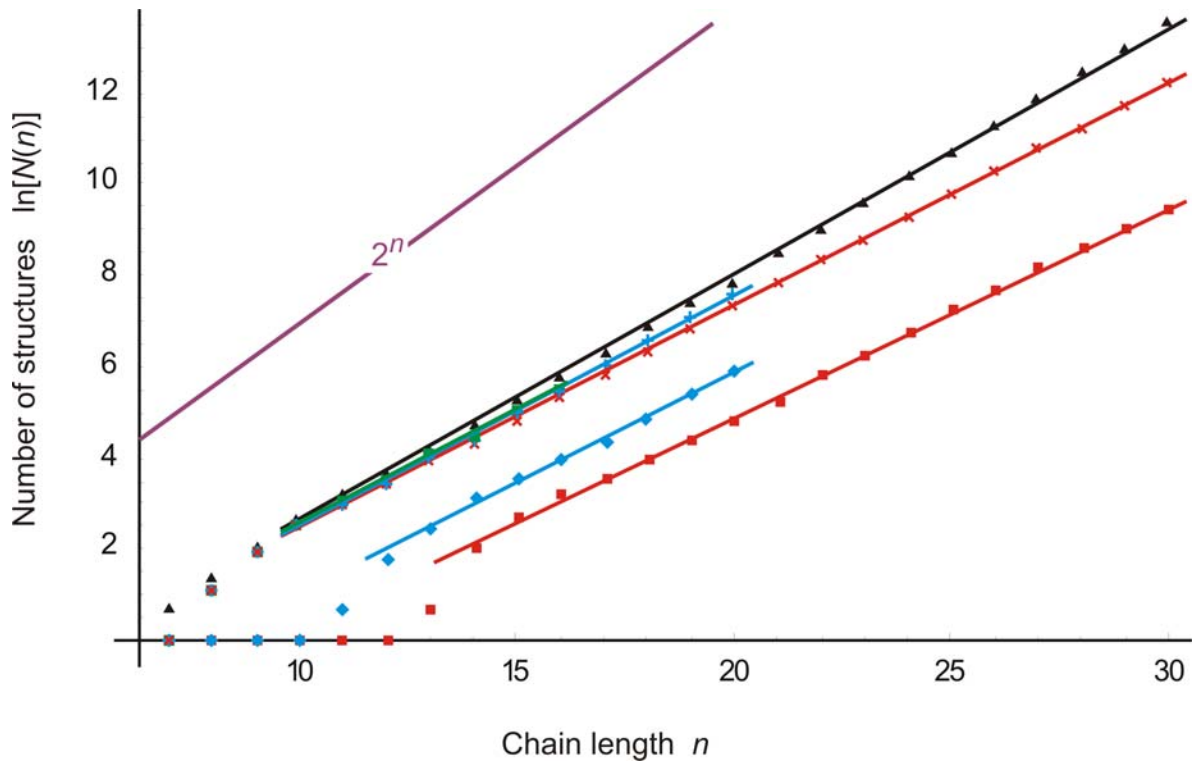
Reference for postulation and ***in silico*** verification of ***neutral networks***

Properties of RNA sequence to secondary structure mapping

1. More sequences than structures

Properties of RNA sequence to secondary structure mapping

1. More sequences than structures

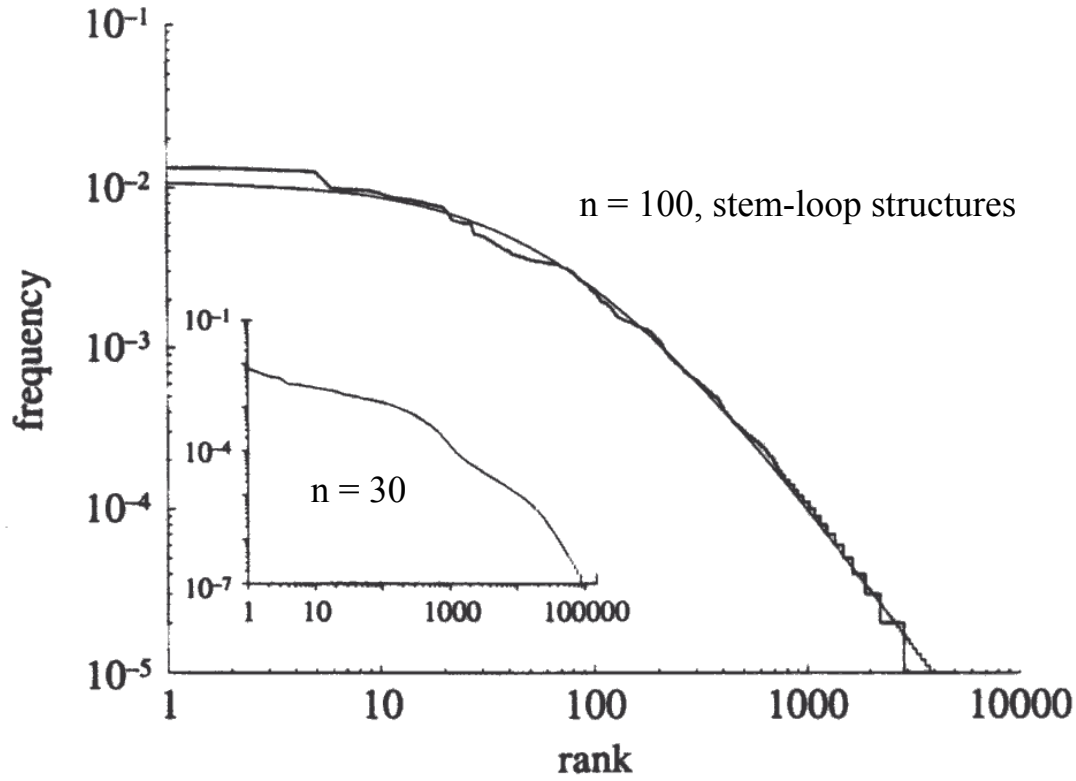


Properties of RNA sequence to secondary structure mapping

1. More sequences than structures
2. Few common versus many rare structures

Properties of RNA sequence to secondary structure mapping

1. More sequences than structures
2. Few common versus many rare structures



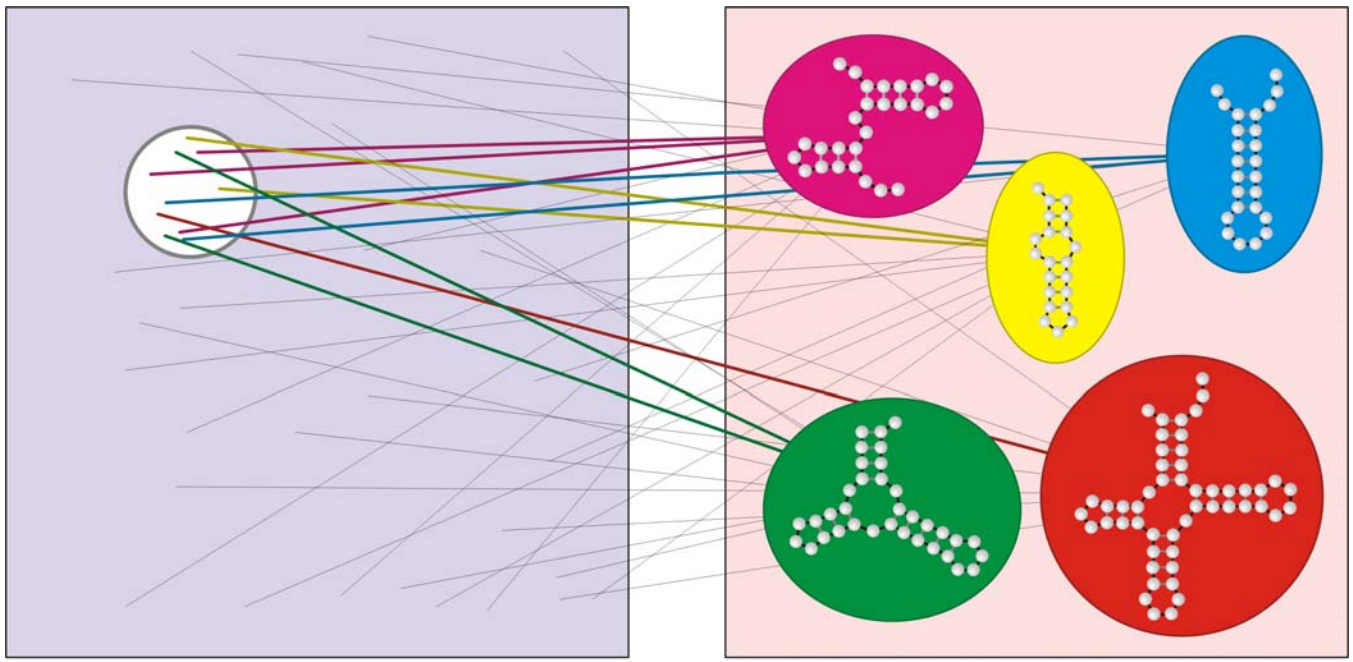
RNA secondary structures and Zipf's law

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

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



Properties of RNA sequence to secondary structure mapping

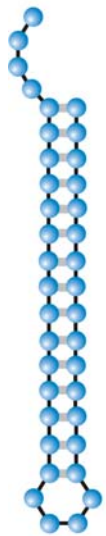
1. More sequences than structures
2. Few common versus many rare structures
3. Shape space covering of common structures
4. Neutral networks of common structures are connected

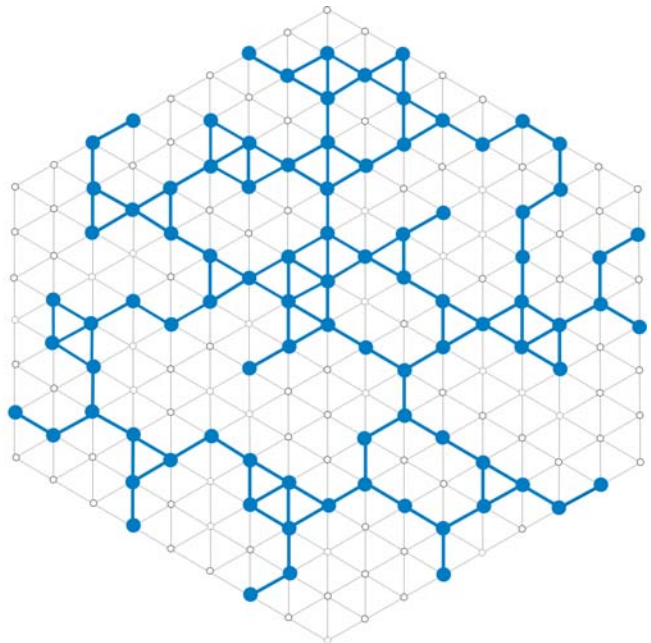
Properties of RNA sequence to secondary structure mapping

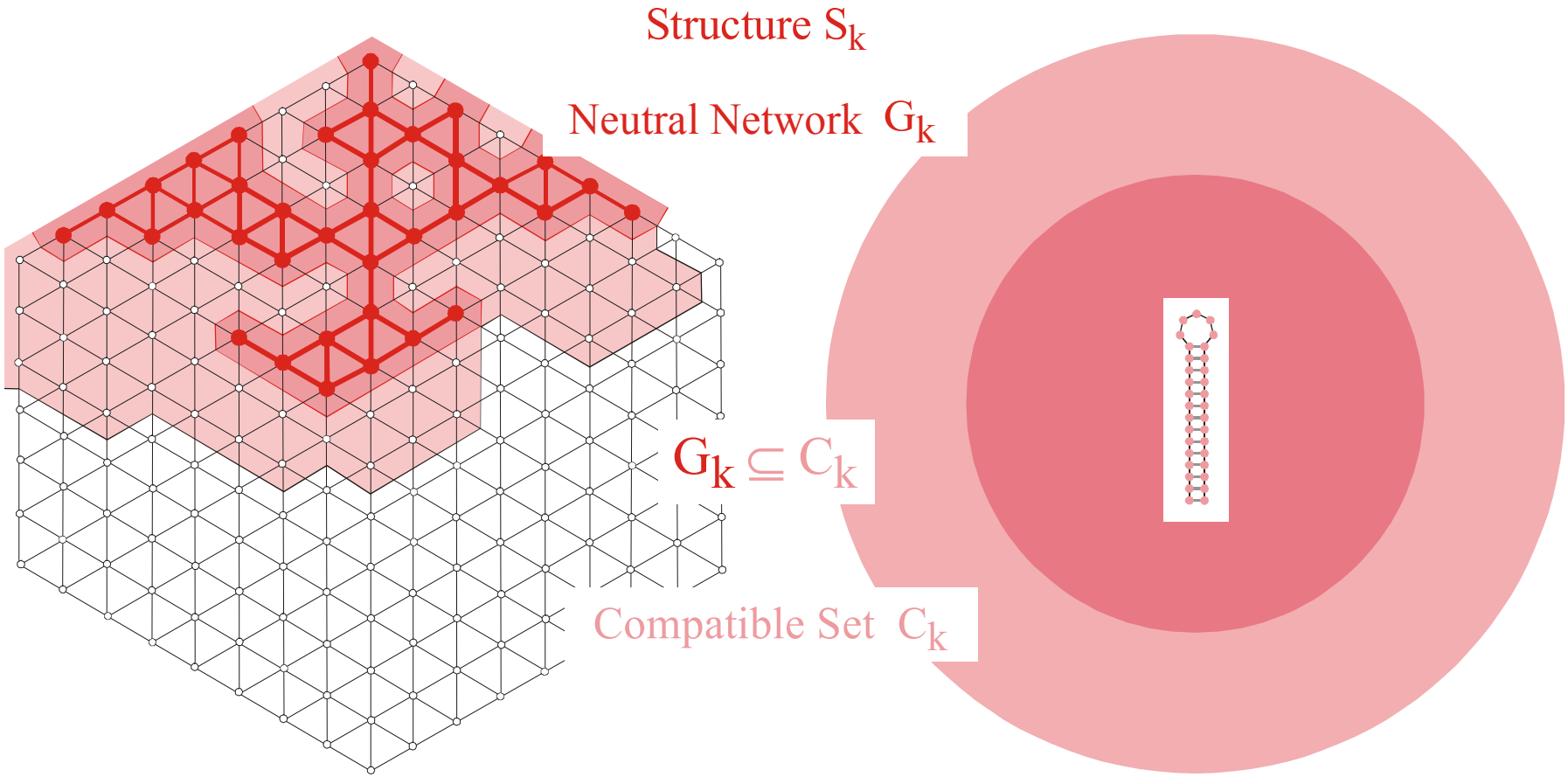
1. More sequences than structures
2. Few common versus many rare structures
3. Shape space covering of common structures
4. Neutral networks of common structures are connected

Alphabet size κ :

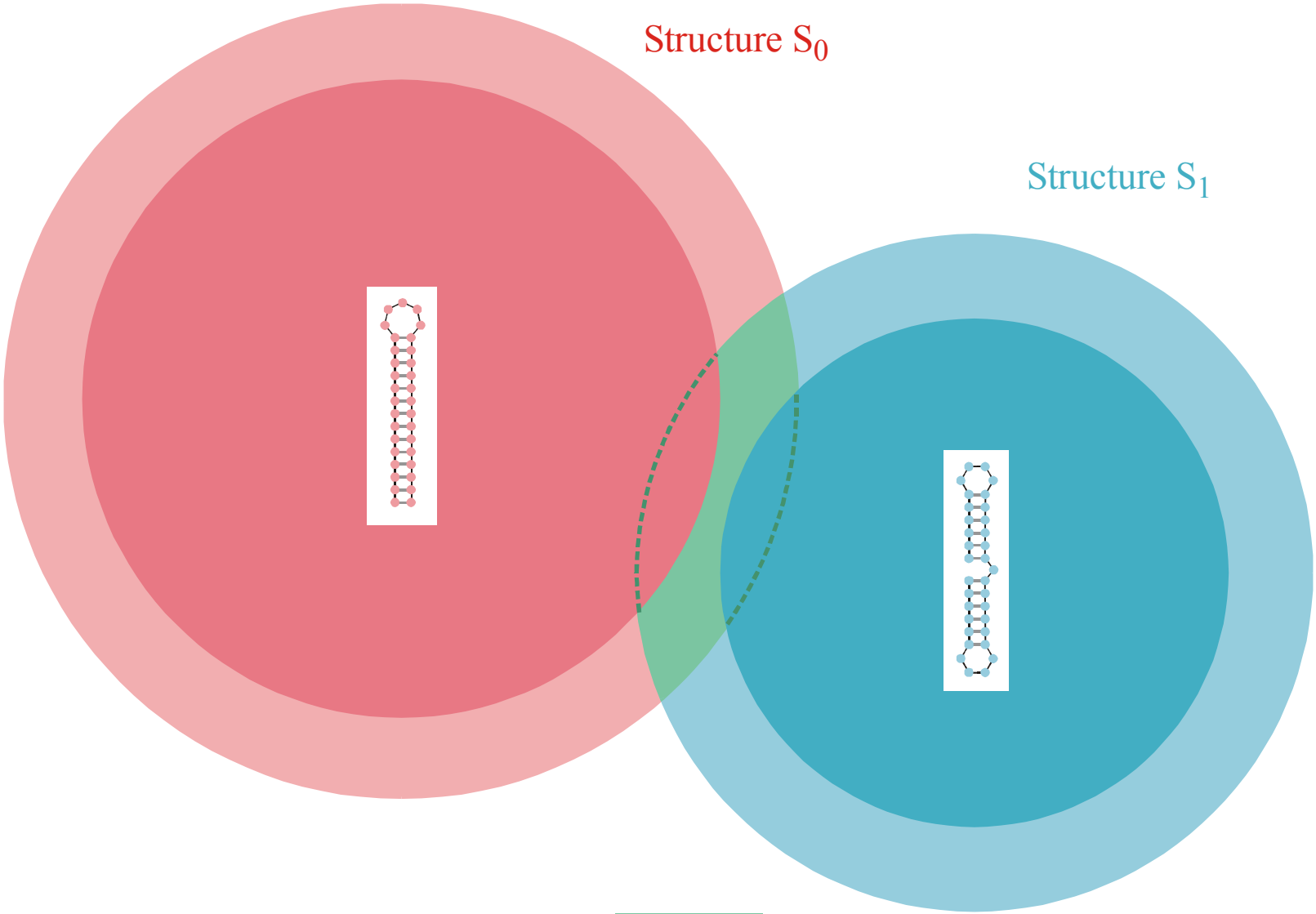
κ	λ_{cr}	
2	0.5	**AU,GC,DU**
3	0.423	**AUG , UGC**
4	0.370	**AUGC**







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: $C_0 \cap C_1 \notin \varnothing$

Bulletin of Mathematical Biology, Vol. 59, No. 2, pp. 339–397, 1997
Elsevier Science Inc.

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,*

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

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 $\gamma(s, s') \cong D_m$ 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**

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*).

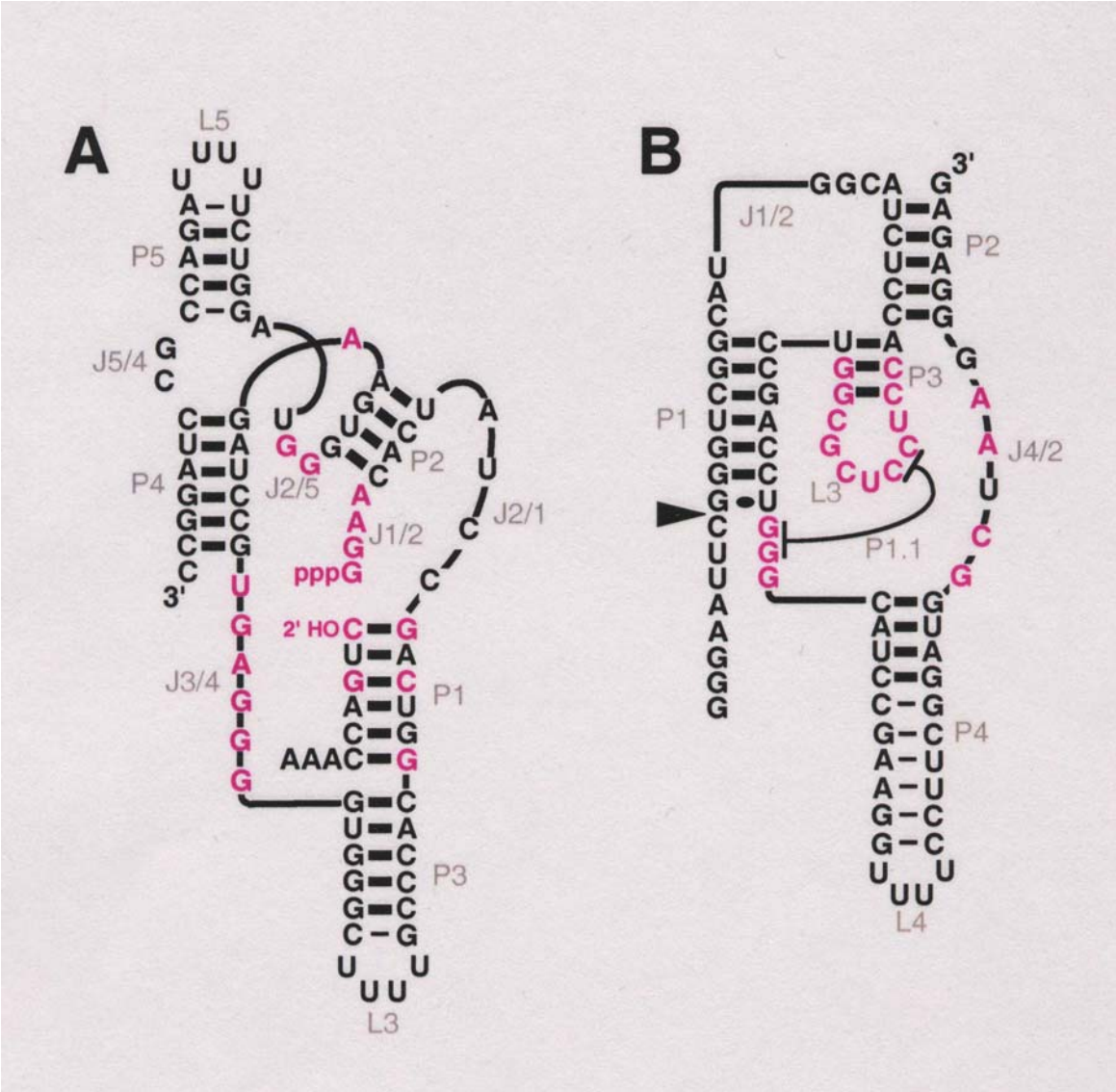
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect.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

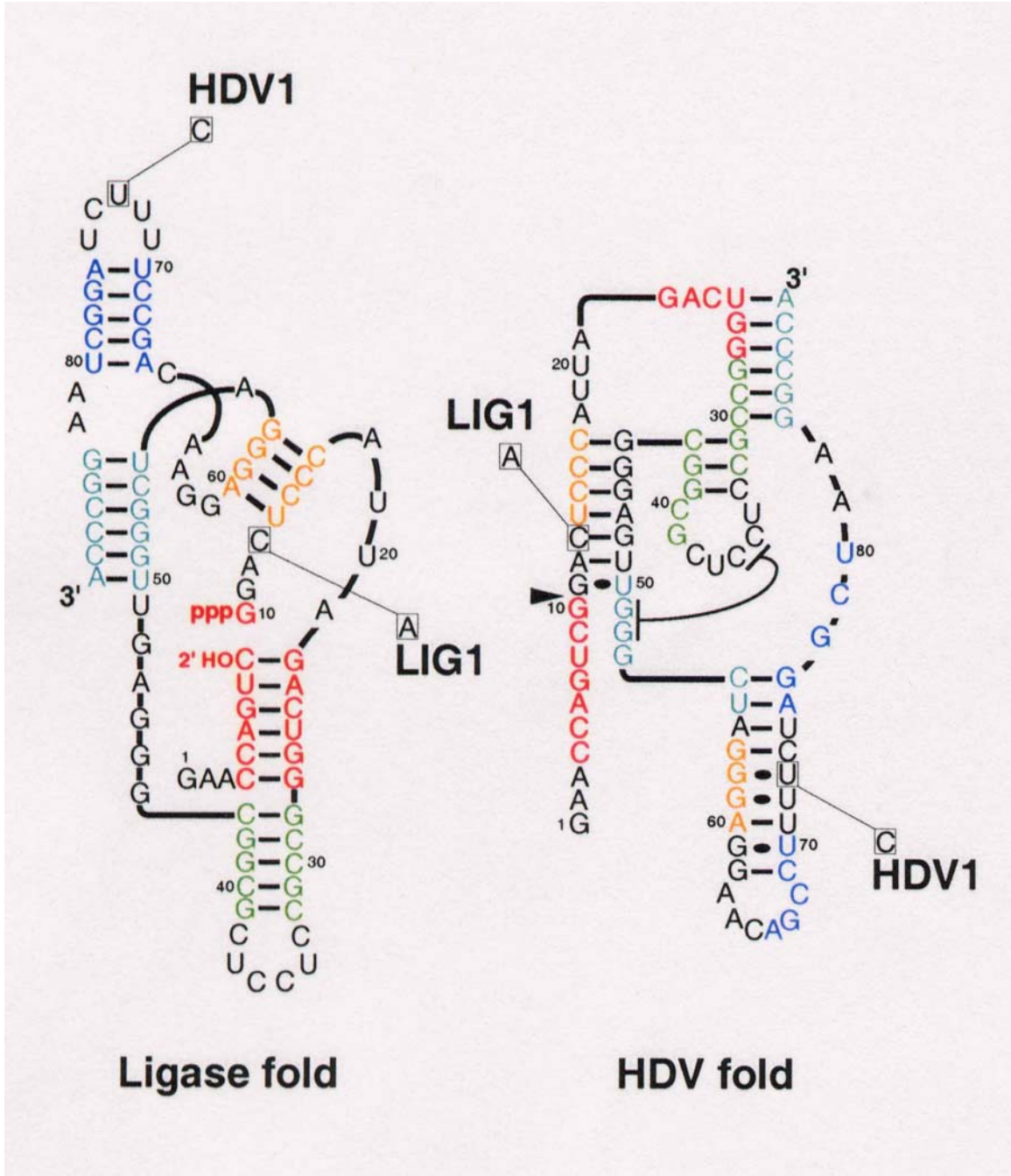
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

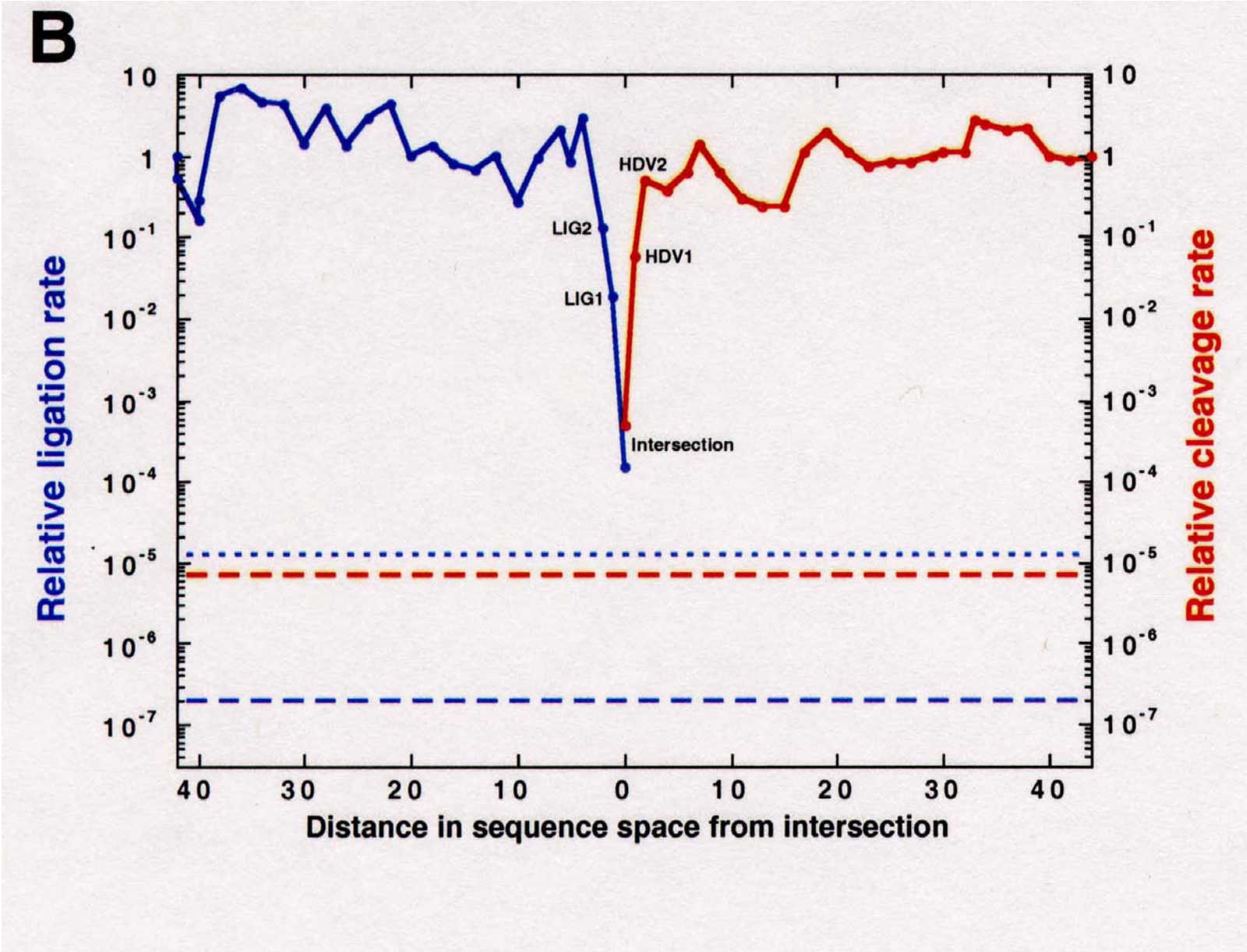


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

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

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



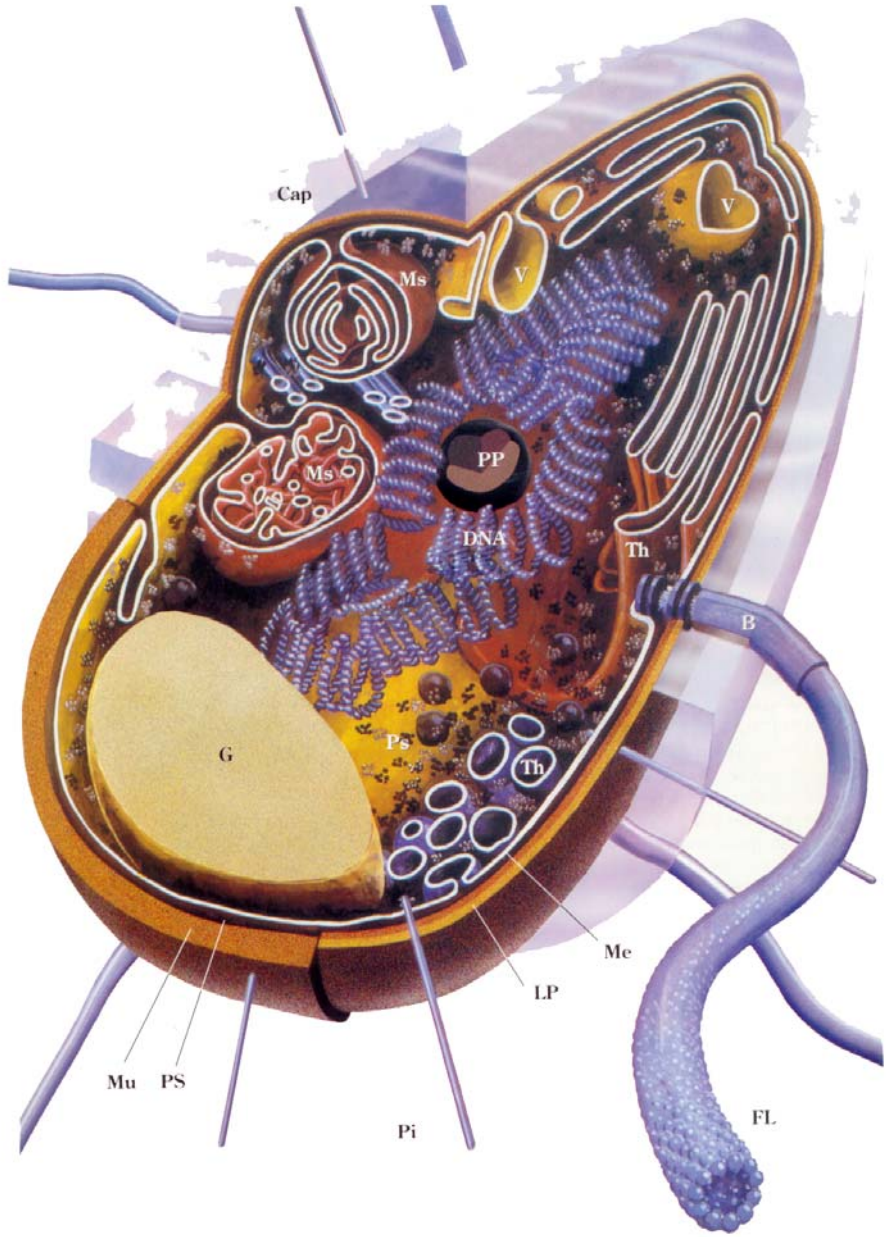


Two neutral walks through sequence space with conservation of structure and catalytic activity

Cell biology

Regulation of cell cycle, metabolic networks, reaction kinetics, homeostasis, ...

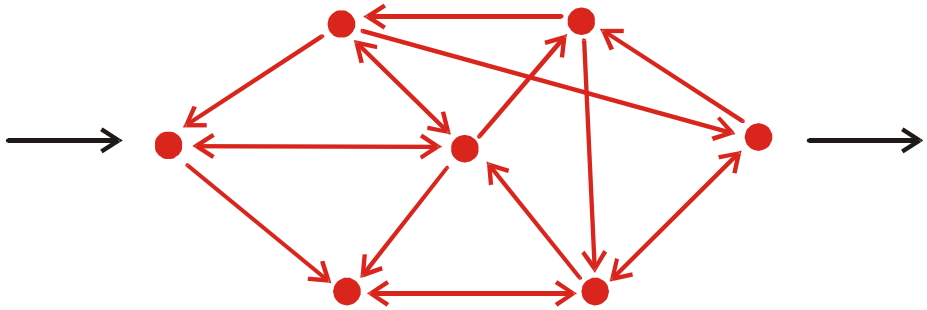
The bacterial cell as an example for the simplest form of autonomous life



The human body:

10^{14} cells = 10^{13} eukaryotic cells +
$\approx 9\times10^{13}$ bacterial (prokaryotic) cells,
and $\approx$ 200 eukaryotic cell types

Linear chain



Network

Processing of information in cascades and networks

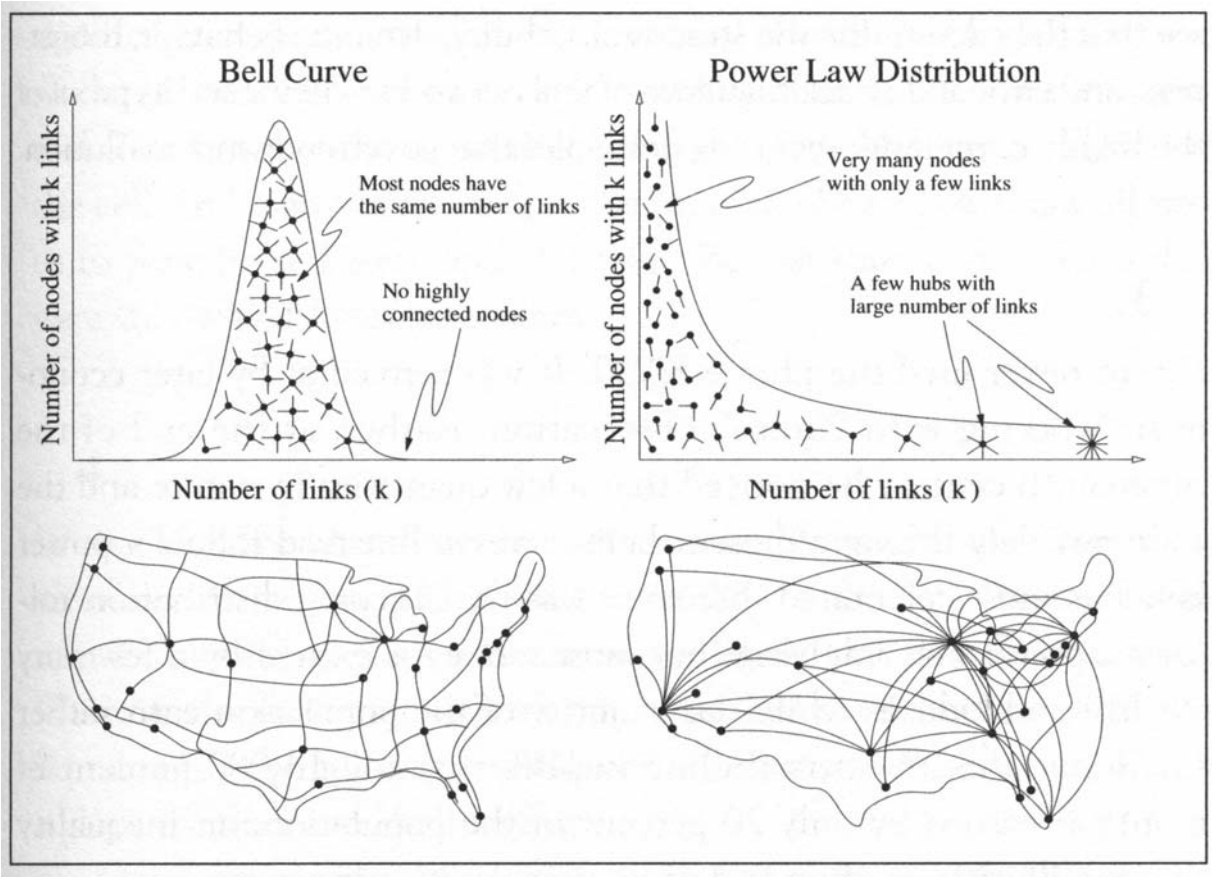
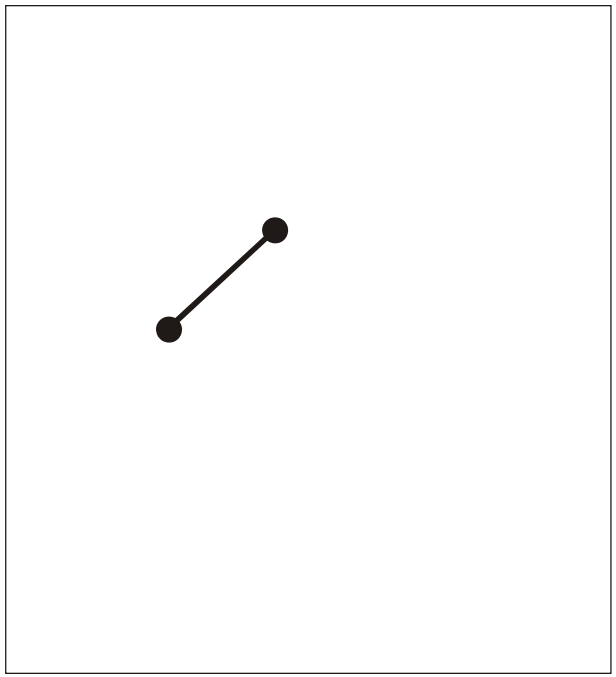
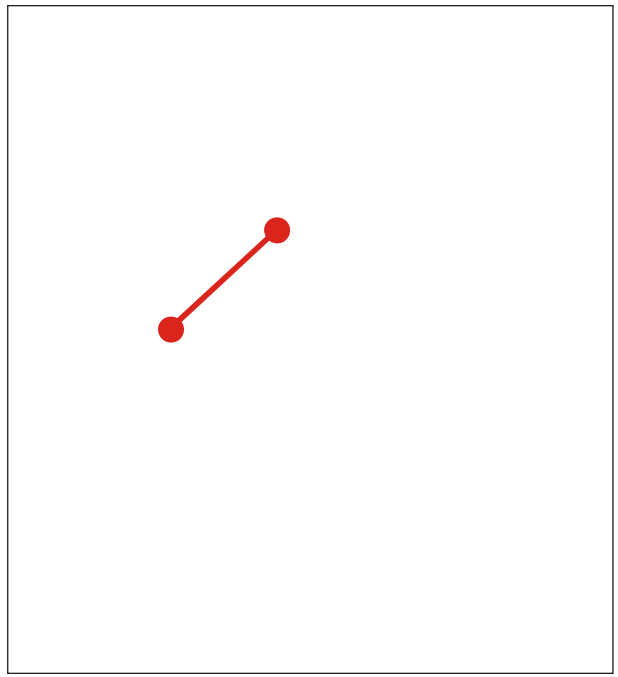


Figure 6.1 Random and Scale-Free Networks. *The degree distribution of a random network follows a bell curve, telling us that most nodes have the same number of links, and nodes with a very large number of links don't exist (top left). Thus a random network is similar to a national highway network, in which the nodes are the cities, and the links are the major highways connecting them. Indeed, most cities are served by roughly the same number of highways (bottom left). In contrast, the power law degree distribution of a scale-free network predicts that most nodes have only a few links, held together by a few highly connected hubs (top right). Visually this is very similar to the air traffic system, in which a large number of small airports are connected to each other via a few major hubs (bottom right).*

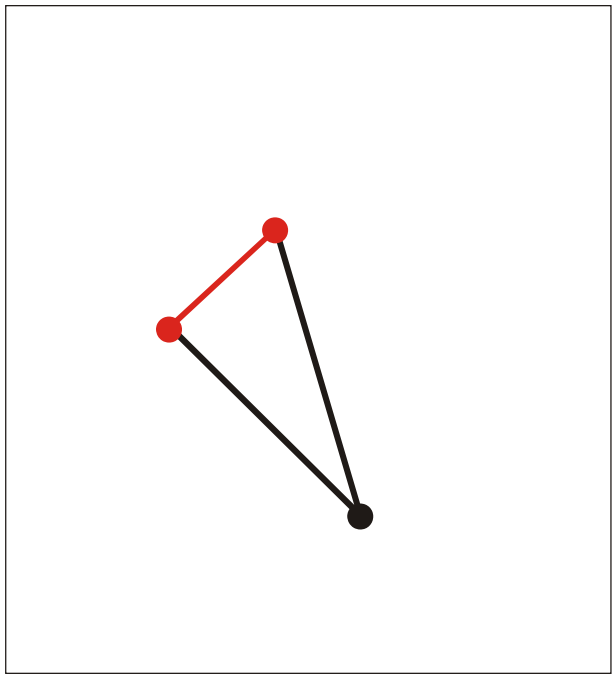
Albert-László Barabási,
Linked – The New Science of Networks
Perseus Publ., Cambridge, MA, 2002



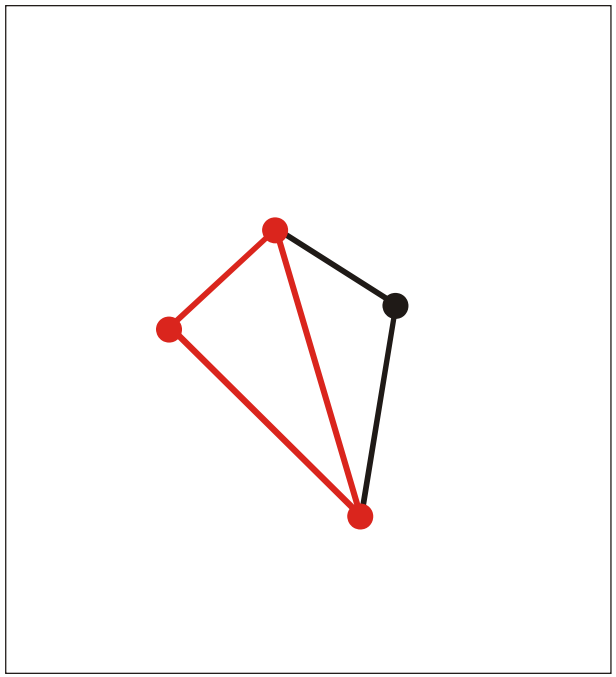
Formation of a scale-free network through evolutionary point by point expansion: Step 000



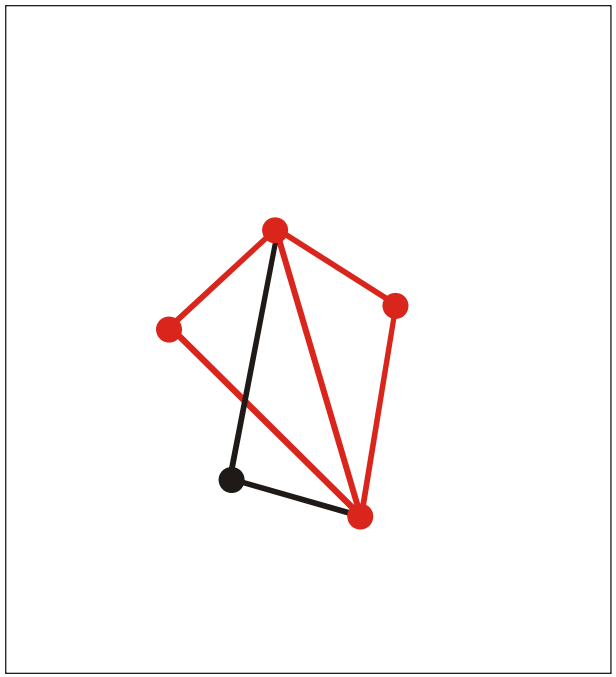
Formation of a scale-free network through evolutionary point by point expansion: Step 001



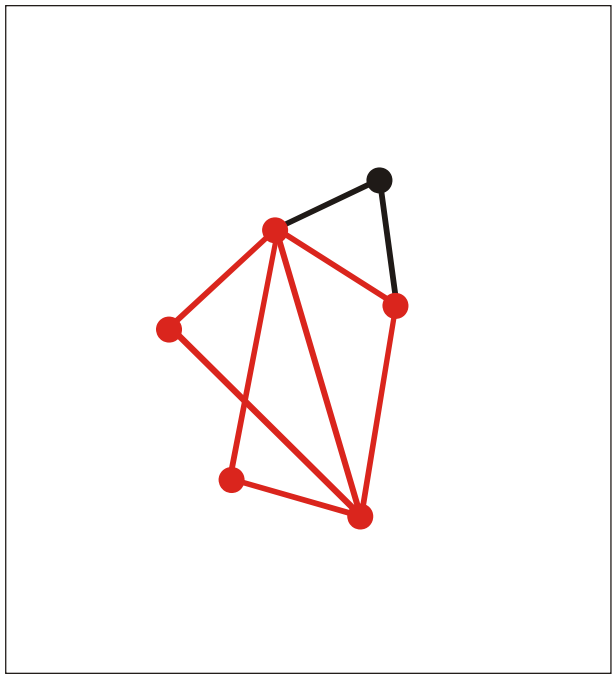
Formation of a scale-free network through evolutionary point by point expansion: Step 002



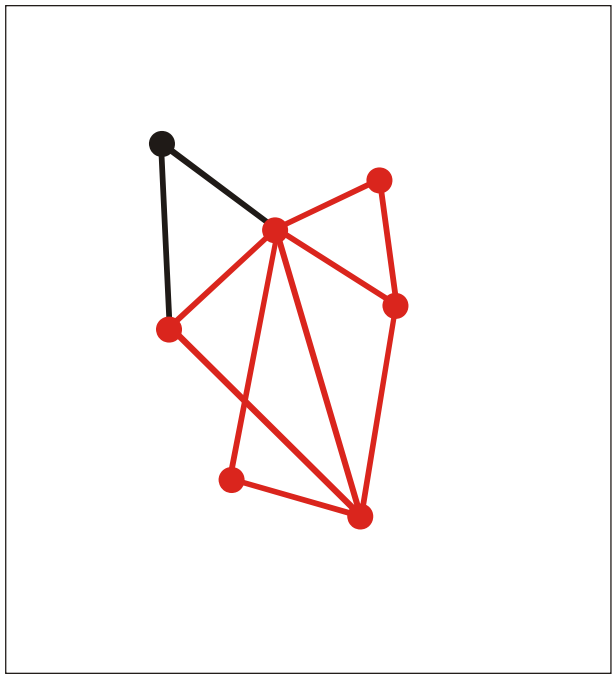
Formation of a scale-free network through evolutionary point by point expansion: Step 003



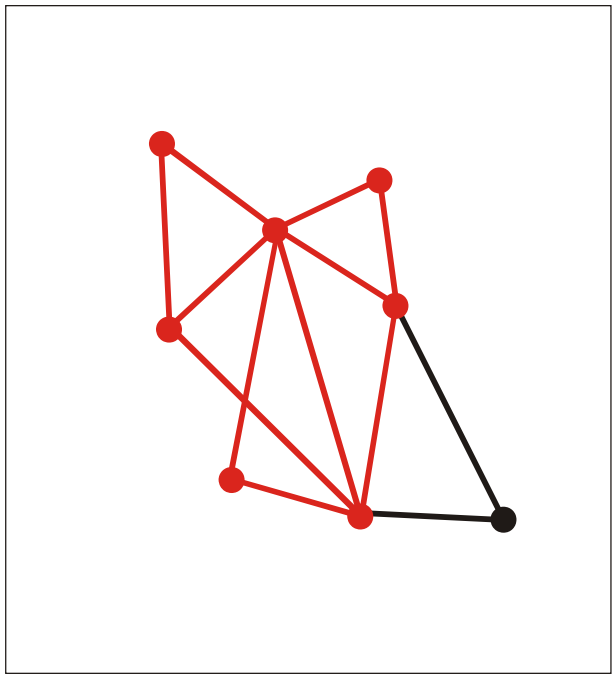
Formation of a scale-free network through evolutionary point by point expansion: Step 004



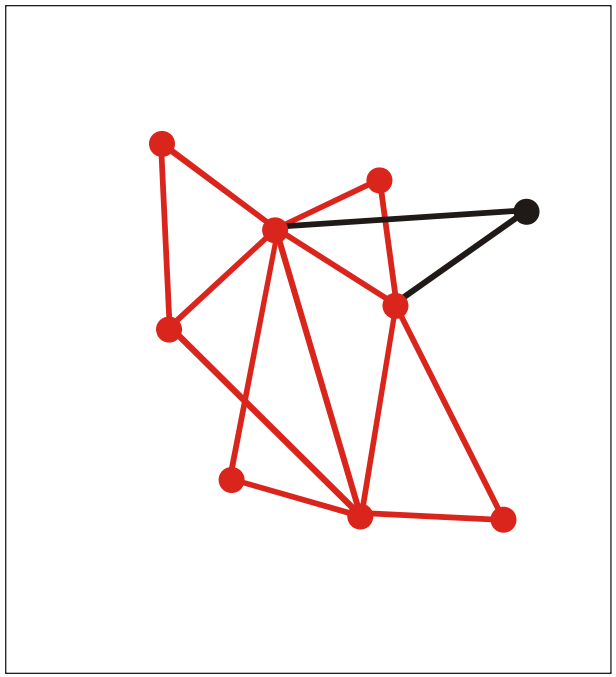
Formation of a scale-free network through evolutionary point by point expansion: Step 005



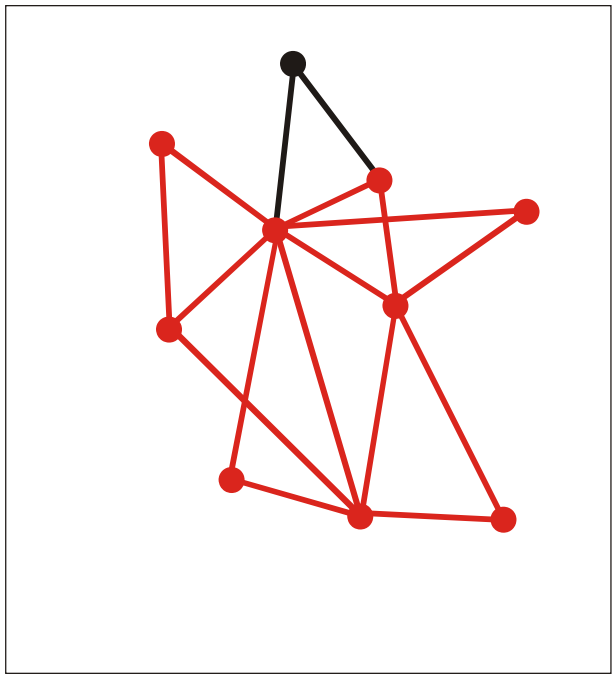
Formation of a scale-free network through evolutionary point by point expansion: Step 006



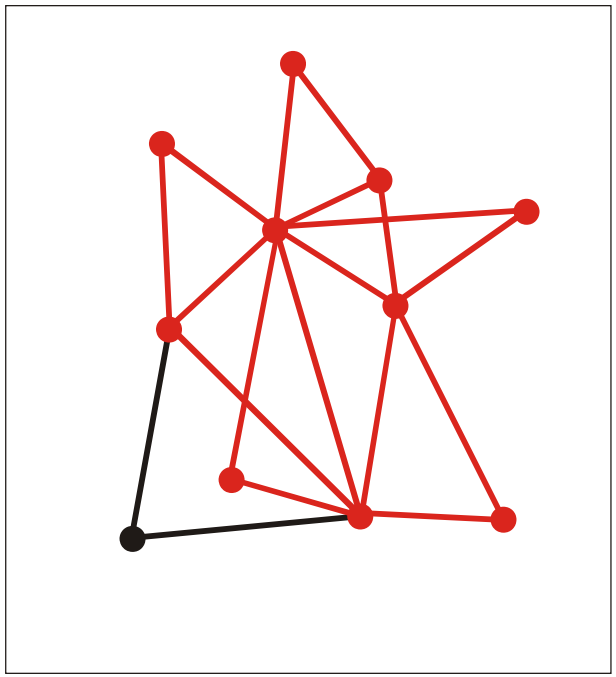
Formation of a scale-free network through evolutionary point by point expansion: Step 007



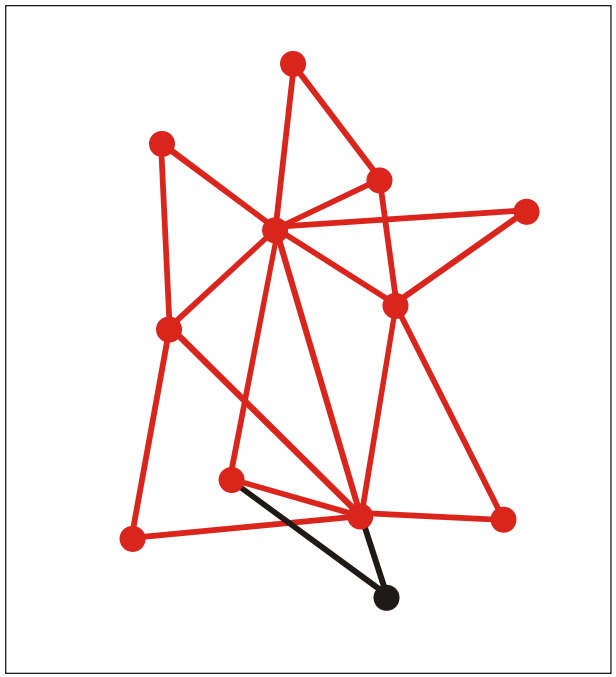
Formation of a scale-free network through evolutionary point by point expansion: Step 008



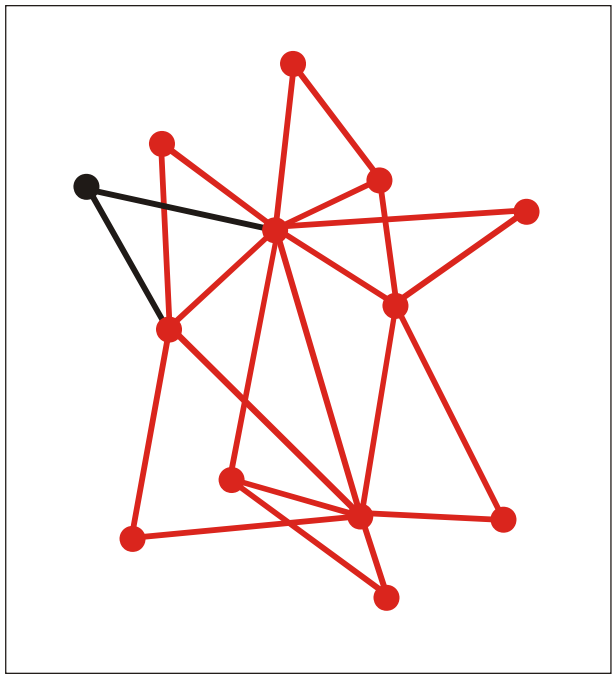
Formation of a scale-free network through evolutionary point by point expansion: Step 009



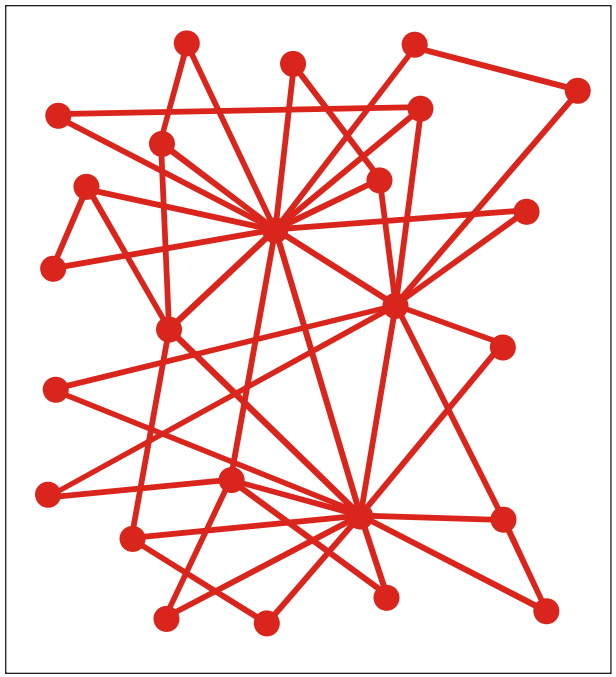
Formation of a scale-free network through evolutionary point by point expansion: Step 010



Formation of a scale-free network through evolutionary point by point expansion: Step 011

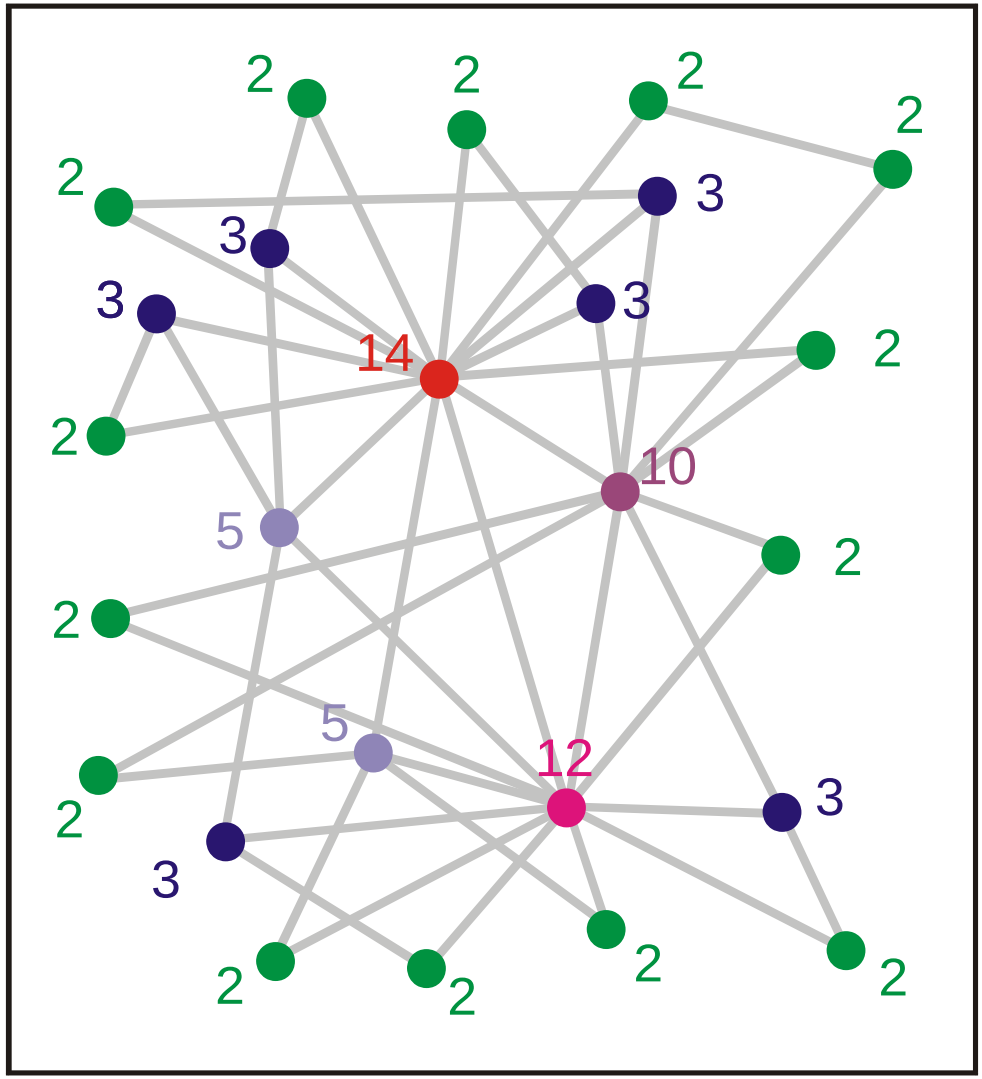


Formation of a scale-free network through evolutionary point by point expansion: Step 012

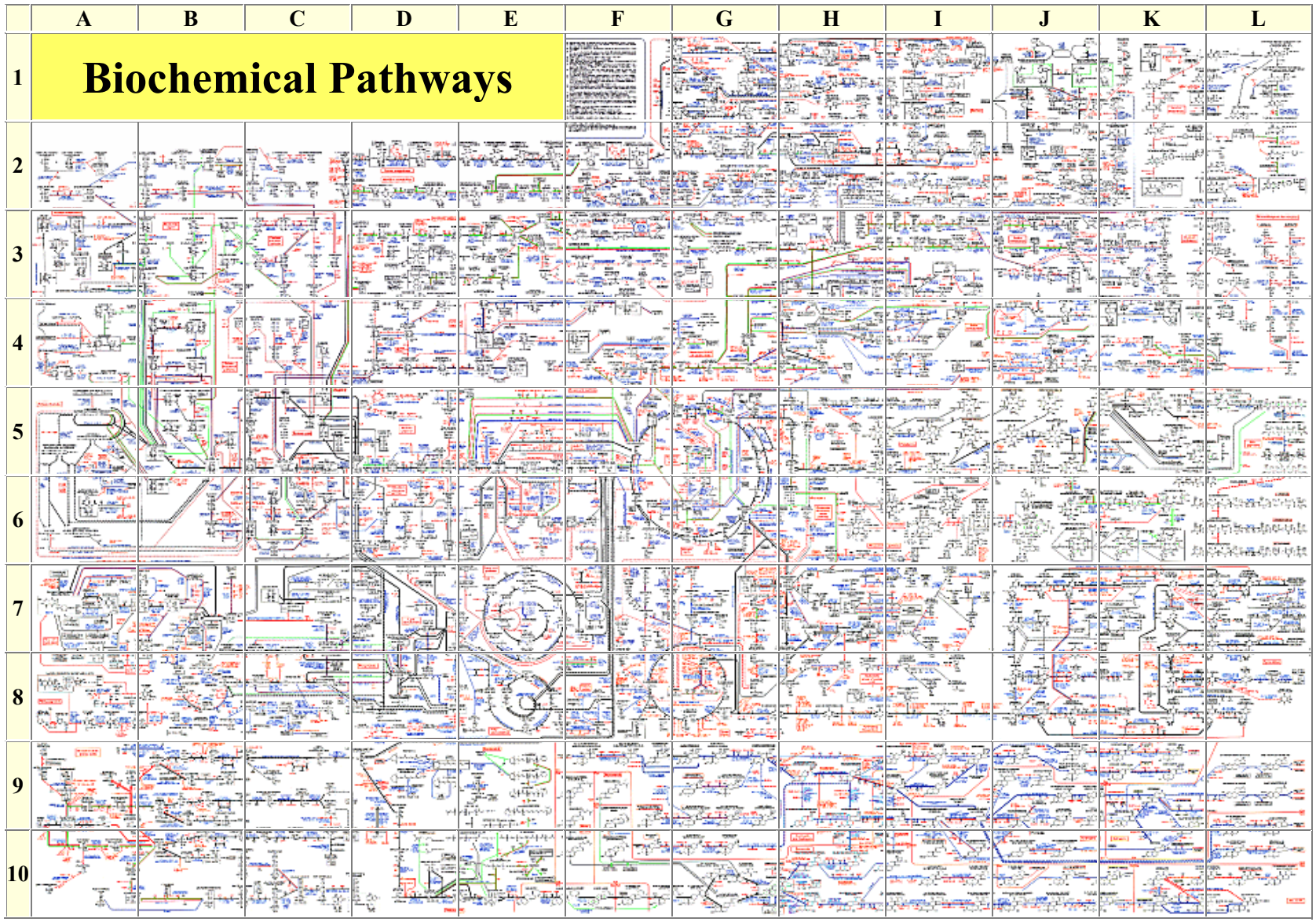


Formation of a scale-free network through evolutionary point by point expansion: Step 024

links	# nodes
2	14
3	6
5	2
10	1
12	1
14	1

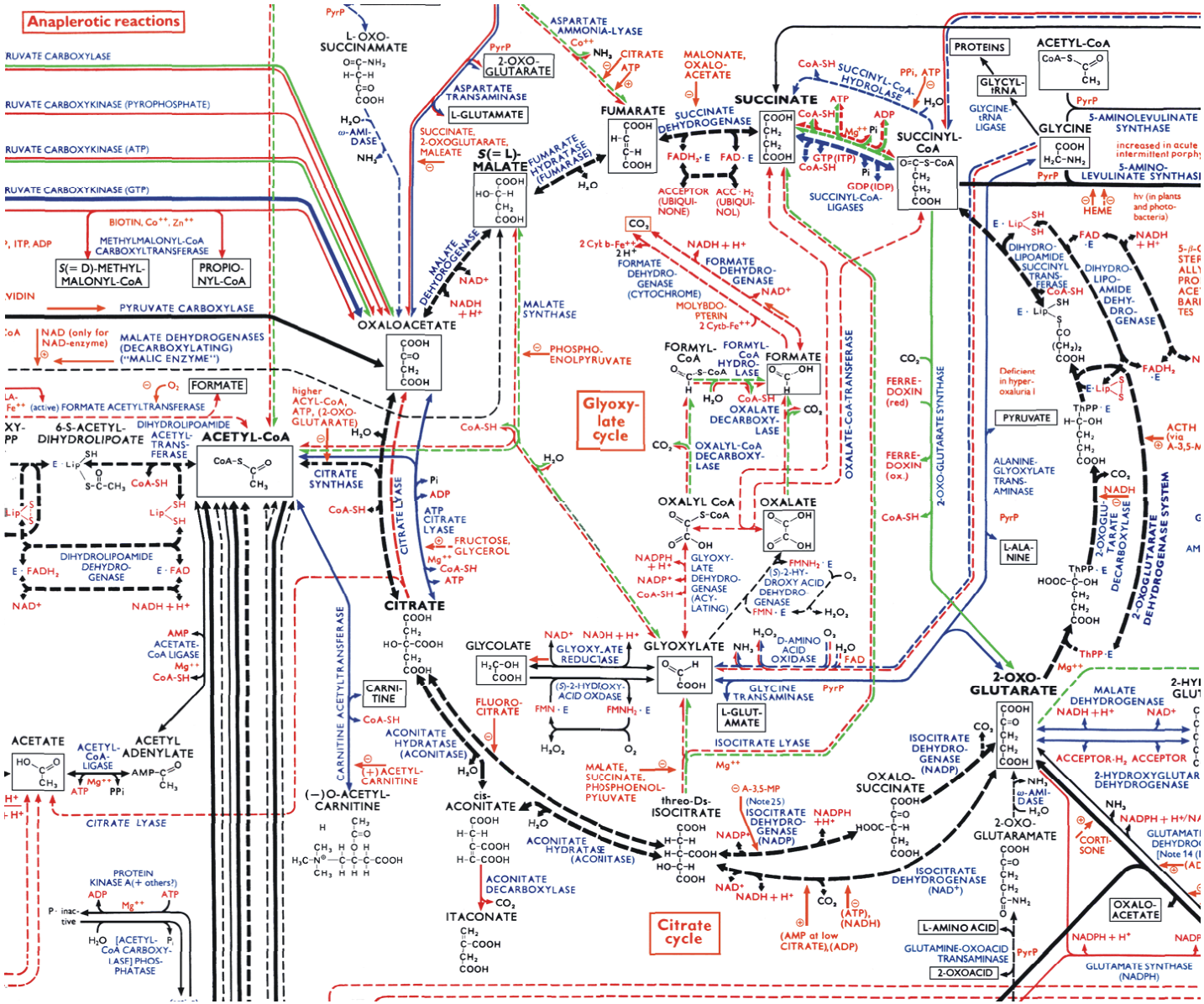


Analysis of nodes and links in a step by step evolved network



The reaction network of cellular metabolism published by Boehringer-Ingelheim.

The citric acid or Krebs cycle (enlarged from previous slide).



Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

Parameter set

$k_j(\mathrm{T},\mathrm{p},\mathrm{pH},\mathrm{I},\ldots);\ \mathrm{j}=1,2,\ldots,\mathrm{m}$

General conditions: T, p, pH, I, ...

Initial conditions: $x(0)$

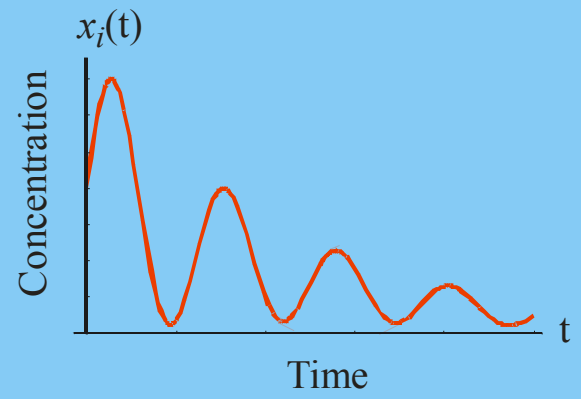
Boundary conditions:

boundary ... $\boldsymbol{S}$, normal unit vector ... $\hat{\boldsymbol{u}}$

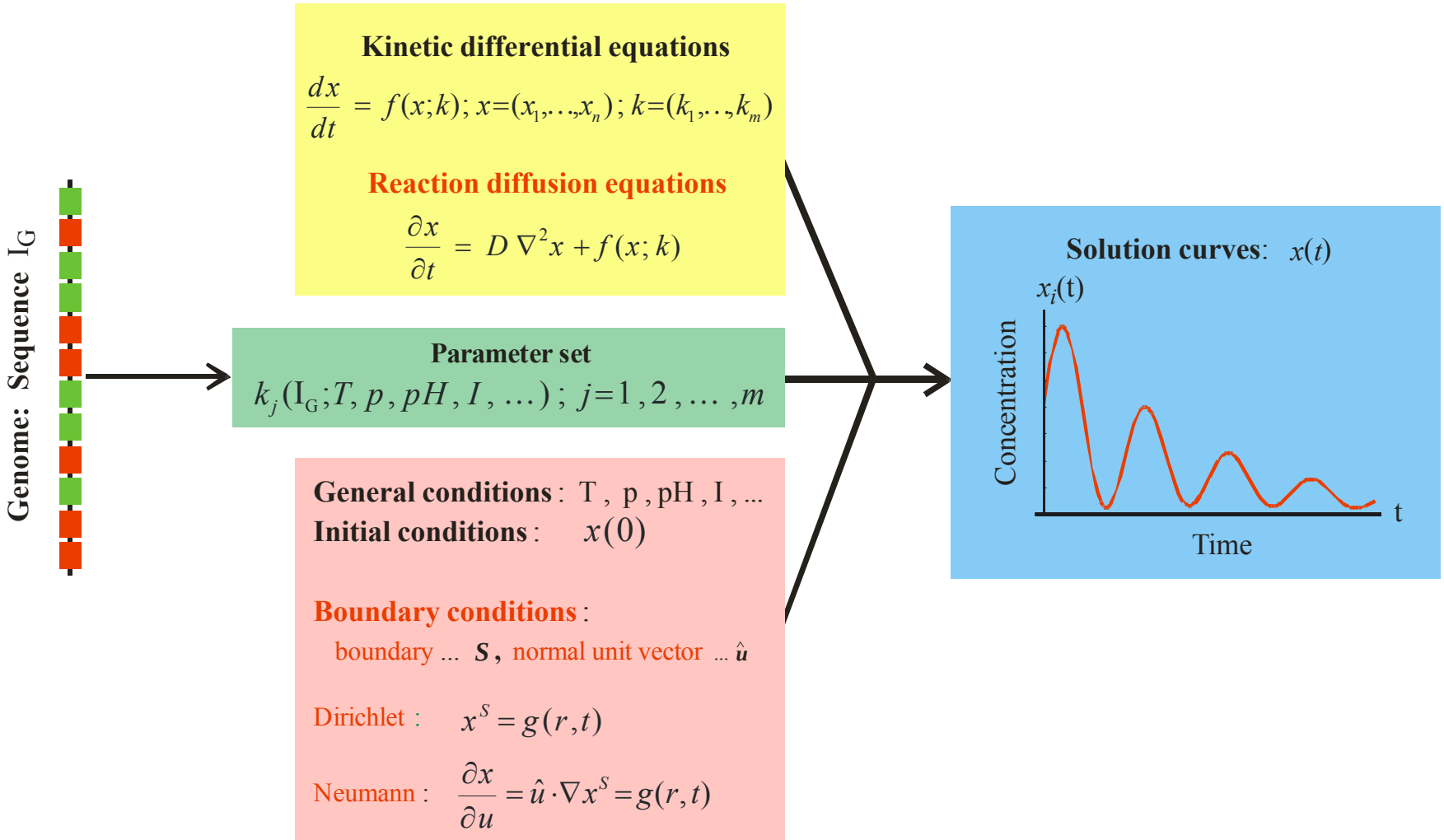
Dirichlet : $x^S = g(r,t)$

Neumann : $\frac{\partial x}{\partial u} = \hat{u}\cdot\nabla x^S = g(r,t)$

Solution curves: $x(t)$



The forward problem of chemical reaction kinetics (Level I)



The forward problem of cellular reaction kinetics (Level I)

Parameter set

$$k_j(I_G; T, p, pH, I, \ldots);\ j = 1, 2, \ldots, m$$

Kinetic differential equations

$$\frac{dx}{dt} = f(x;k);\ x=(x_1,\ldots,x_n);\ k=(k_1,\ldots,k_m)$$

Reaction diffusion equations

$$\frac{\partial x}{\partial t} = D\,\nabla^2 x + f(x;k)$$

General conditions : T , p , pH , I , ...

Initial conditions : $x(0)$

Boundary conditions :

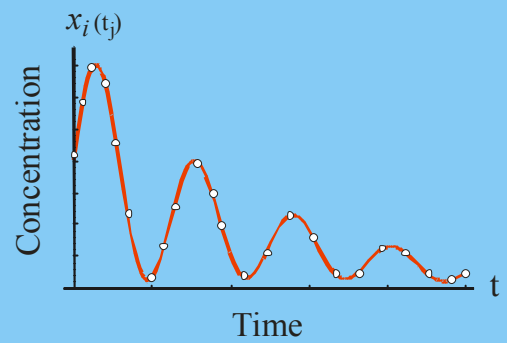
boundary ... S, normal unit vector... $\hat{u}$

Dirichlet : $x^S = g(r,t)$

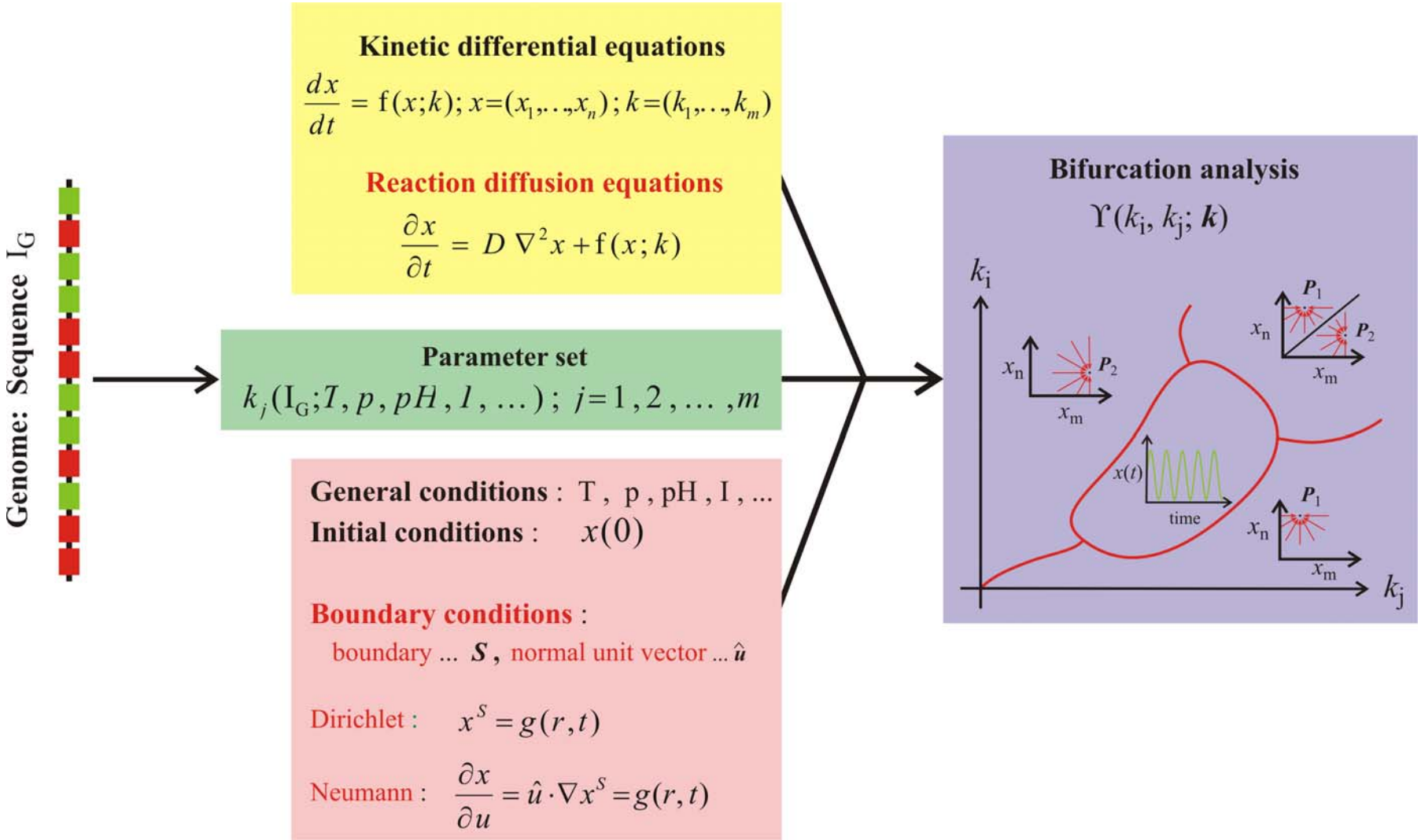
Neumann : $\frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r,t)$

Data from measurements

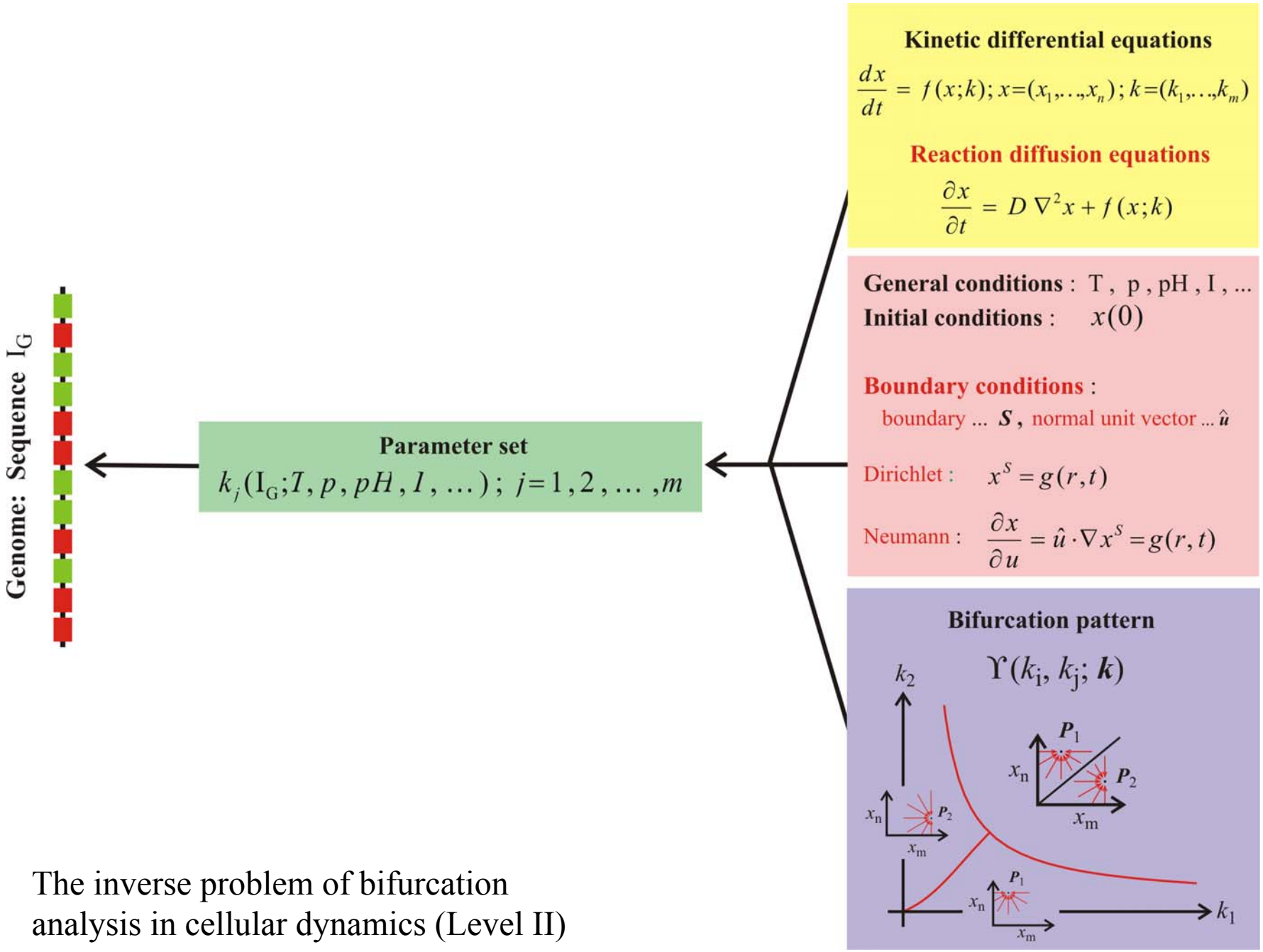
$x(t_j);\ j = 1, 2, \ldots, N$



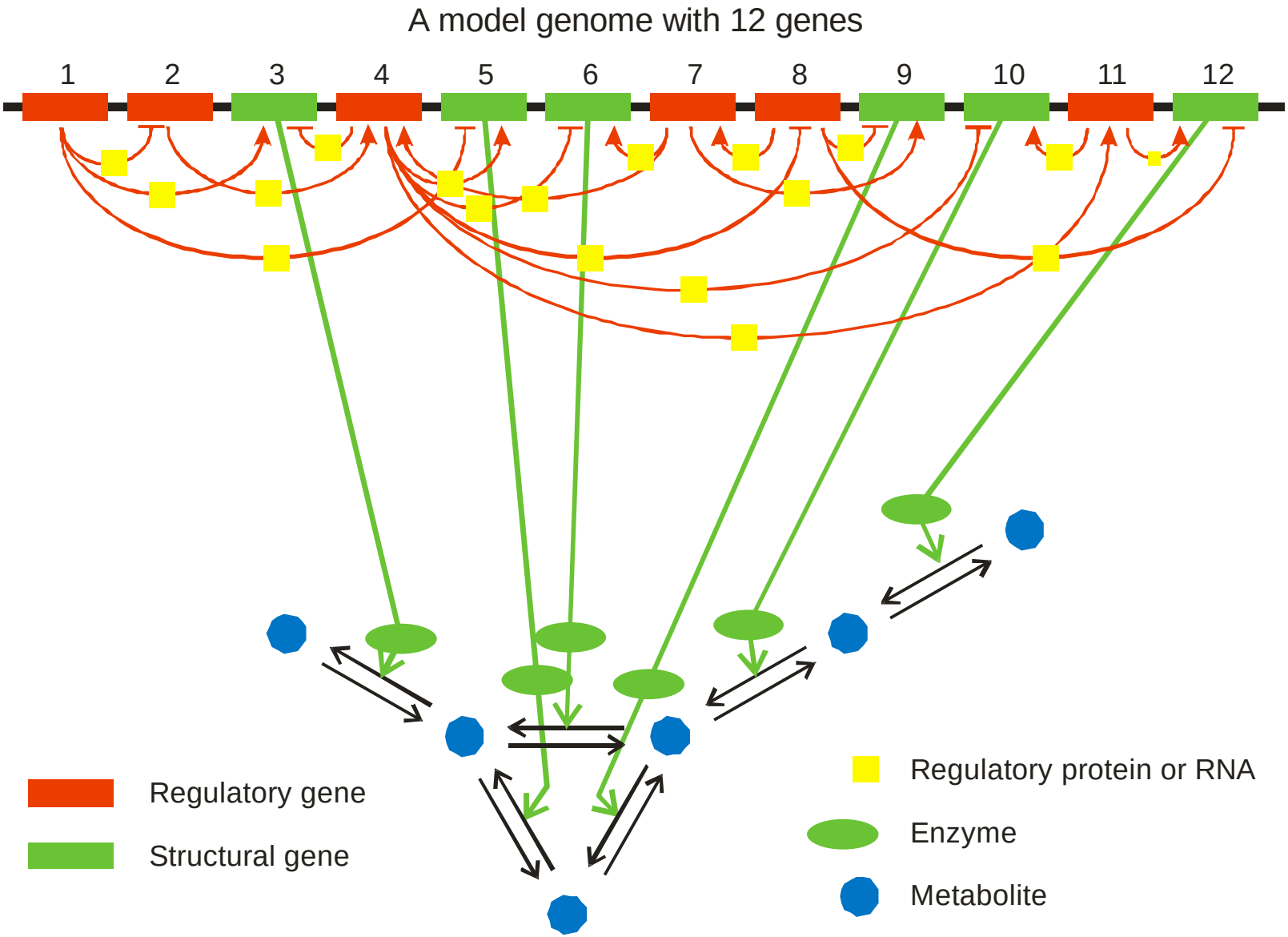
The inverse problem of cellular reaction kinetics (Level I)



The forward problem of bifurcation analysis in cellular dynamics (Level II)



The inverse problem of bifurcation analysis in cellular dynamics (Level II)



Sketch of a genetic and metabolic network

$$A + B \rightarrow X$$
$$2\,X \rightarrow Y$$
$$Y + X \rightarrow D$$

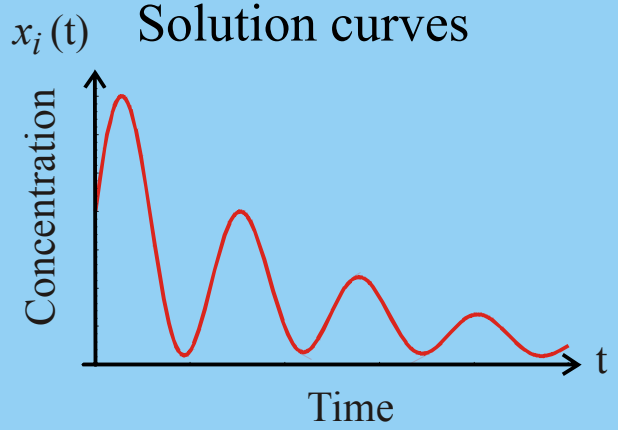
Stoichiometric equations

SBML – systems biology markup language

Kinetic differential equations

ODE Integration by means of *CVODE*

Solution curves

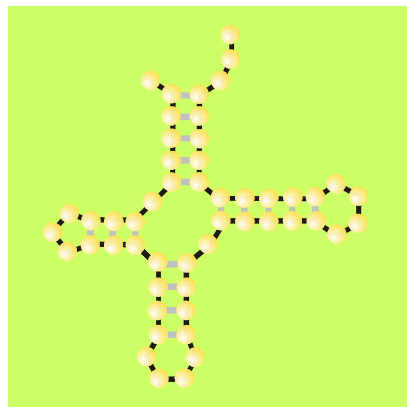


$$\frac{da}{dt} = \frac{db}{dt} = -k_1 a b$$
$$\frac{dx}{dt} = k_1 a b - k_2 x^2 - k_3 x y$$
$$\frac{dy}{dt} = k_2 x^2 - k_3 x y$$
$$\frac{dd}{dt} = k_3 x y$$

Sequences

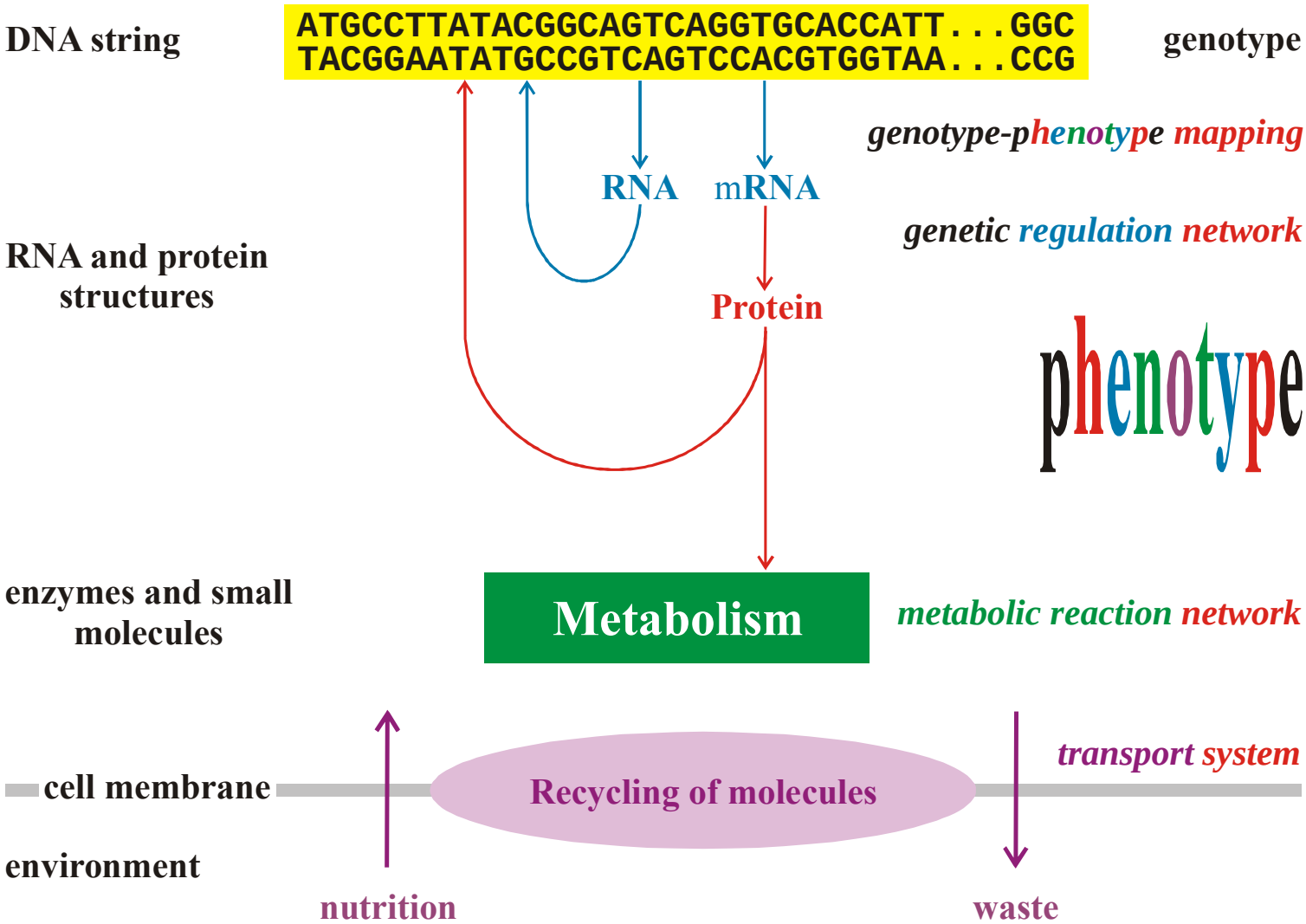
Vienna RNA Package

Structures and kinetic parameters

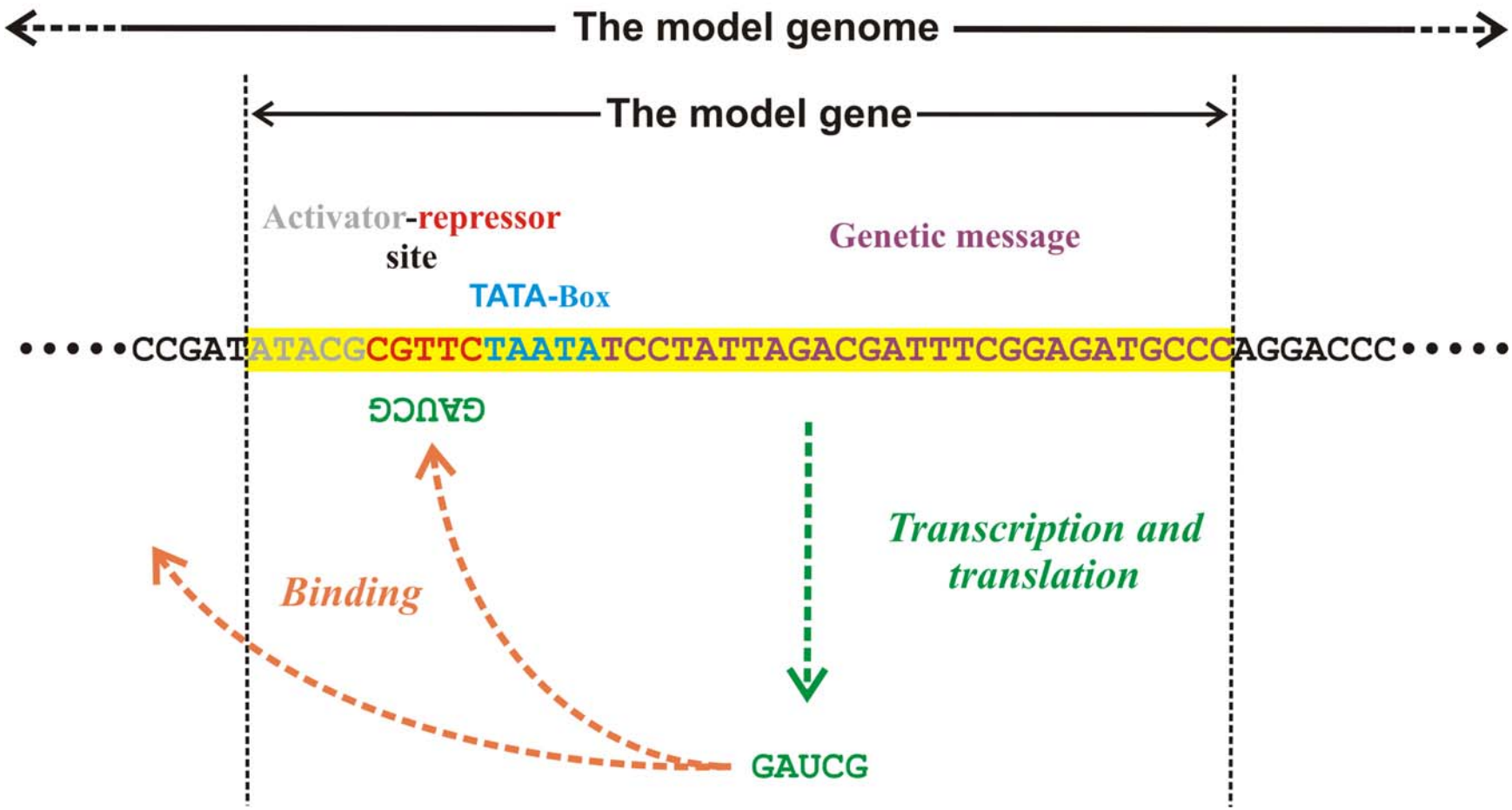


The elements of the simulation tool **MiniCellSim**

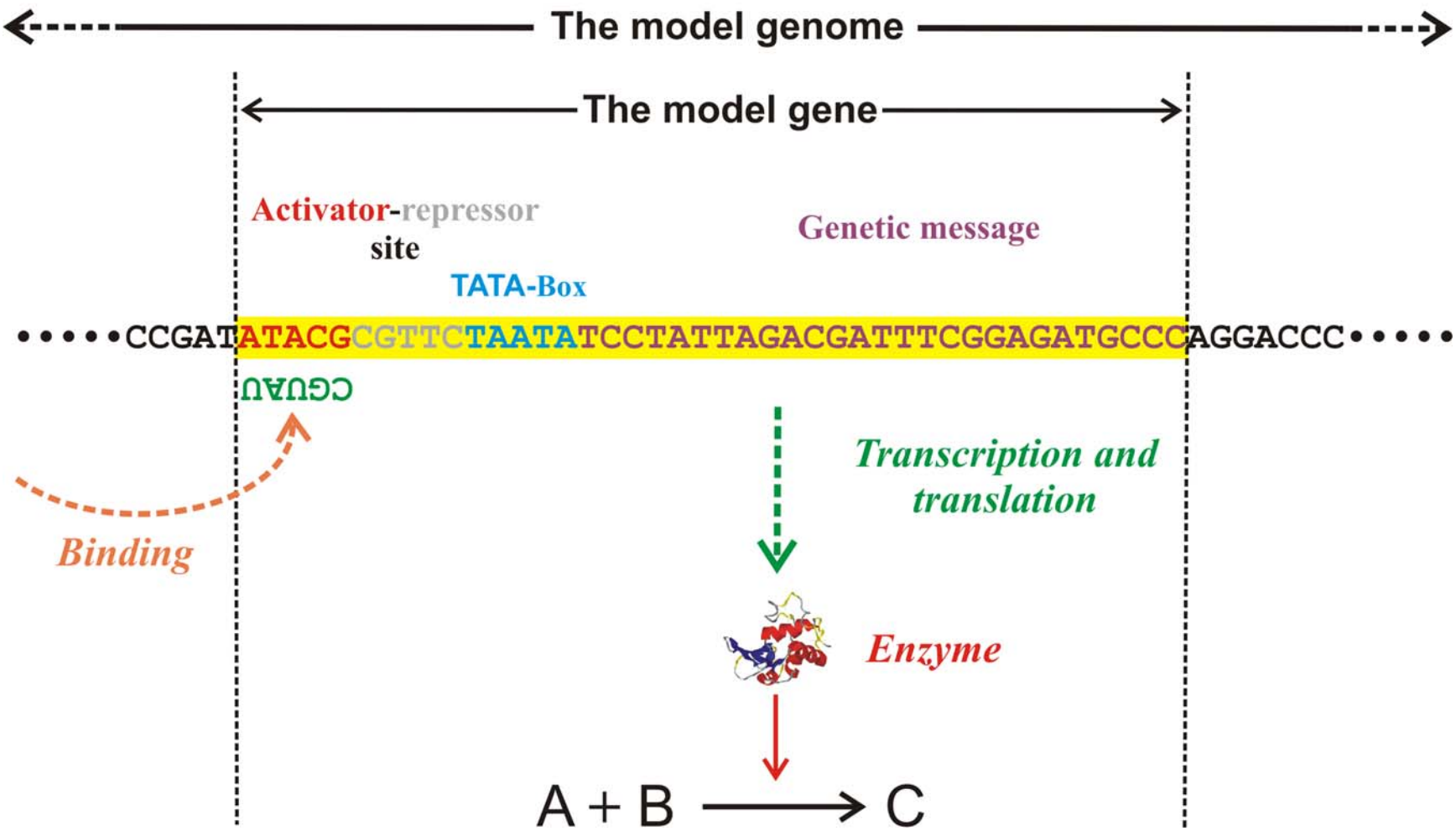
SBML: *Bioinformatics* **19**:524-531, 2003; *CVODE*: *Computers in Physics* **10**:138-143, 1996



The regulatory logic of MiniCellSym



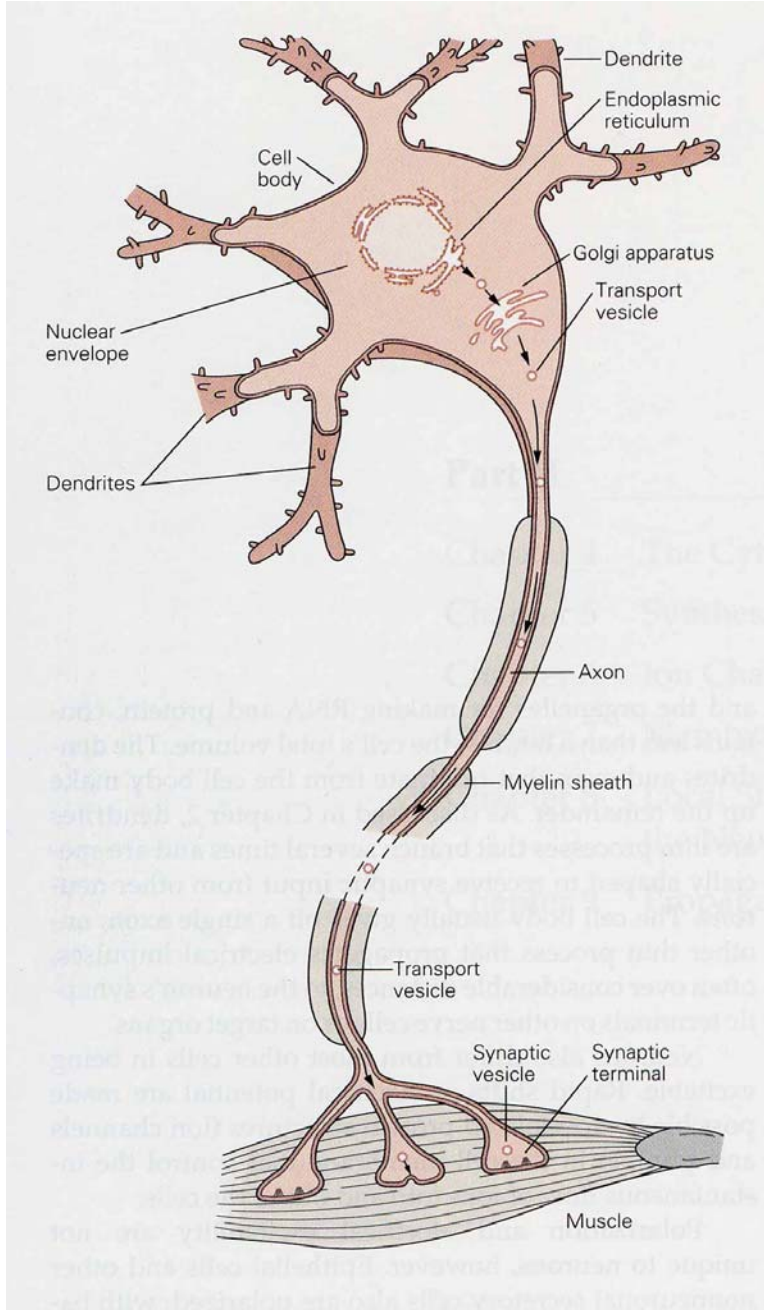
The model regulatory gene in MiniCellSim



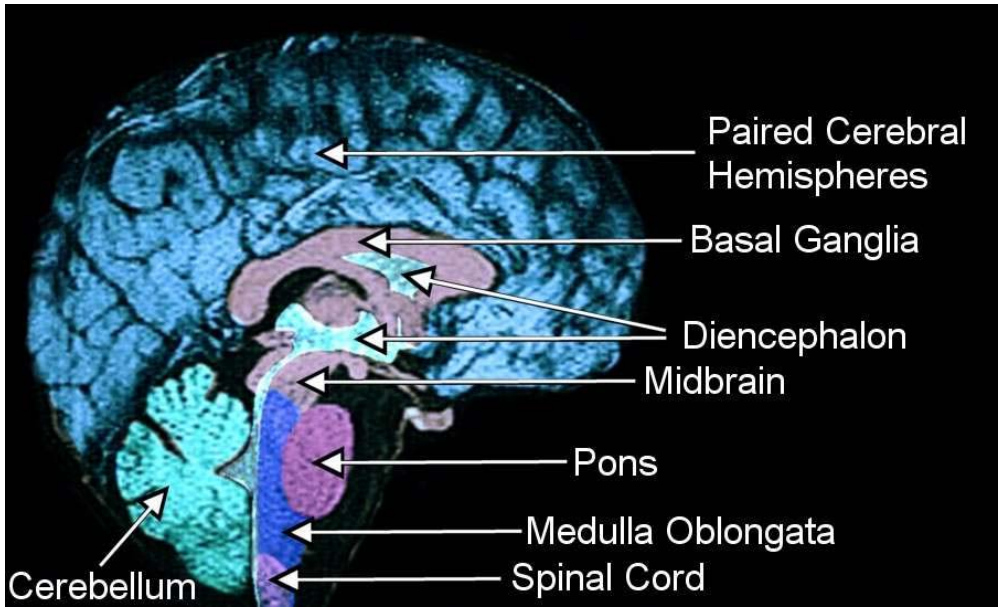
The model structural gene in MiniCellSim

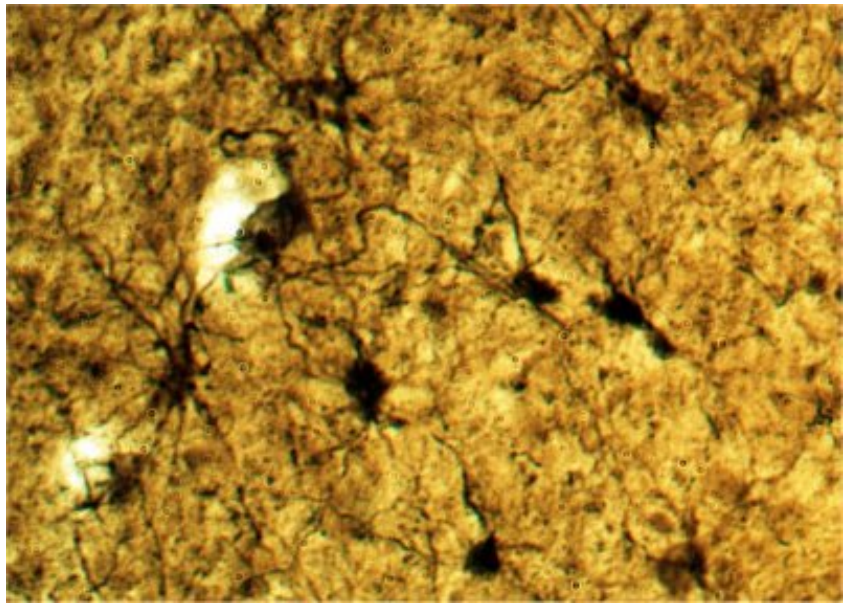
Neurobiology

Neural networks, collective properties, nonlinear dynamics, signalling, ...



A single neuron signaling to a muscle fiber





The human brain

10^{11} neurons connected by $\approx 10^{13}$ to 10^{14} synapses

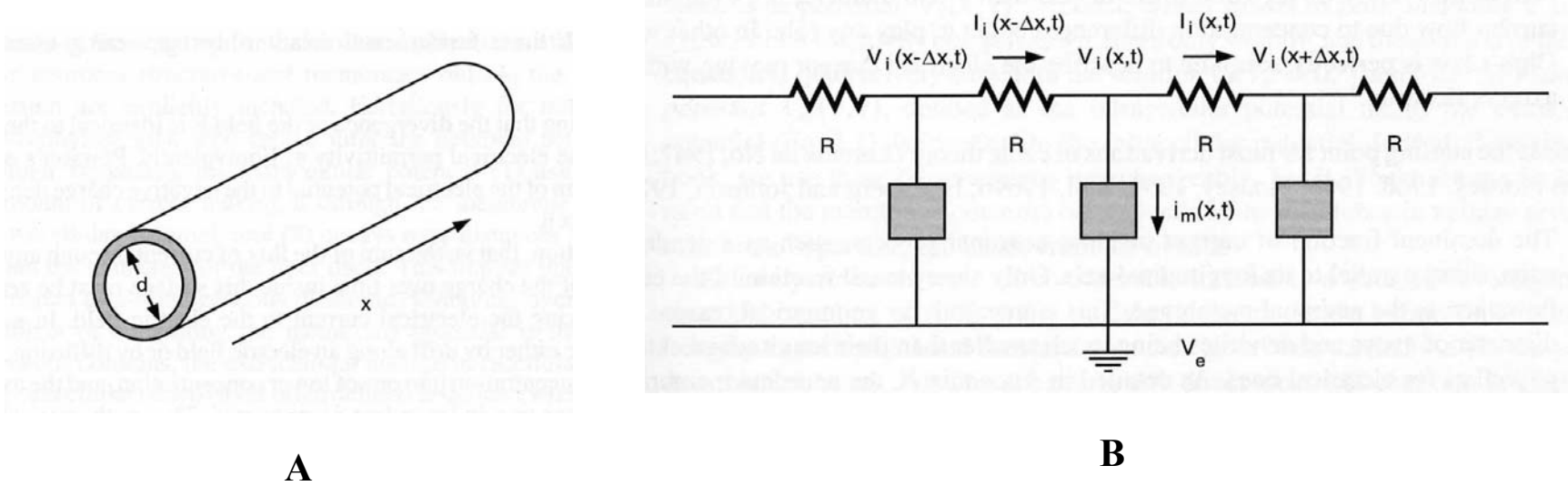


Fig. 2.2 ELECTRICAL STRUCTURE OF A CABLE **(A)** Idealized cylindrical axon or dendrite at the heart of one-dimensional cable theory. Almost all of the current inside the cylinder is longitudional due to geometrical (the radius is much smaller than the length of the cable) and electrical factors (the membrane covering the axon or dendrite possesses a very high resistivity compared to the intracellular cytoplasm). As a consequence, the radial and angular components of the current can be neglected, and the problem of determining the potential in these structures can be reduced from three spatial dimensions to a single one. On the basis of the bidomain approximation, gradients in the extracellular potentials are neglected and the cable problem is expressed in terms of the transmembrane potential $V_m(x,t) = V_i(x,t) - V_e$. **(B)** Equivalent electrical structure of an arbitrary neuronal process. The intracellular cytoplasm is modeled by the purely ohmic resistance R. This tacitly assumes that movement of carriers is exclusively due to drift along the voltage gradient and not to diffusion. Here and in the following the extracellular resistance is assumed to be negligible and V_e is set to zero. The current per unit length across the membrane, whether it is passive or contains voltage-dependent elements, is described by i_m and the system is characterized by the second-order differential equation, Eq. 2.5.

Christof Koch, Biophysics of Computation. Information Processing in single neurons.
Oxford University Press, New York 1999.

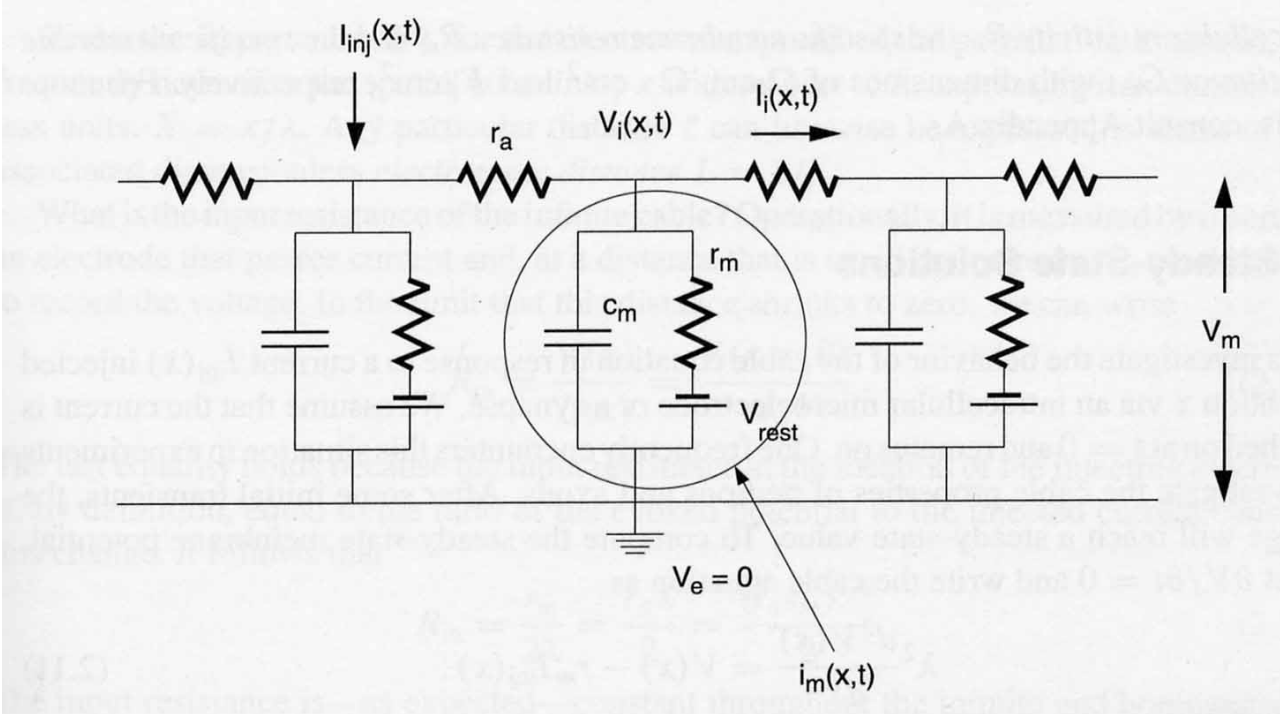


Fig. 2.3 A SINGLE PASSIVE CABLE Equivalent lumped electrical circuit of an elongated neuronal fiber with passive membrane. The intracellular cytoplasm is described by an ohmic resistance per unit length r_a and the membrane by a capacitance c_m in parallel with a passive membrane resistance r_m and a battery V_{rest}. The latter two components are frequently referred to as *leak resistance* and *leak battery*. An external current $I_{\text{inj}}(x,t)$ is injected into the cable. The associated linear cable equation (Eq. 2.7) describes the dynamics of the electrical potential $V_m = V_i - V_e$ along the cable.

Christof Koch, Biophysics of Computation. Information Processing in single neurons. Oxford University Press, New York 1999.

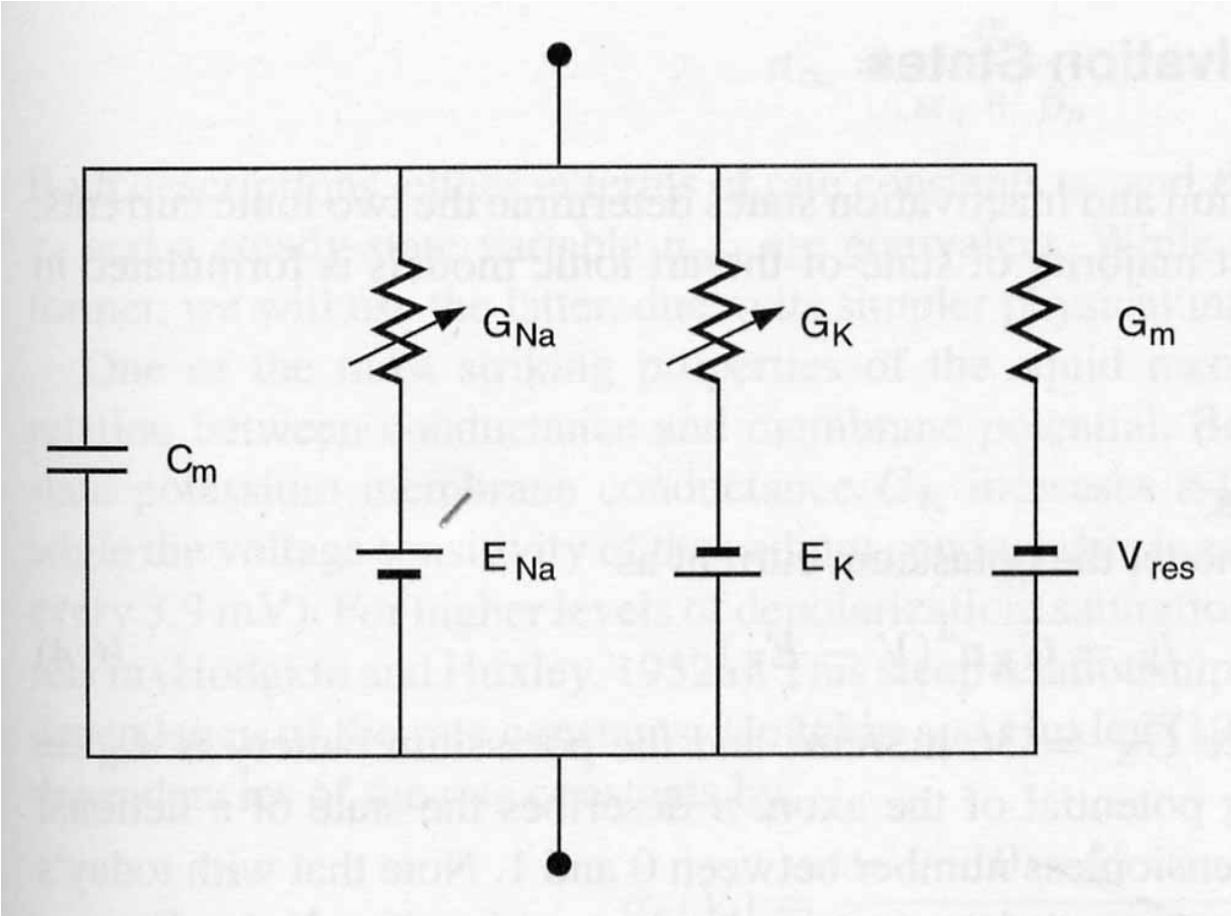


Fig. 6.2 Electrical Circuit for a Patch of Squid Axon Hodgkin and Huxley modeled the membrane of the squid axon using four parallel branches: two passive ones (membrane capacitance C_m and the leak conductance $G_m = 1/R_m$) and two time- and voltage-dependent ones representing the sodium and potassium conductances.

Christof Koch, Biophysics of Computation. Information Processing in single neurons. Oxford University Press, New York 1999.

Neurobiology

Neural networks, collective properties, nonlinear dynamics, signalling, ...

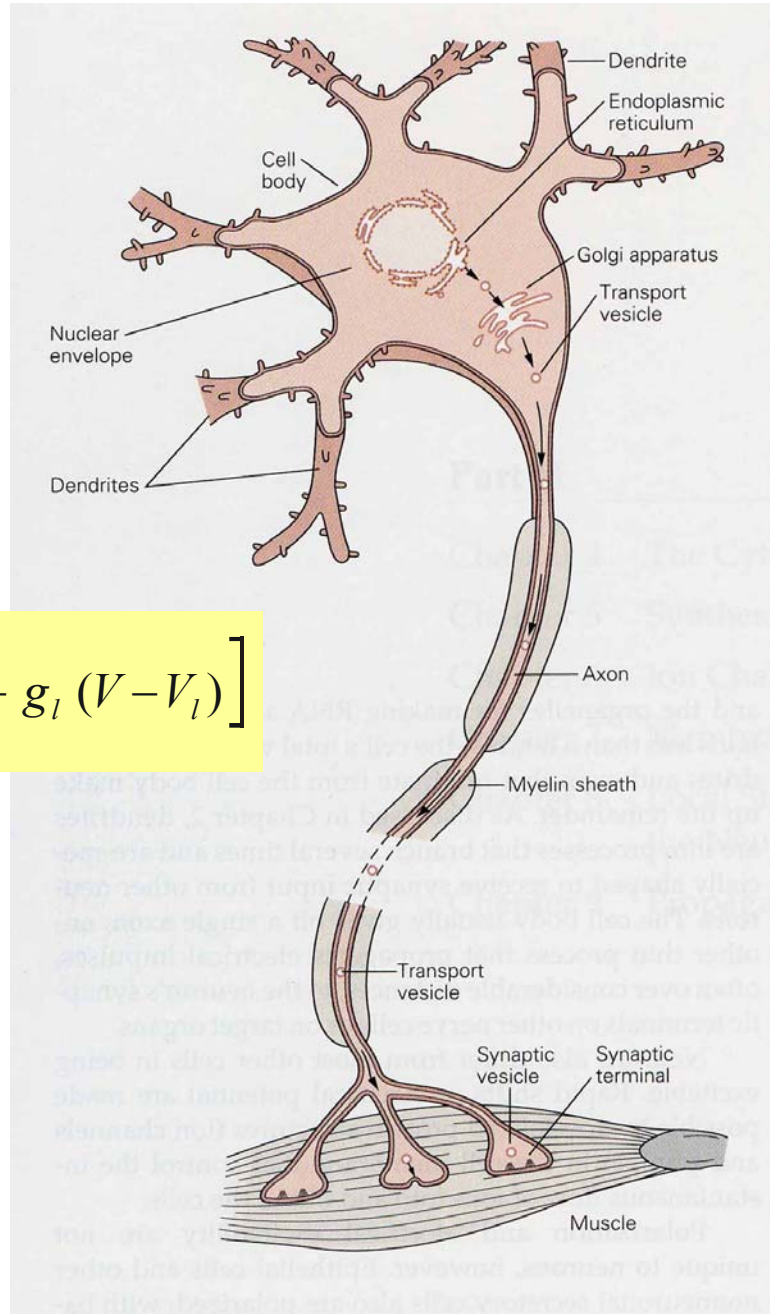
$$\frac{dV}{dt} = \frac{1}{C_M}\left[I - g_{Na}\, m^3\, h\,(V - V_{Na}) - g_K\, n^4\,(V - V_K) - g_l\,(V - V_l)\right]$$

$$\frac{dm}{dt} = \alpha_m\,(1-m) - \beta_m\, m$$

$$\frac{dh}{dt} = \alpha_h\,(1-h) - \beta_h\, h$$

$$\frac{dn}{dt} = \alpha_n\,(1-n) - \beta_n\, n$$

Hogdkin-Huxley OD equations



A single neuron signaling to a muscle fiber

$$\alpha_m = \frac{x}{e^x - 1}, \quad x = \frac{25 - V}{10} \; ; \quad \beta_m = 4 \exp\left[- \frac{V}{18} \right]$$

$$\alpha_h = 0.07 \exp\left[- \frac{V}{20} \right] \; ; \quad \beta_h = \frac{1}{e^x - 1}, \quad x = \frac{30 - V}{10}$$

$$\alpha_n = \frac{x}{10(e^x - 1)}, \quad x = \frac{10 - V}{10} \; ; \quad \beta_n = 0.125 \exp\left[- \frac{V}{80} \right]$$

Gating functions of the Hodgkin-Huxley equations

$$\frac{\partial m}{\partial t} = \Theta(T)\big[\alpha_m(1-m) - \beta_m m\big]$$

$$\frac{\partial h}{\partial t} = \Theta(T)\big[\alpha_h(1-h) - \beta_h h\big]$$

$$\frac{\partial n}{\partial t} = \Theta(T)\big[\alpha_n(1-n) - \beta_n n\big] \ ,$$

$$\text{where } \ \Theta(T) = 3^{(T-6.3)/10}$$

Temperature dependence of the Hodgkin-Huxley equations

$$\frac{dV}{dt} = \frac{1}{C_M}\left[I - \bar{g}_{Na}\, m^3\, h\,(V - V_{Na}) - \bar{g}_K\, n^4\,(V - V_K) - \bar{g}_l\,(V - V_l)\right]$$

$$\frac{dm}{dt} = \alpha_m\,(1-m) - \beta_m\, m$$

$$\frac{dh}{dt} = \alpha_h\,(1-h) - \beta_h\, h$$

$$\frac{dn}{dt} = \alpha_n\,(1-n) - \beta_n\, n$$

Hogdkin-Huxley OD equations

Hhsim.lnk

Simulation of space independent Hodgkin-Huxley equations:
Voltage clamp and constant current

$$\frac{1}{R}\frac{\partial^2 V}{\partial x^2} = C\frac{\partial V}{\partial t} + [g_{Na}\, m^3\, h\,(V - V_{Na}) + g_K\, n^4\,(V - V_K) + g_l\,(V - V_l)]\,2\pi\, r\, L$$

$$\frac{\partial m}{\partial t} = \alpha_m (1 - m) - \beta_m m$$

$$\frac{\partial h}{\partial t} = \alpha_h (1 - h) - \beta_h h$$

$$\frac{\partial n}{\partial t} = \alpha_n (1 - n) - \beta_n n$$

Hodgkin-Huxley partial differential equations (PDE)

Hodgkin-Huxley equations describing pulse propagation along nerve fibers

$$\frac{1}{R}\frac{\partial^2 V}{\partial \xi^2} = C_M \theta \frac{\partial V}{\partial \xi} + [g_{Na}\, m^3\, h\,(V - V_{Na}) + g_K\, n^4\,(V - V_K) + g_l\,(V - V_l)]\, 2\pi\, r\, L$$

$$\theta \frac{\partial m}{\partial \xi} = \alpha_m (1 - m) - \beta_m m$$

$$\theta \frac{\partial h}{\partial \xi} = \alpha_h (1 - h) - \beta_h h$$

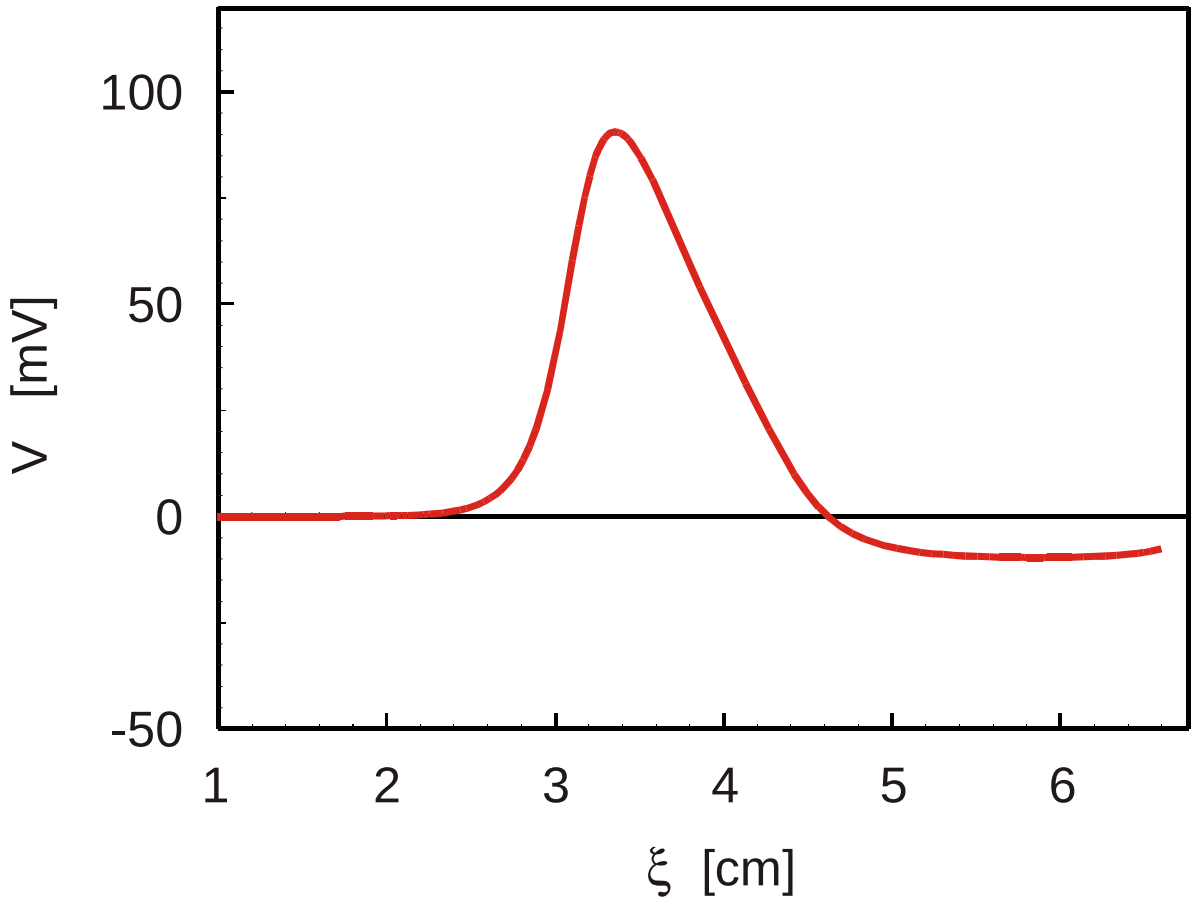
$$\theta \frac{\partial n}{\partial \xi} = \alpha_n (1 - n) - \beta_n n$$

Hodgkin-Huxley ordinary differential equations (ODE)

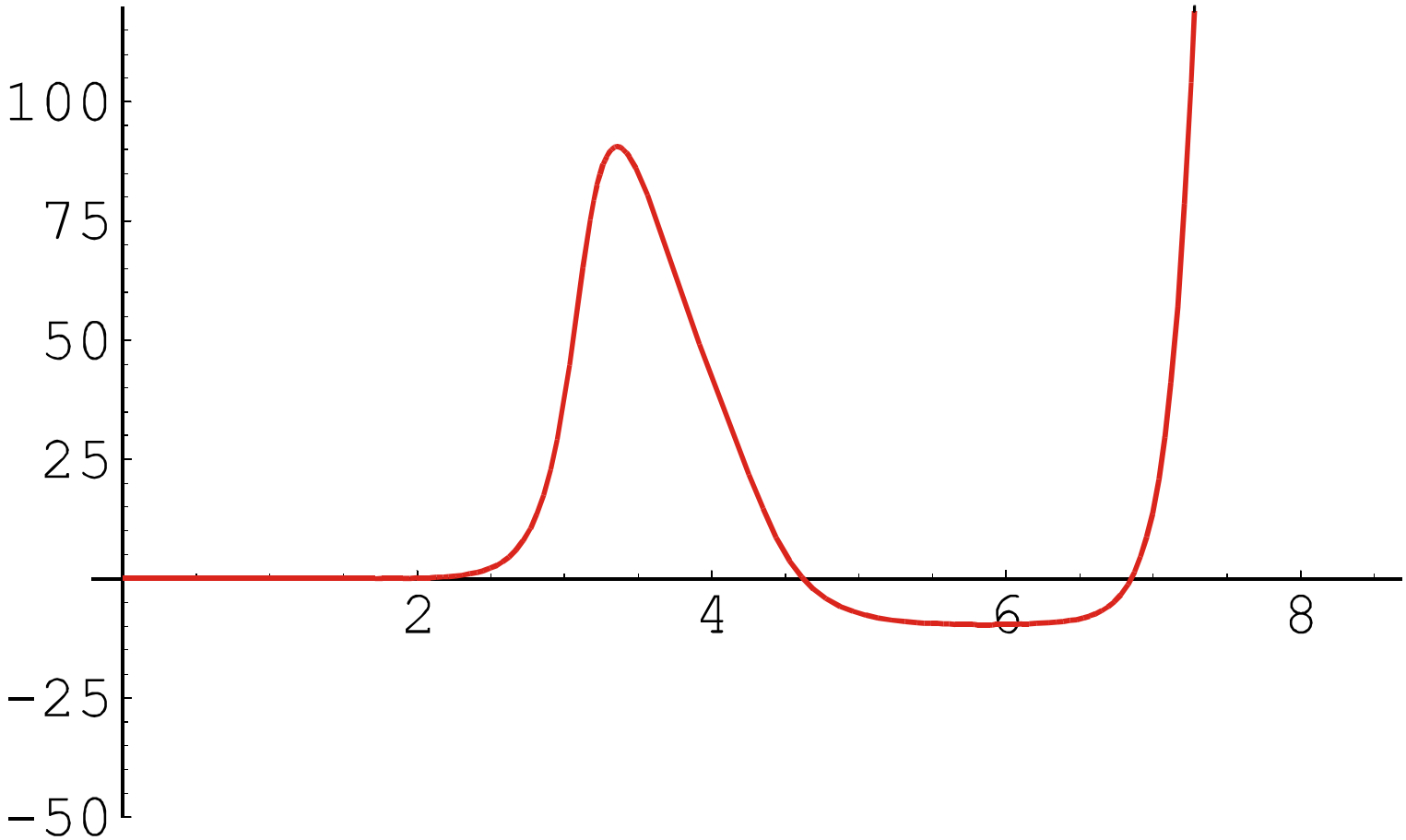
Travelling pulse solution: $V(x,t) = V(\xi)$ with

$$\xi = x + \theta\, t$$

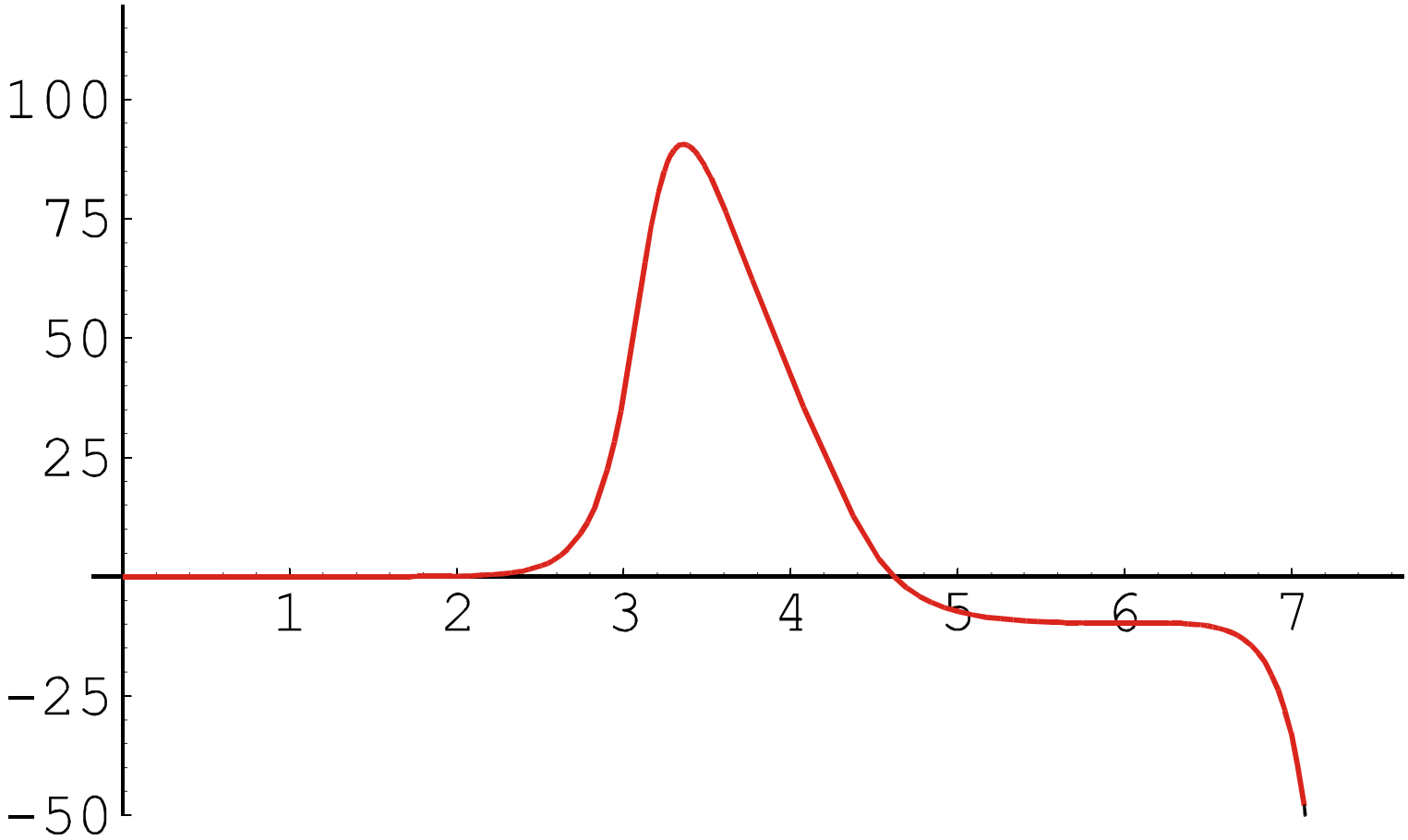
Hodgkin-Huxley equations describing pulse propagation along nerve fibers



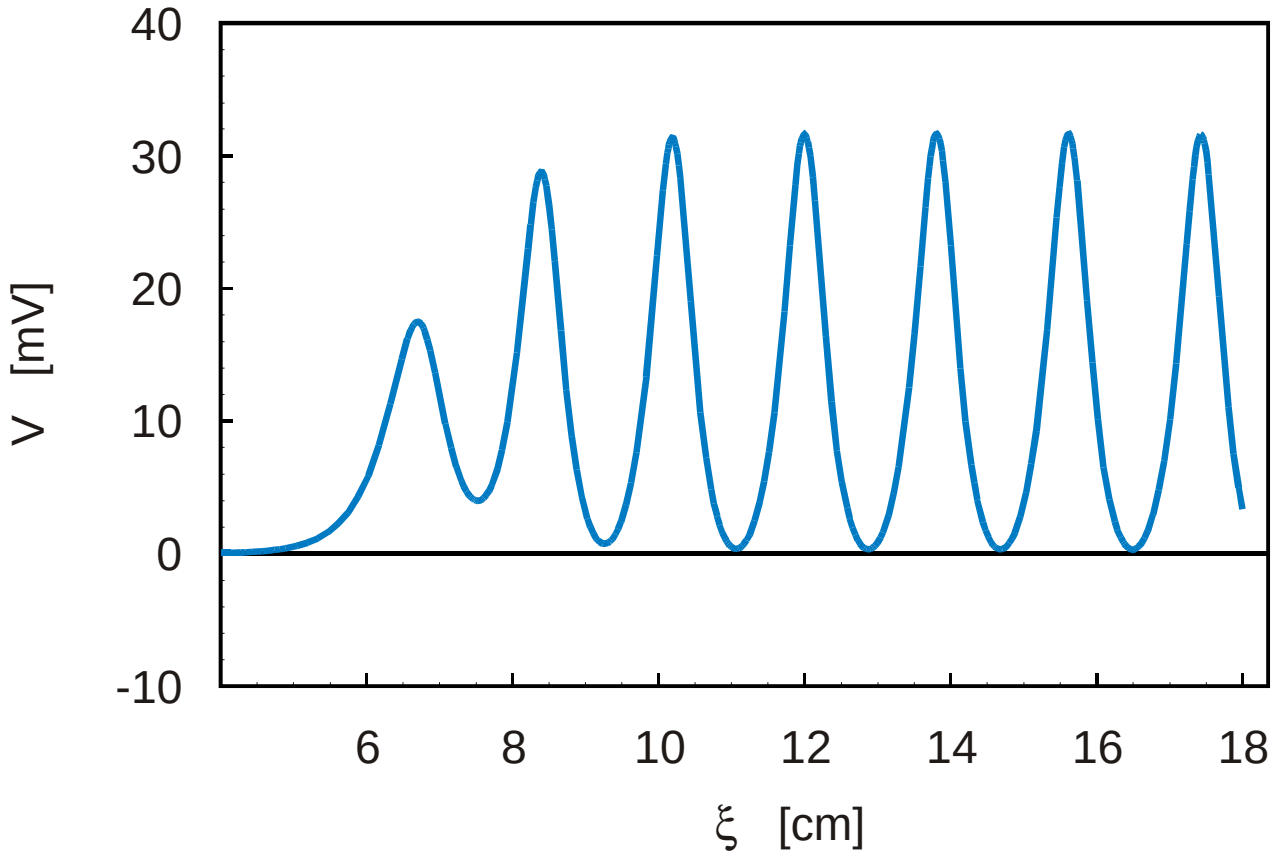
T = 18.5 C; θ = 1873.33 cm / sec



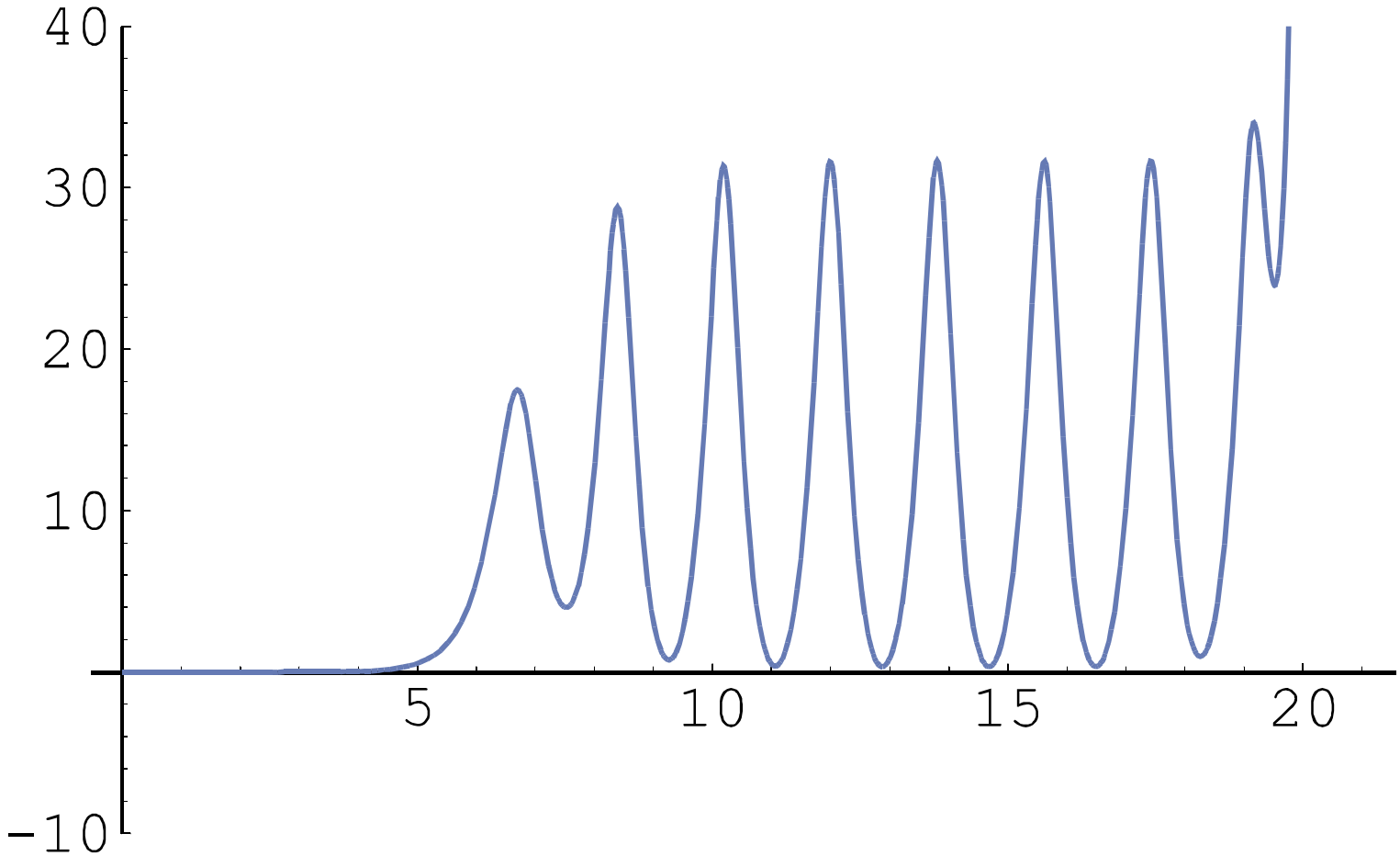
T = 18.5 C; θ = 1873.3324514717698 cm / sec



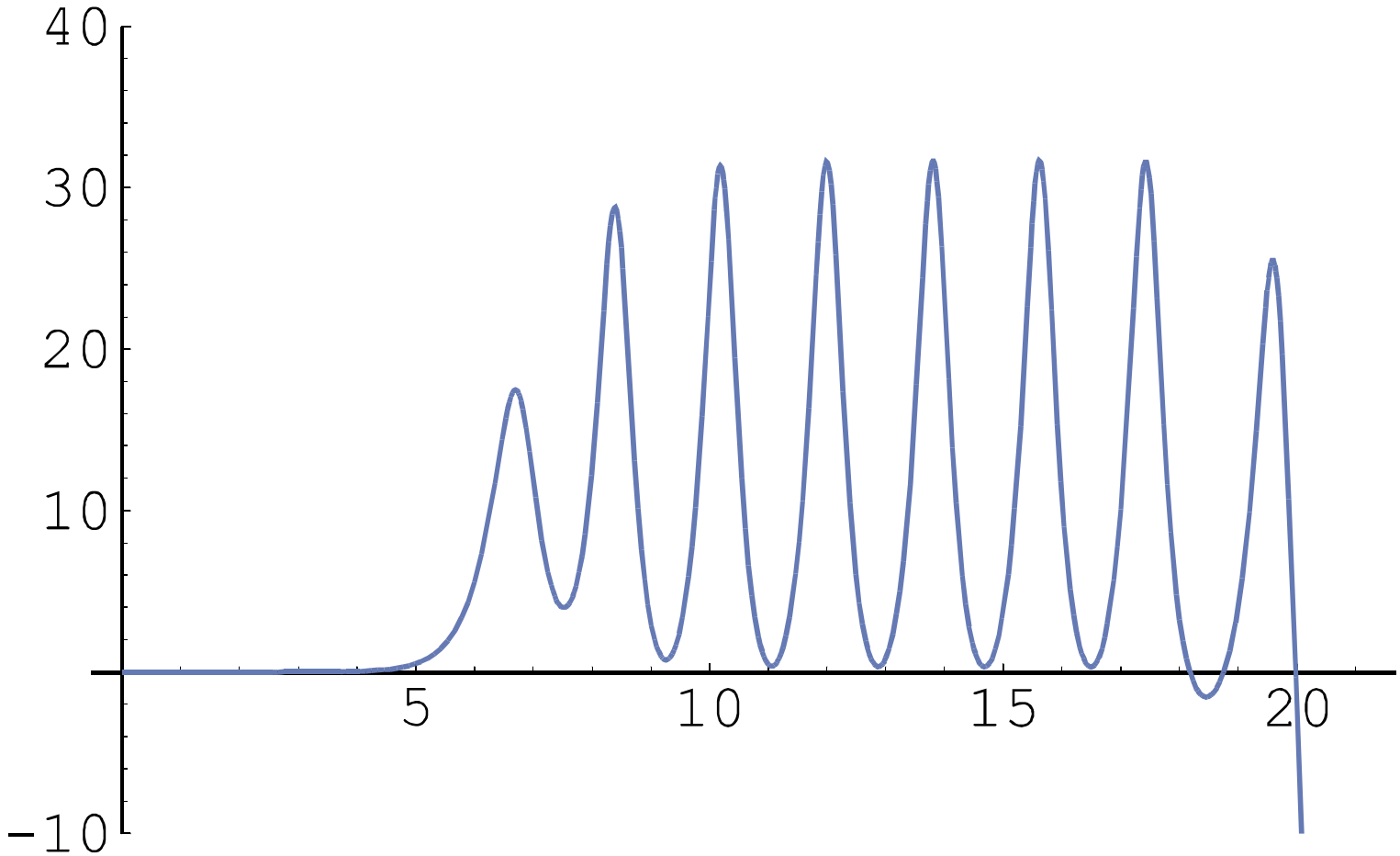
T = 18.5 C; θ = 1873.3324514717697 cm / sec



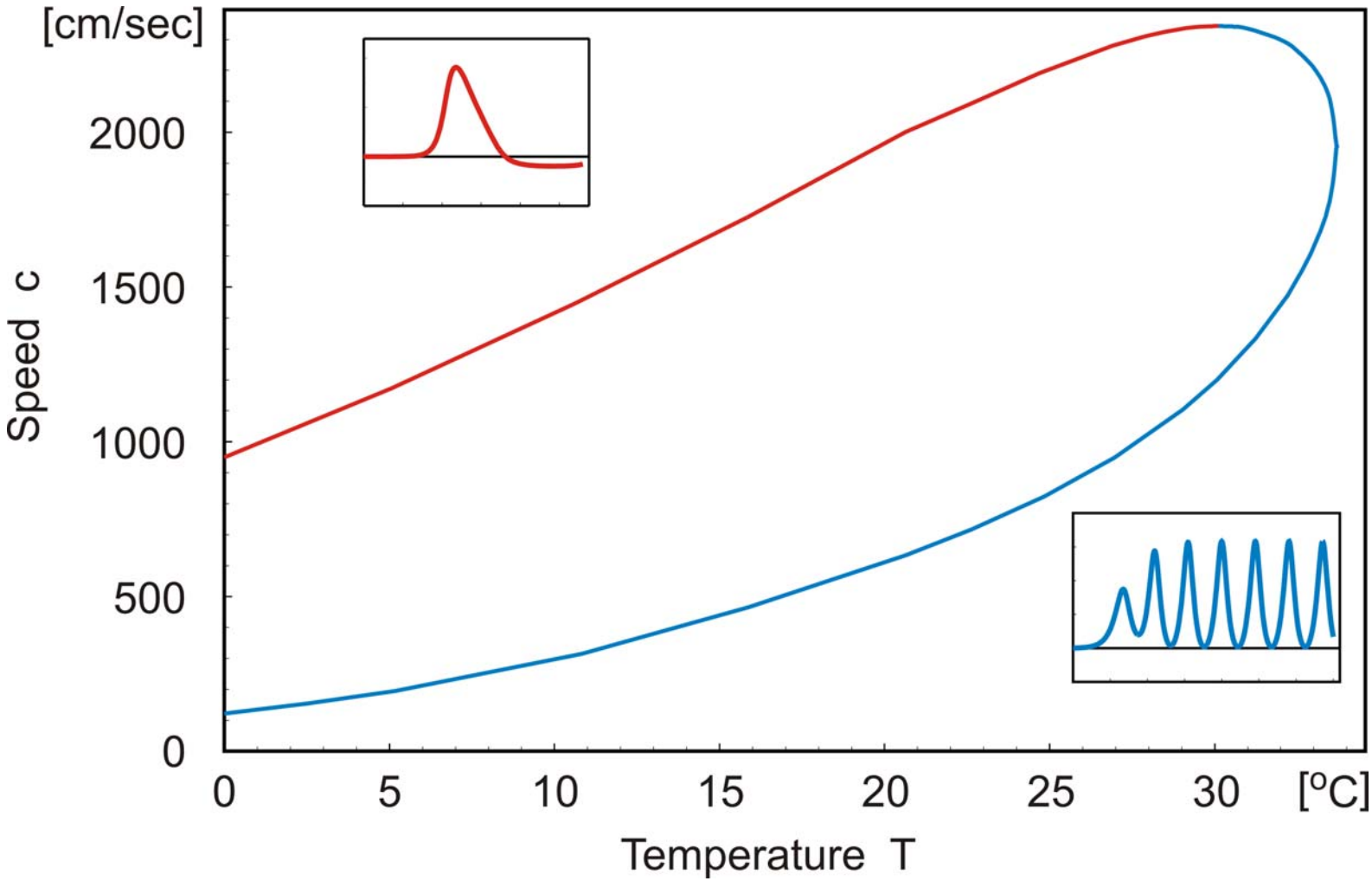
T = 18.5 C; θ = 544.070 cm / sec



T = 18.5 C; θ = 554.070286919319 cm/sec



T = 18.5 C; θ = 554.070286919320 cm/sec



Propagating wave solutions of the Hodgkin-Huxley equations

$$\frac{1}{R}\frac{\partial^2 V}{\partial \xi^2} = C_M \theta \frac{\partial V}{\partial \xi} + [g_{Na}\, m^3 (h_0 + n_0 - n)(V - V_{Na}) + g_K\, n^4 (V - V_K) + g_l (V - V_l)]\, 2\pi r L$$

$$\theta \frac{\partial m}{\partial \xi} = \alpha_m (1 - m) - \beta_m m$$

$$\theta \frac{\partial n}{\partial \xi} = \alpha_n (1 - n) - \beta_n n$$

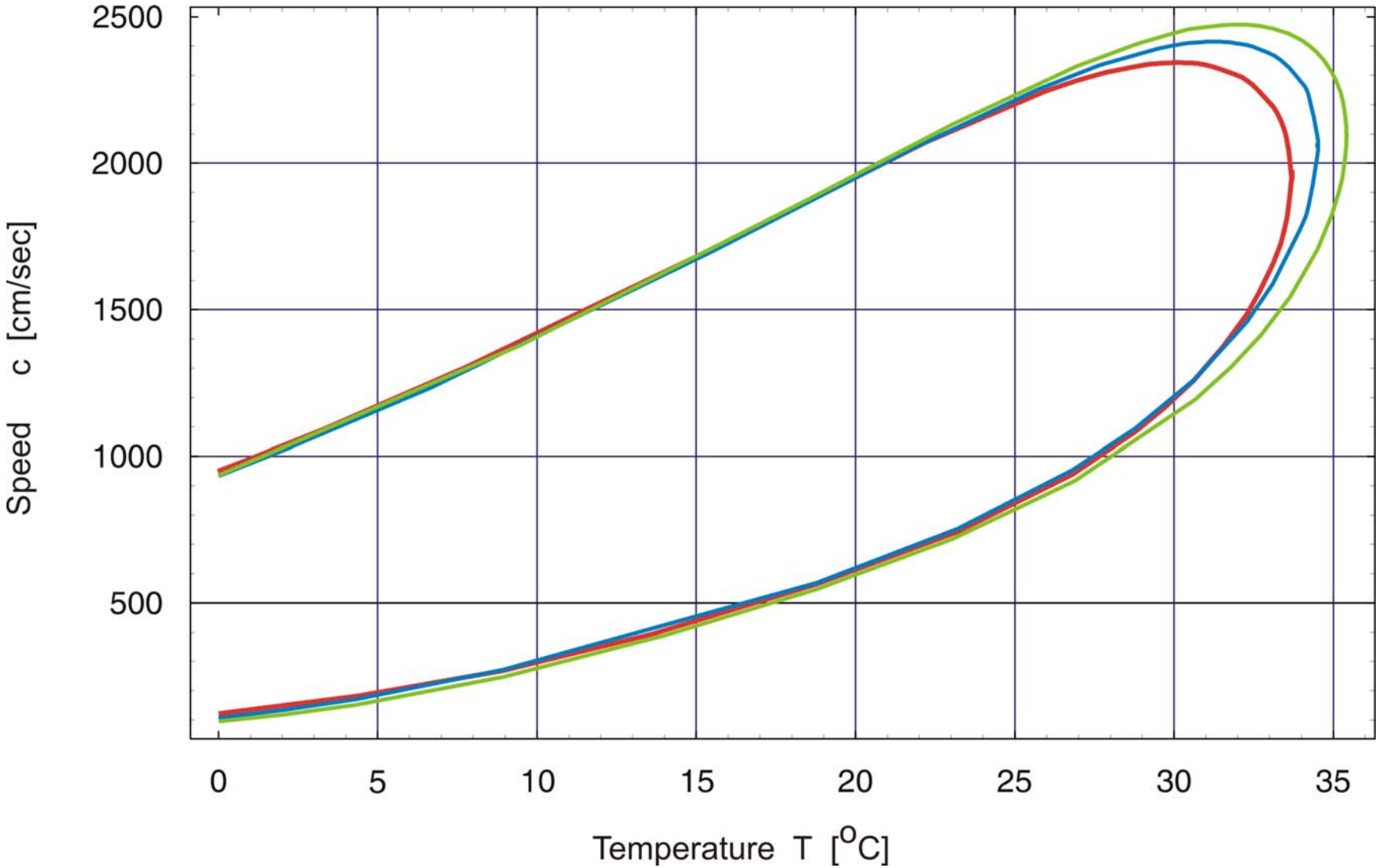
Hodgkin-Huxley ordinary differential equations (ODE)

Travelling pulse solution: $V(x,t) = V(\xi)$ with

$$\xi = x + \theta\, t$$

$$\alpha_n = \frac{V}{E_{Na}} + \alpha_0 \; ; \quad \beta_n = 0.125 \exp(-V/80) \approx 0.125\,(1 - \tfrac{V}{80})$$

An approximation to the Hodgkin-Huxley equations



Propagating wave solutions of approximations to the Hodgkin-Huxley equations

Evolutionary biology

Optimization through variation and selection, relation between genotype, phenotype, and function, ...

	Generation time	Selection and adaptation 10 000 generations	Genetic drift in small populations 10^6 generations	Genetic drift in large populations 10^7 generations
RNA molecules	10 sec 1 min	27.8 h = 1.16 d 6.94 d	115.7 d 1.90 a	3.17 a 19.01 a
Bacteria	20 min 10 h	138.9 d 11.40 a	38.03 a 1 140 a	380 a 11 408 a
Multicelluar organisms	10 d 20 a	274 a 200 000 a	27 380 a 2×10^7 a	273 800 a 2×10^8 a

Time scales of evolutionary change

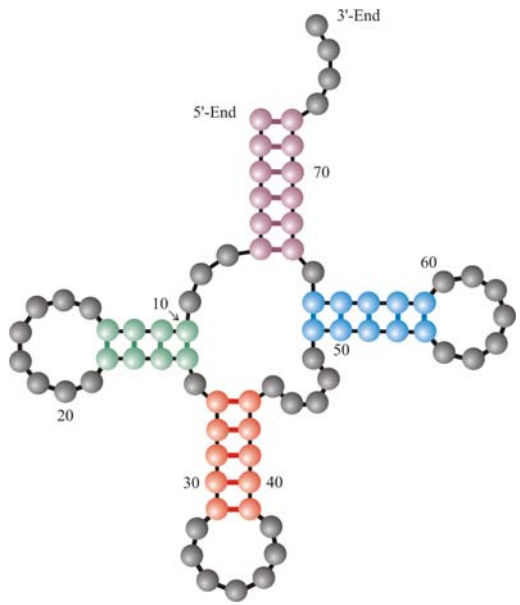
Genotype = Genome

Mutation → GGCUAUCGUACGUUUACCCAAAAAGUCUACGUUGGACCCAGGCAUUGGAC.....G

Unfolding of the genotype:
RNA structure formation

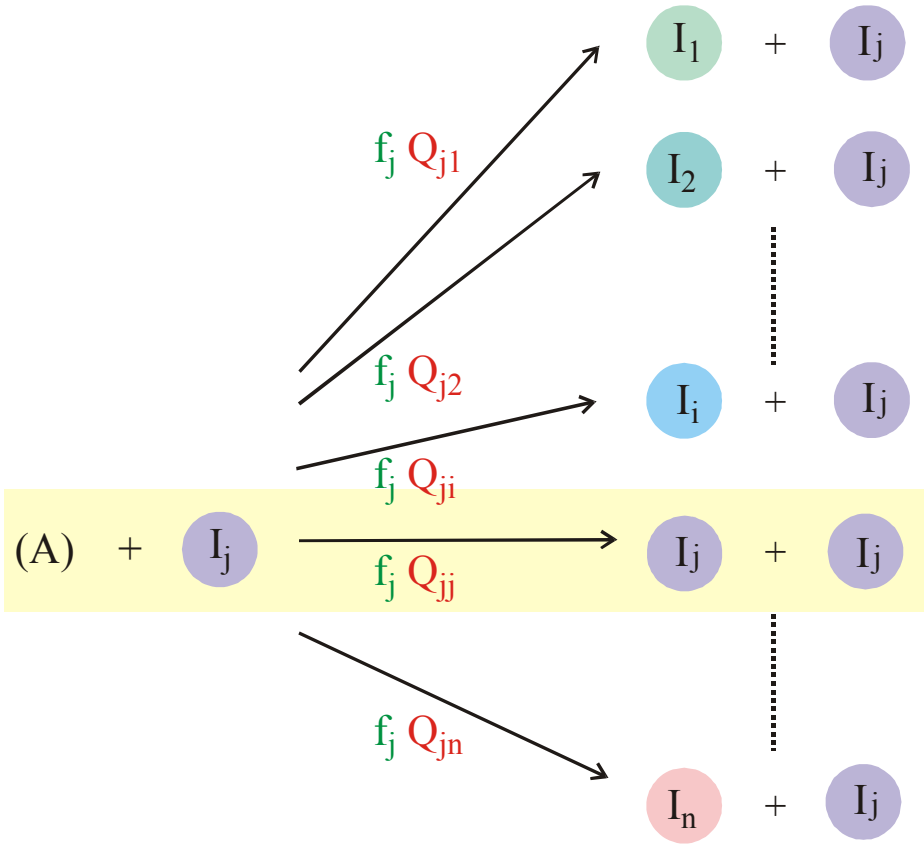
Fitness in reproduction:
Number of genotypes in the next generation

Phenotype



Selection →

Evolution of phenotypes



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

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

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

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

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

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

Mutation-selection equation: $[\mathrm{I}_i] = x_i \geq 0,\ f_i > 0,\ Q_{ij} \geq 0$

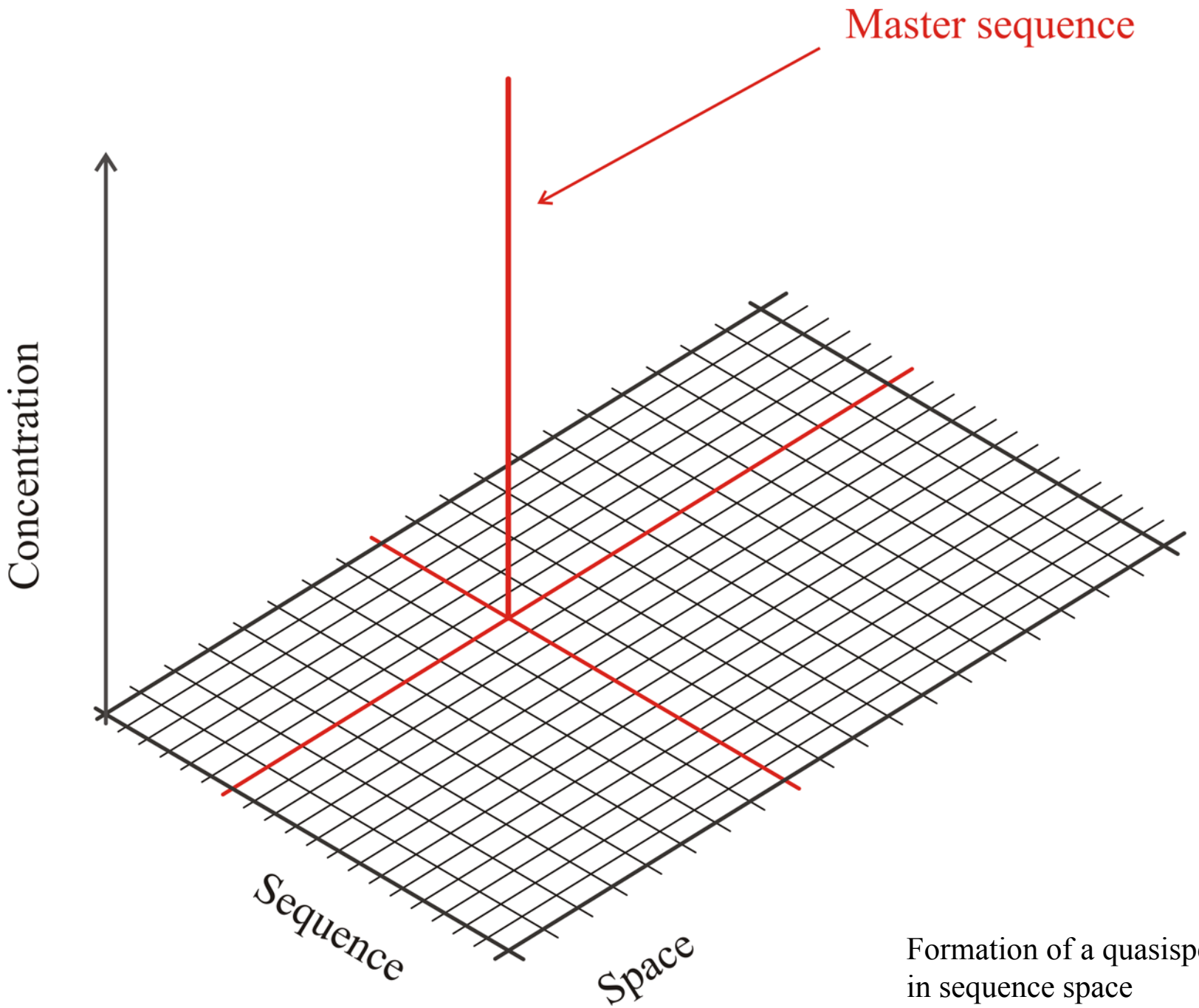
$$\frac{dx_i}{dt} = \sum_{j=1}^{n} f_j Q_{ji}\, x_j - x_i\, \phi, \quad i = 1,2,\cdots,n; \quad \sum_{i=1}^{n} x_i = 1; \quad \phi = \sum_{j=1}^{n} f_j x_j = \overline{f}$$

Solutions are obtained after integrating factor transformation by means of an eigenvalue problem

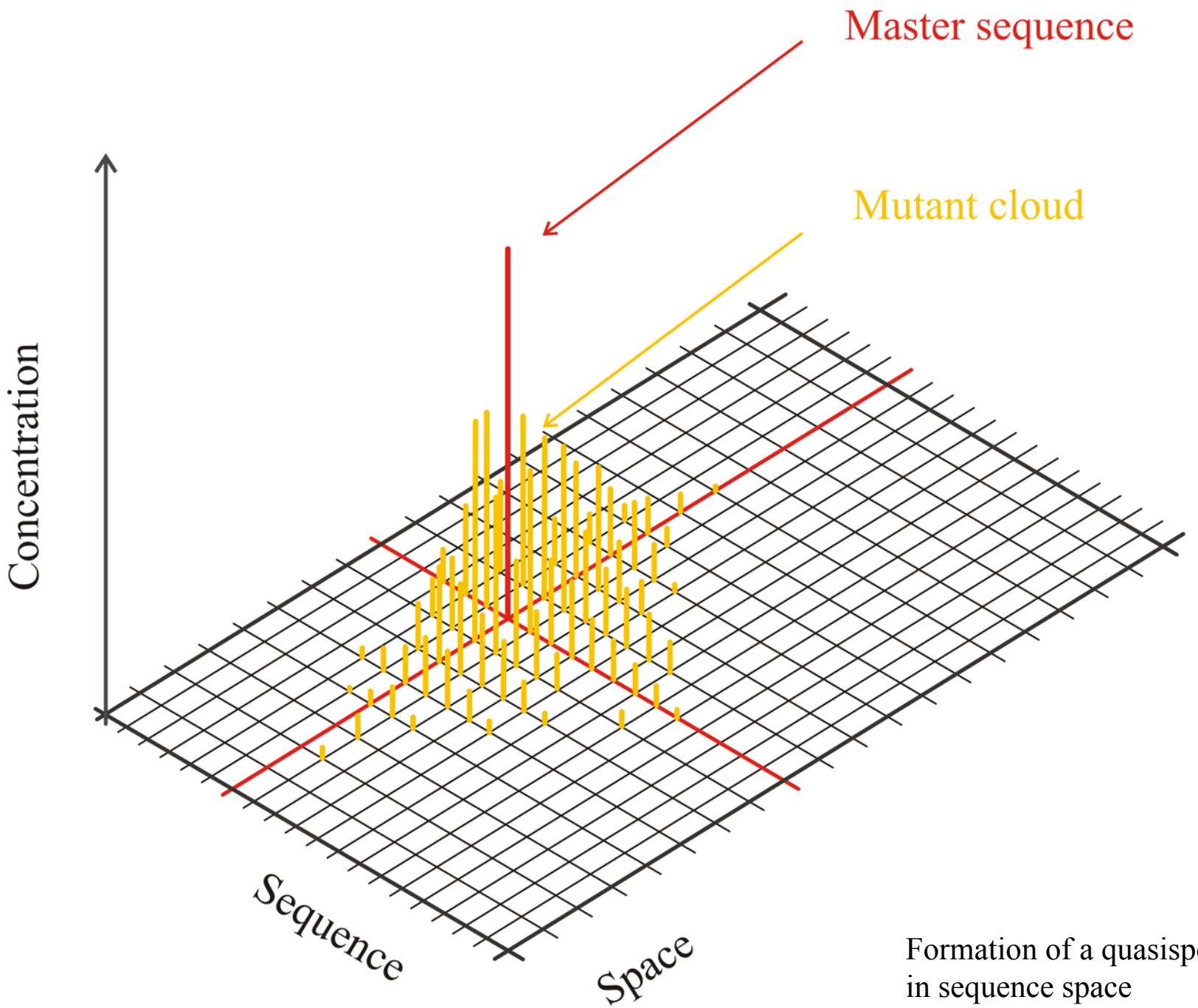
$$x_i(t) = \frac{\sum_{k=0}^{n-1} \ell_{ik} \cdot c_k(0) \cdot \exp(\lambda_k t)}{\sum_{j=1}^{n} \sum_{k=0}^{n-1} \ell_{jk} \cdot c_k(0) \cdot \exp(\lambda_k t)}; \quad i = 1,2,\cdots,n; \quad c_k(0) = \sum_{i=1}^{n} h_{ki}\, x_i(0)$$

$$W \div \left\{ f_i Q_{ij};\ i, j = 1,2,\cdots,n \right\};\ L = \left\{ \ell_{ij};\ i, j = 1,2,\cdots,n \right\};\ L^{-1} = H = \left\{ h_{ij};\ i, j = 1,2,\cdots,n \right\}$$

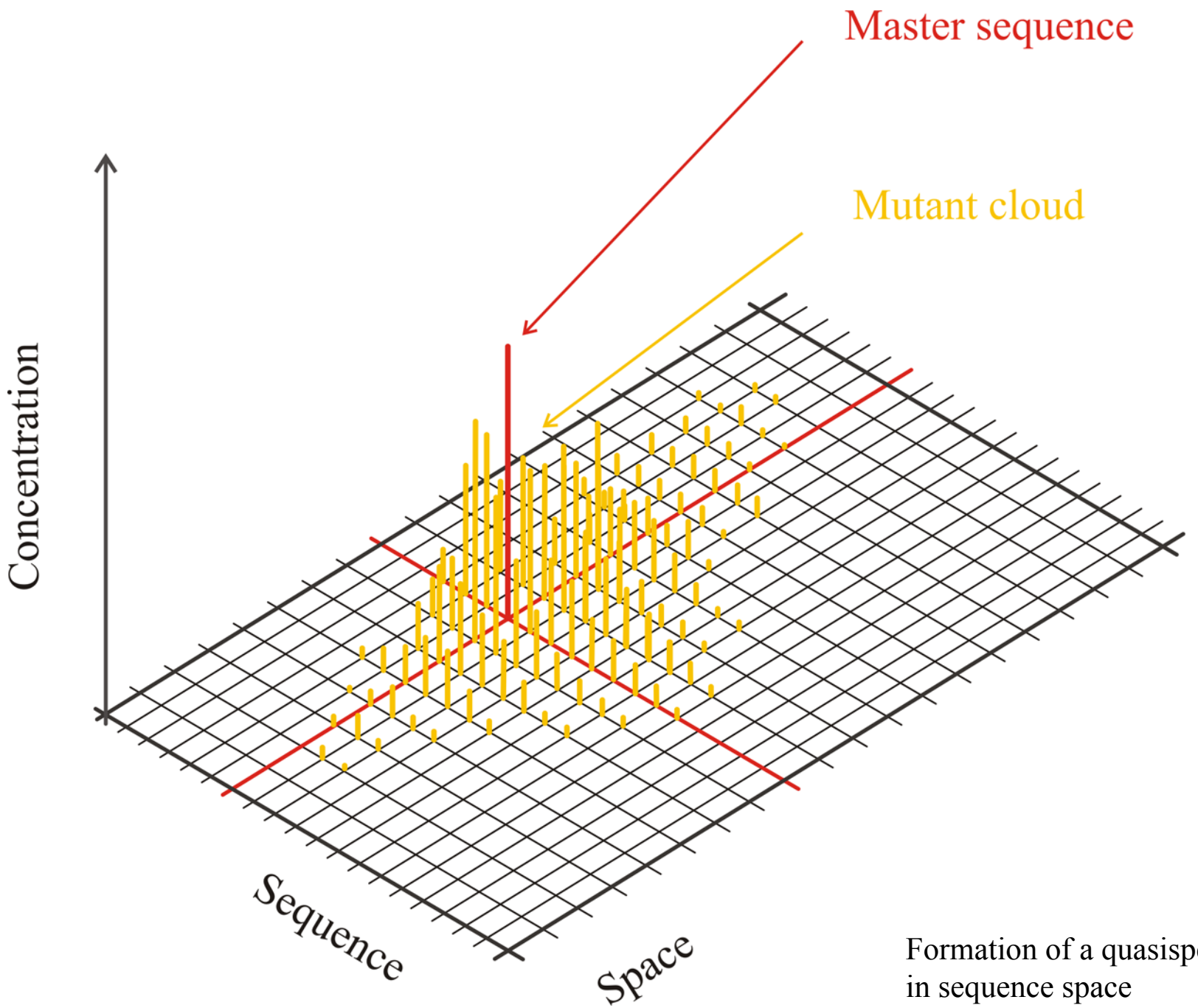
$$L^{-1} \cdot W \cdot L \quad = \quad \Lambda \quad = \quad \left\{ \lambda_k; k = 0, 1, \cdots, n-1 \right\}$$



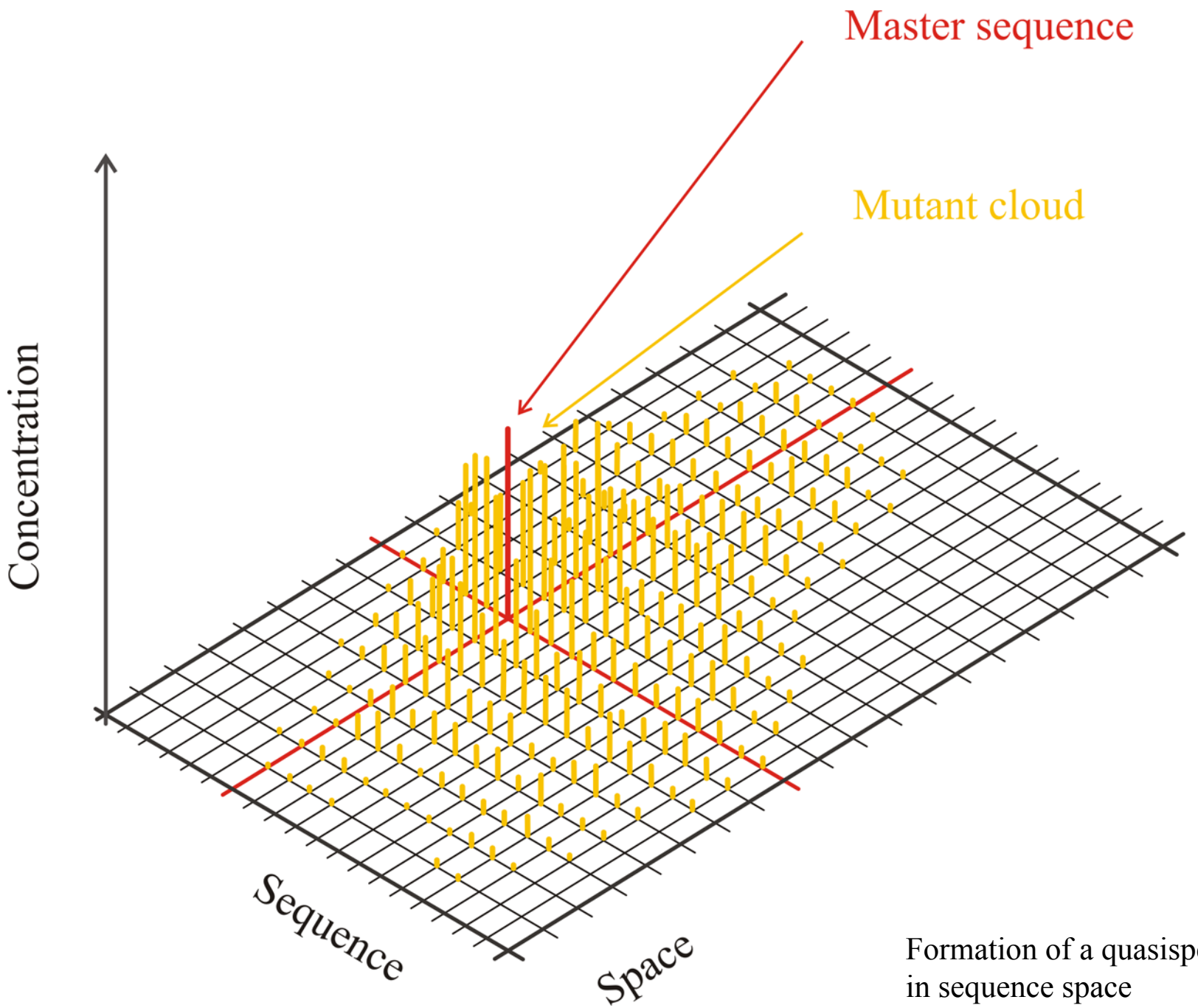
Formation of a quasispecies in sequence space



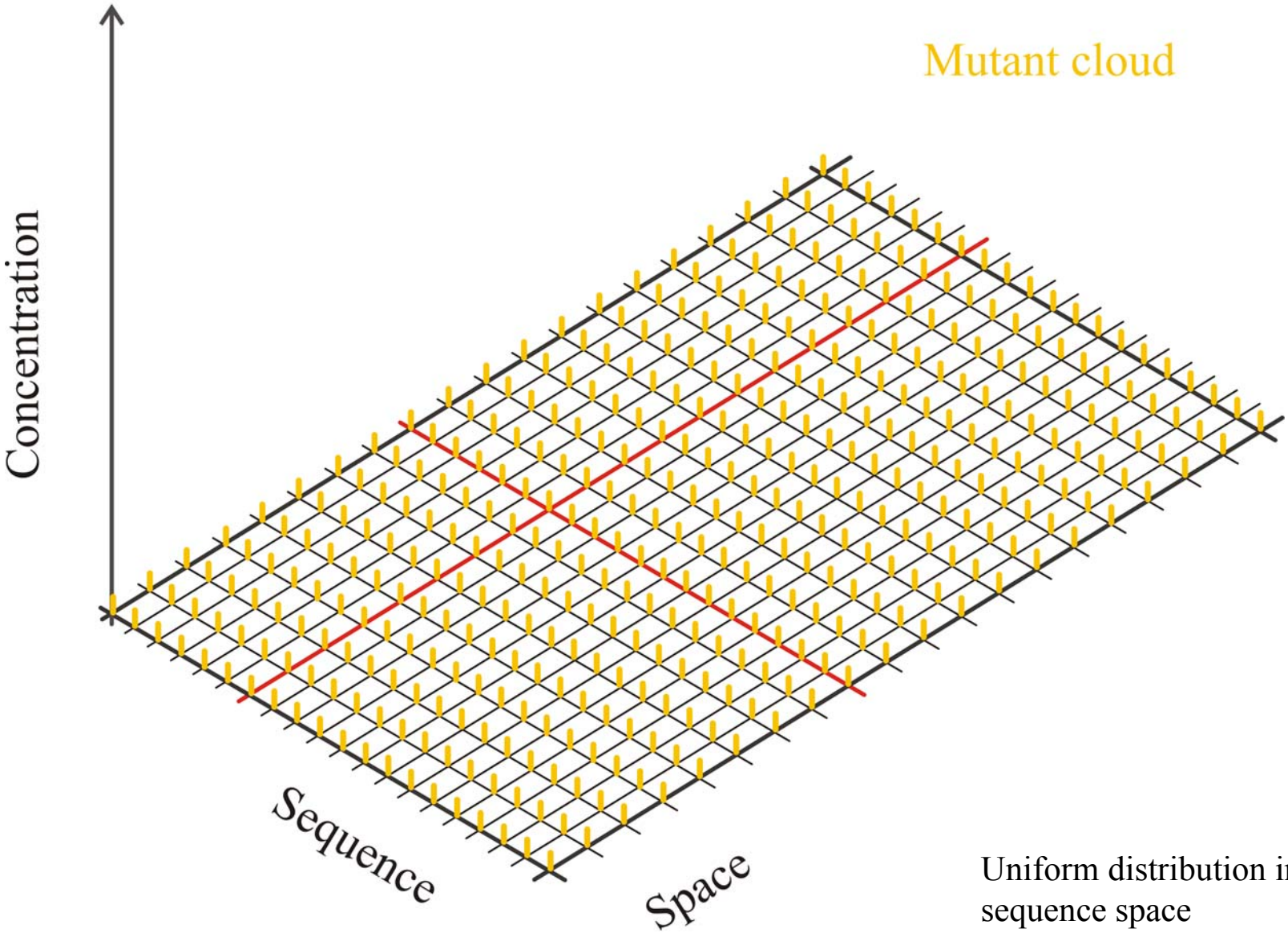
Formation of a quasispecies in sequence space



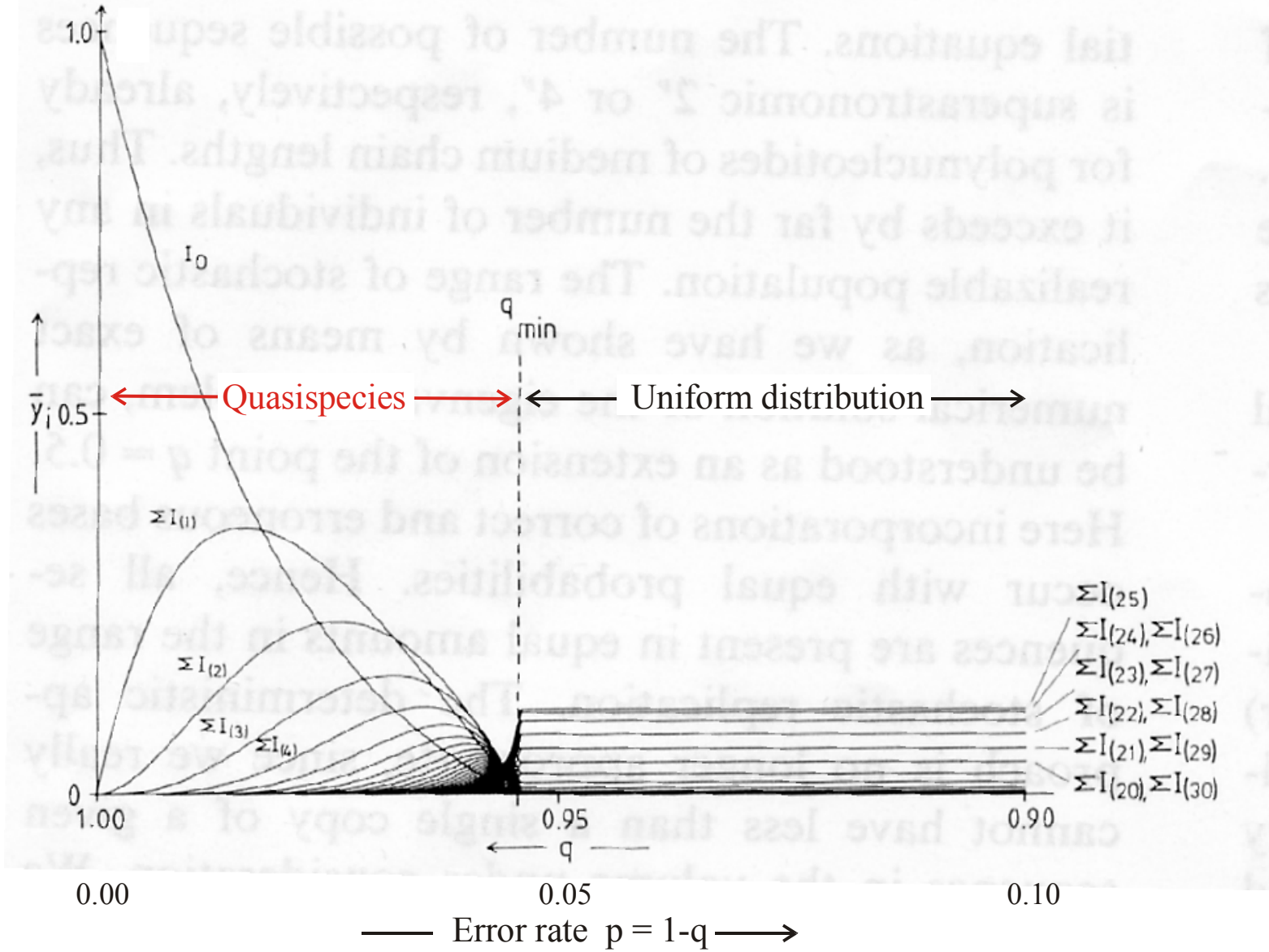
Formation of a quasispecies in sequence space



Formation of a quasispecies in sequence space



Uniform distribution in sequence space



Quasispecies as a function of the replication accuracy q

Chain length and error threshold

$$Q \cdot \sigma \;=\; (1-p)^n \cdot \sigma \;\geq\; 1 \quad\Rightarrow\quad n \cdot \ln(1-p) \geq -\ln\sigma$$

p … constant : $n_{\max} \approx \dfrac{\ln\sigma}{p}$

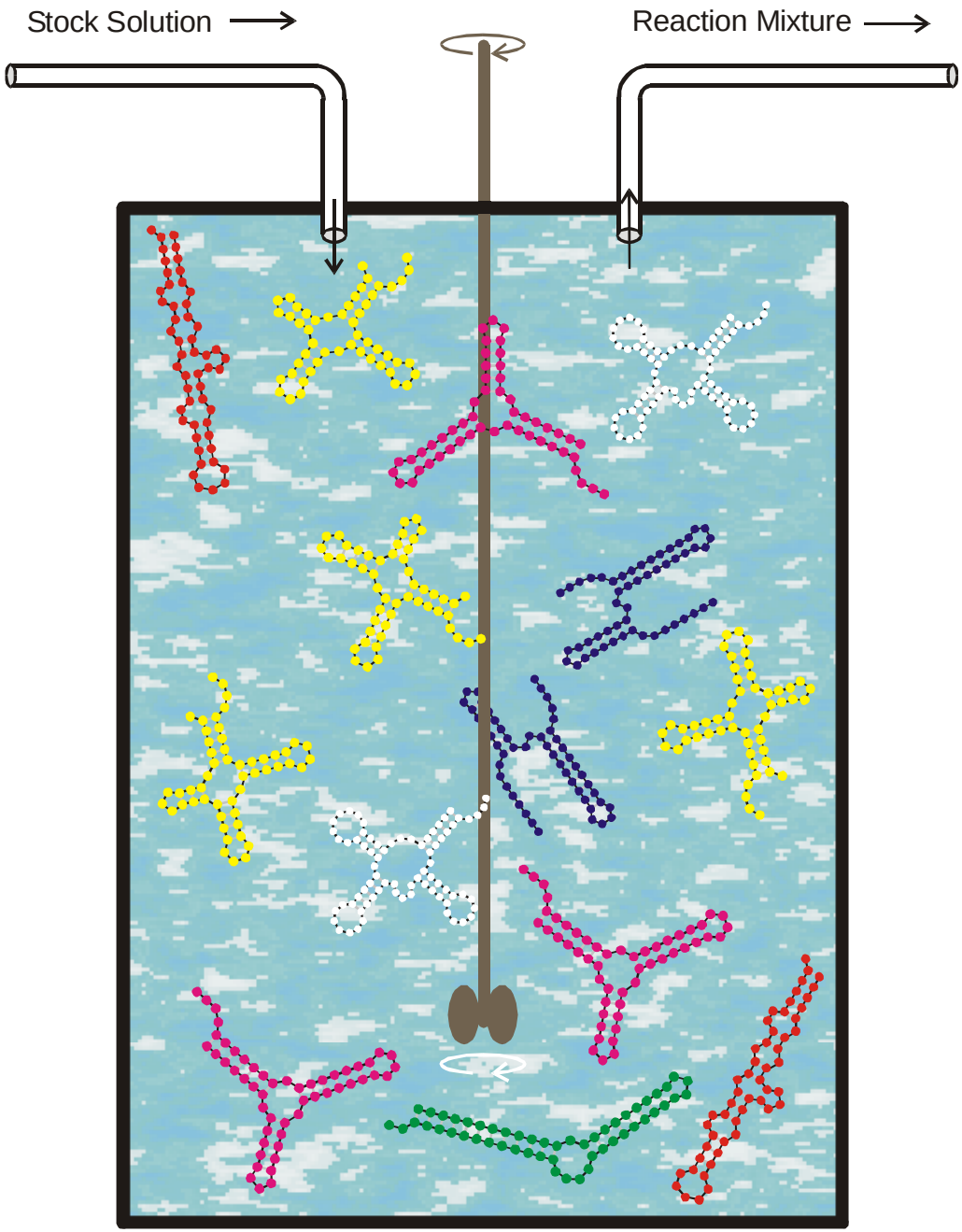
n … constant : $p_{\max} \approx \dfrac{\ln\sigma}{n}$

$Q = (1-p)^n$ … replication accuracy

p … error rate

n … chain length

$\sigma = \dfrac{f_m}{\sum_{j \neq m} f_j}$ … superiority of master sequence



Replication rate constant:

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

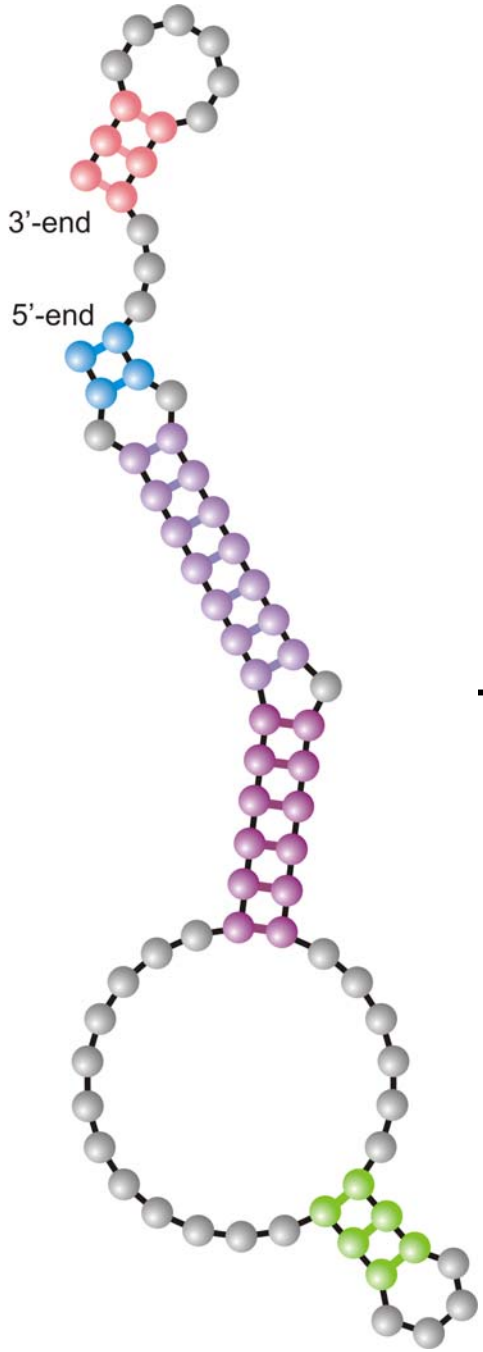
$$\Delta d_S{}^{(k)} = d_H(S_k, S_\tau)$$

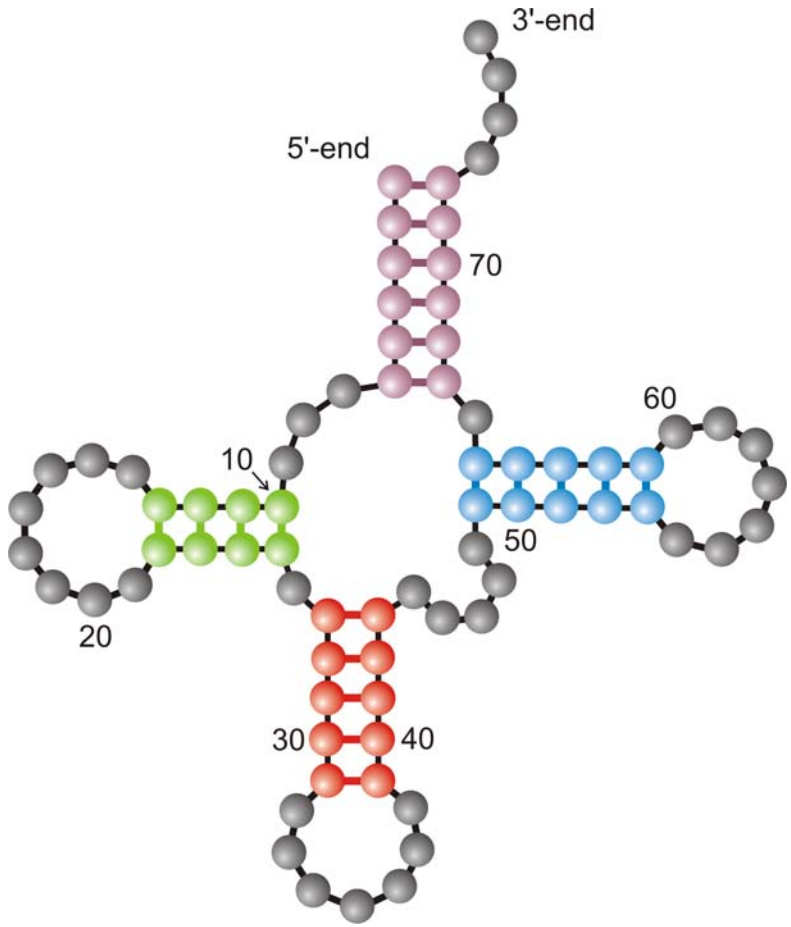
Selection constraint:

RNA molecules is controlled by the flow

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

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





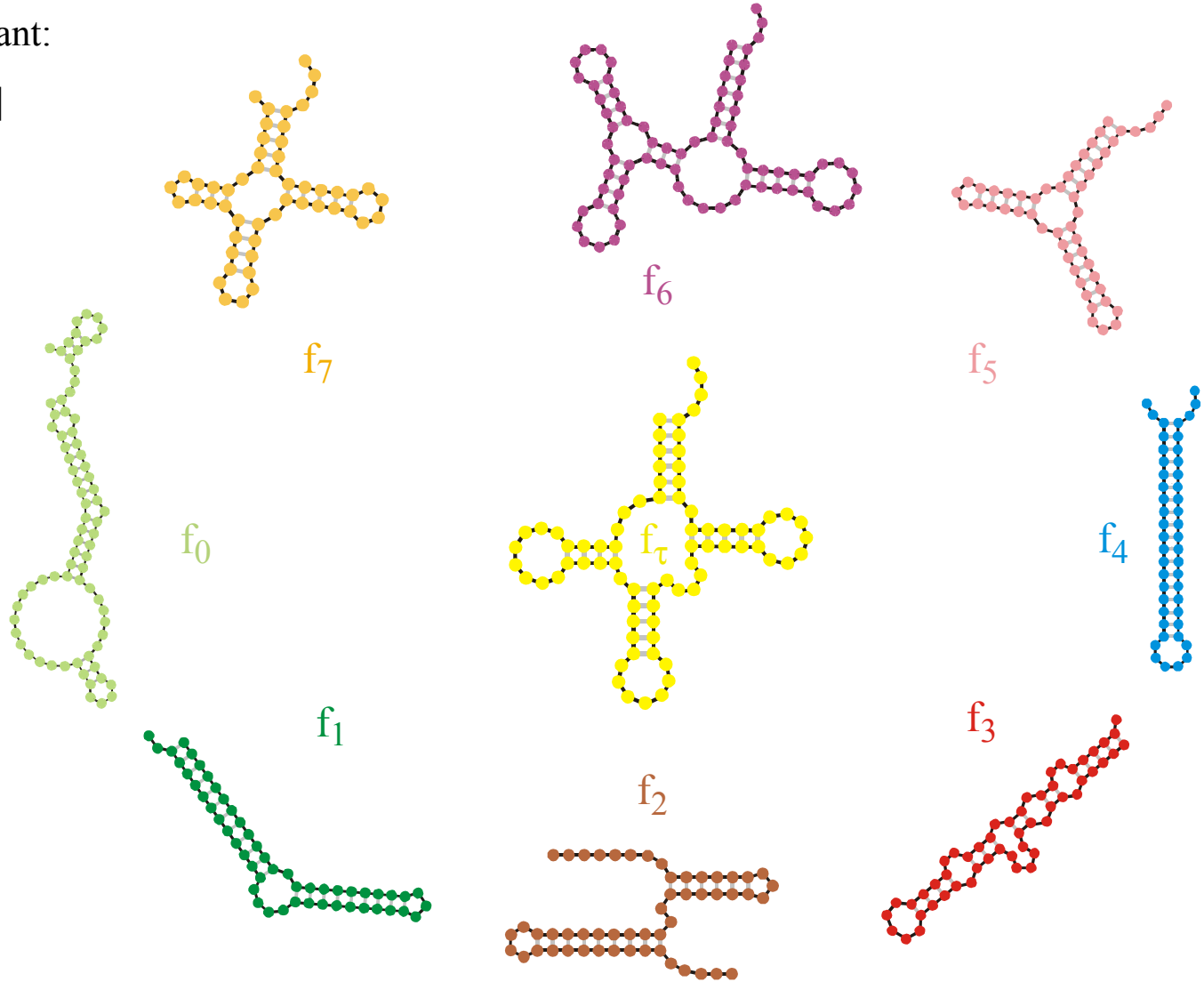
Randomly chosen initial structure

Phenylalanyl-tRNA as target structure

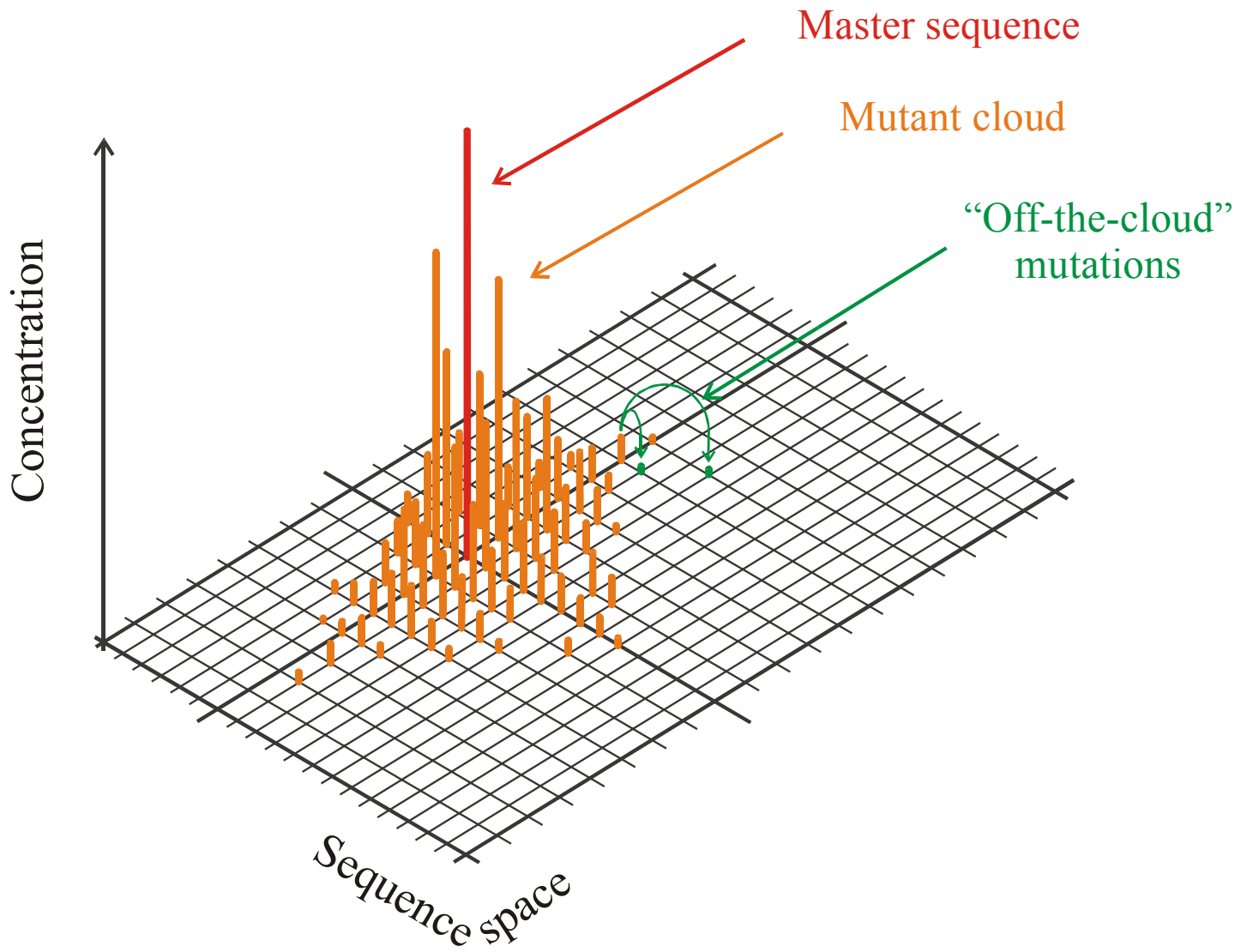
Replication rate constant:

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

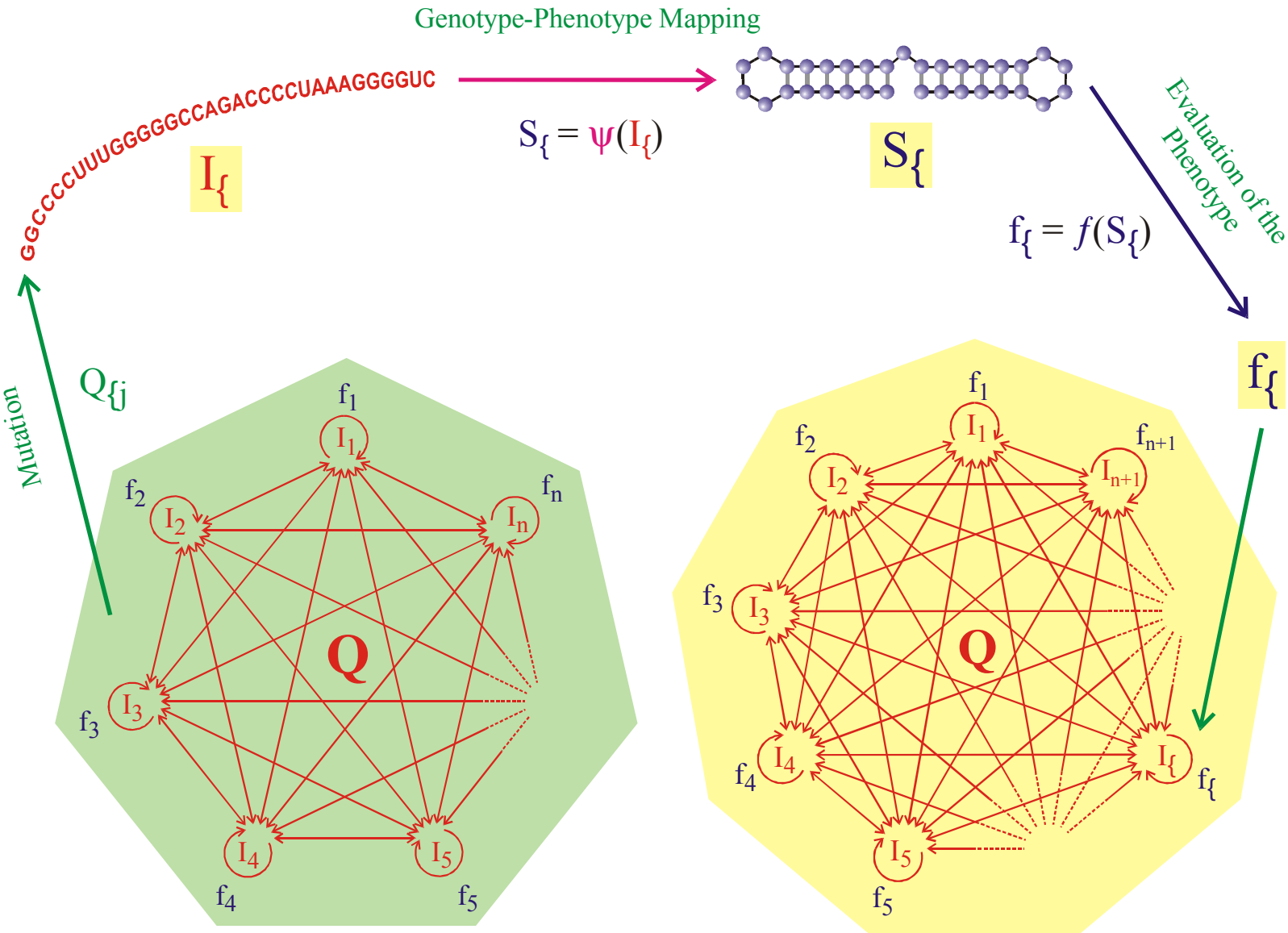
$$\Delta d_S^{(k)} = d_H(S_k, S_\tau)$$



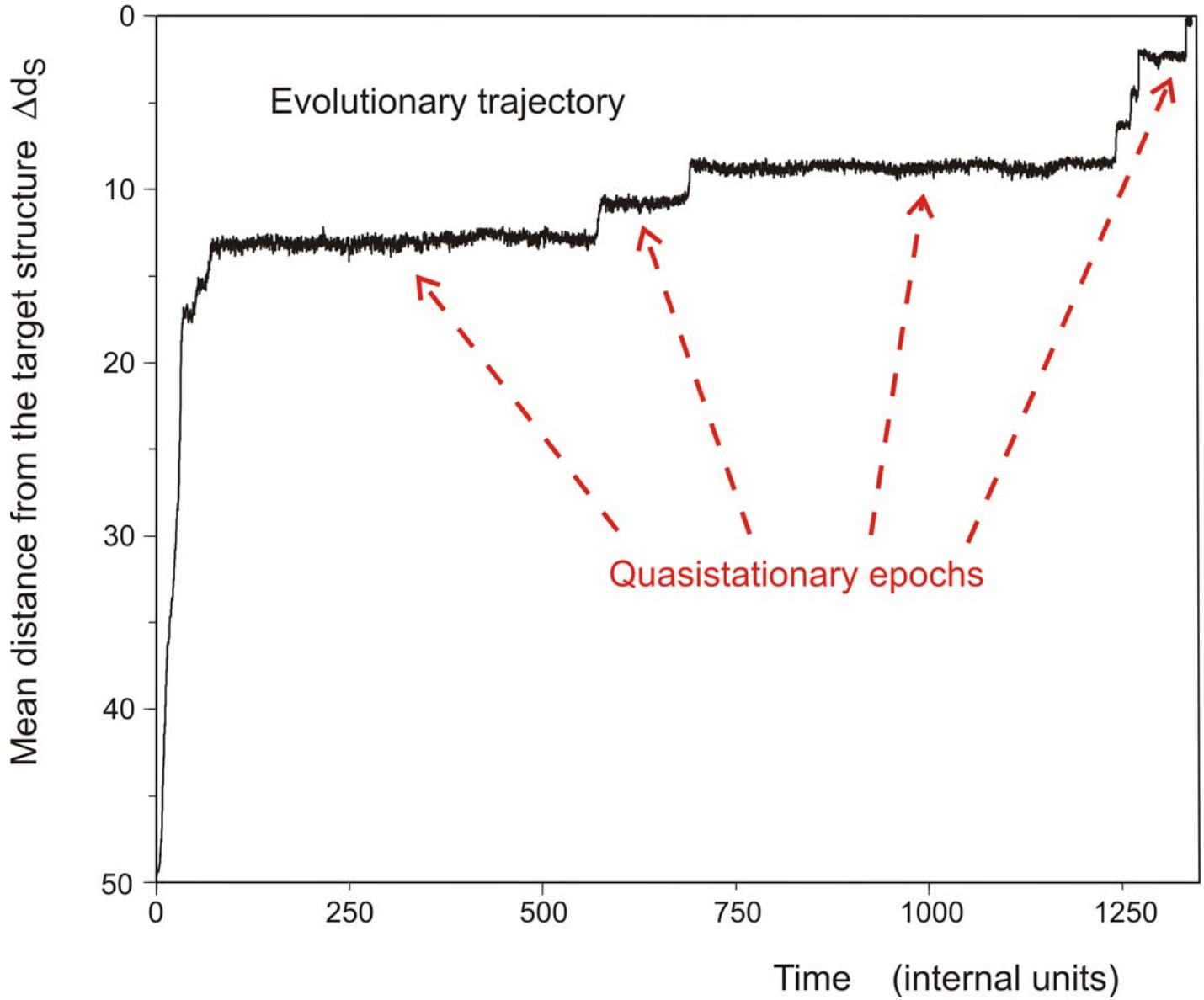
Evaluation of RNA secondary structures yields replication rate constants



The molecular quasispecies in sequence space

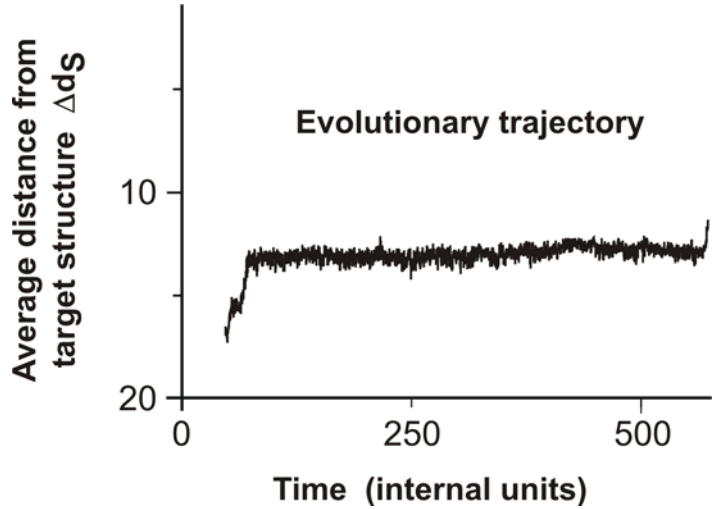


Evolutionary dynamics
including molecular phenotypes



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 change the molecular structure

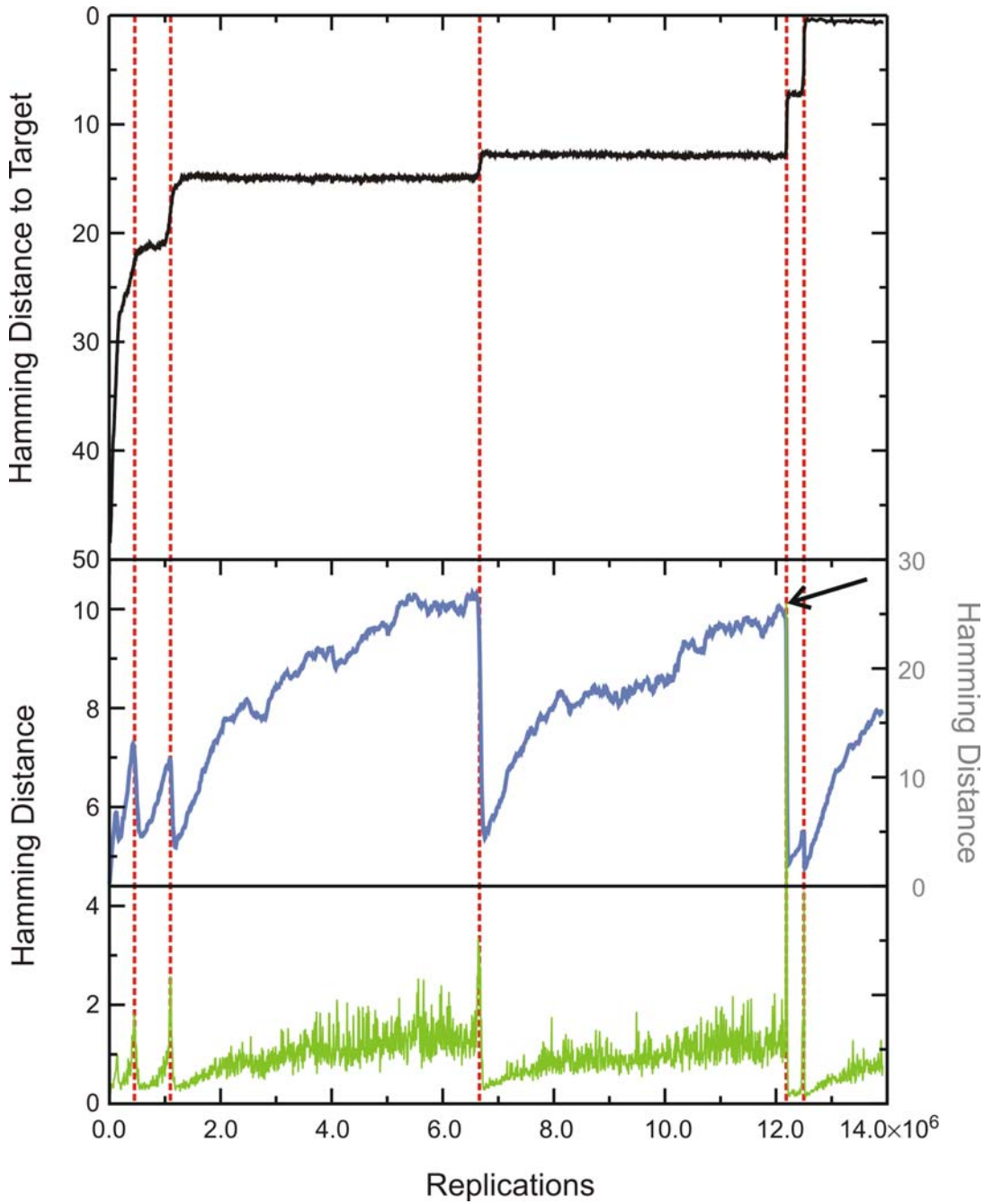
Neutral point mutations leave the molecular structure unchanged

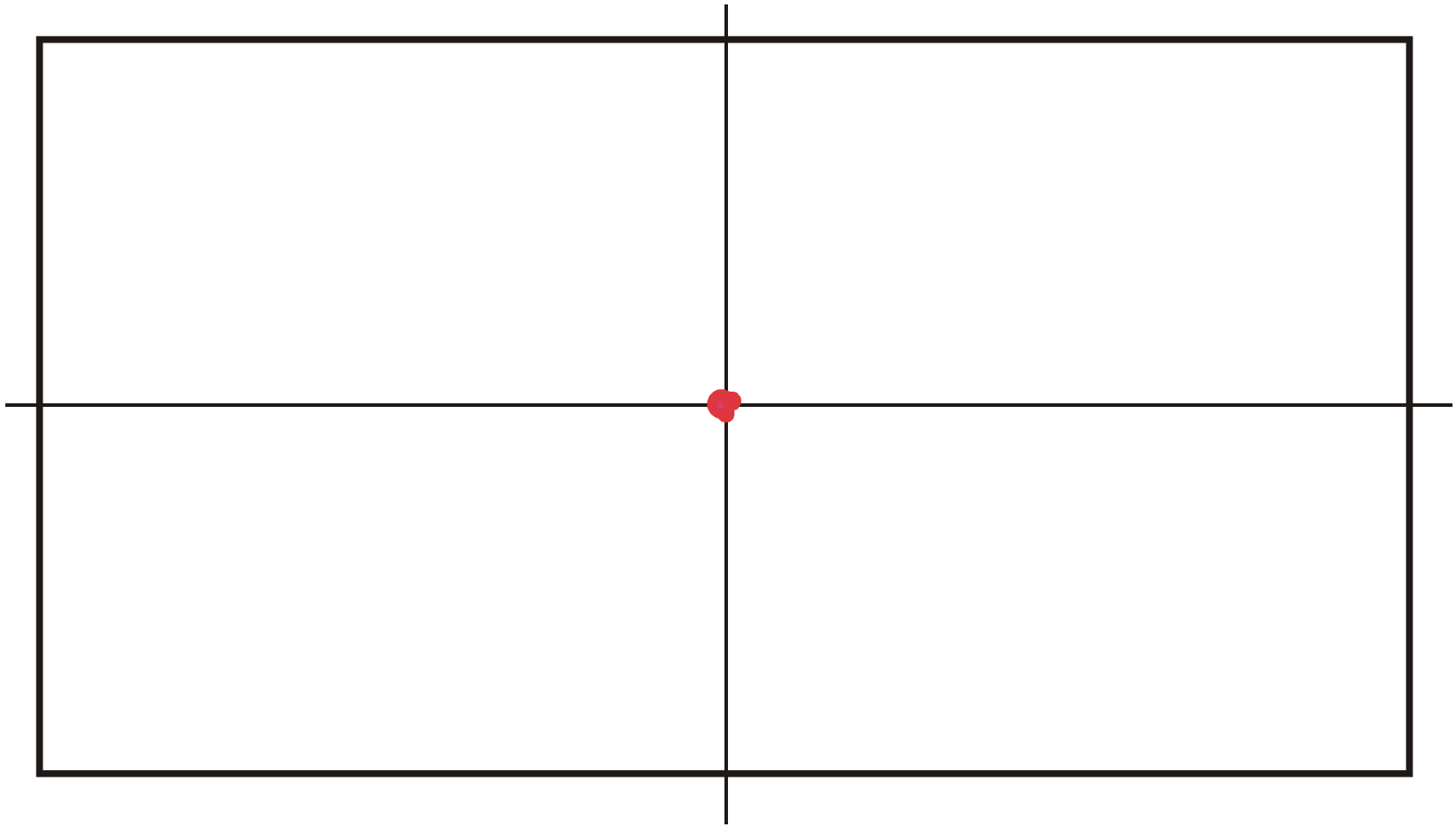
Neutral genotype evolution during phenotypic stasis

Evolutionary trajectory

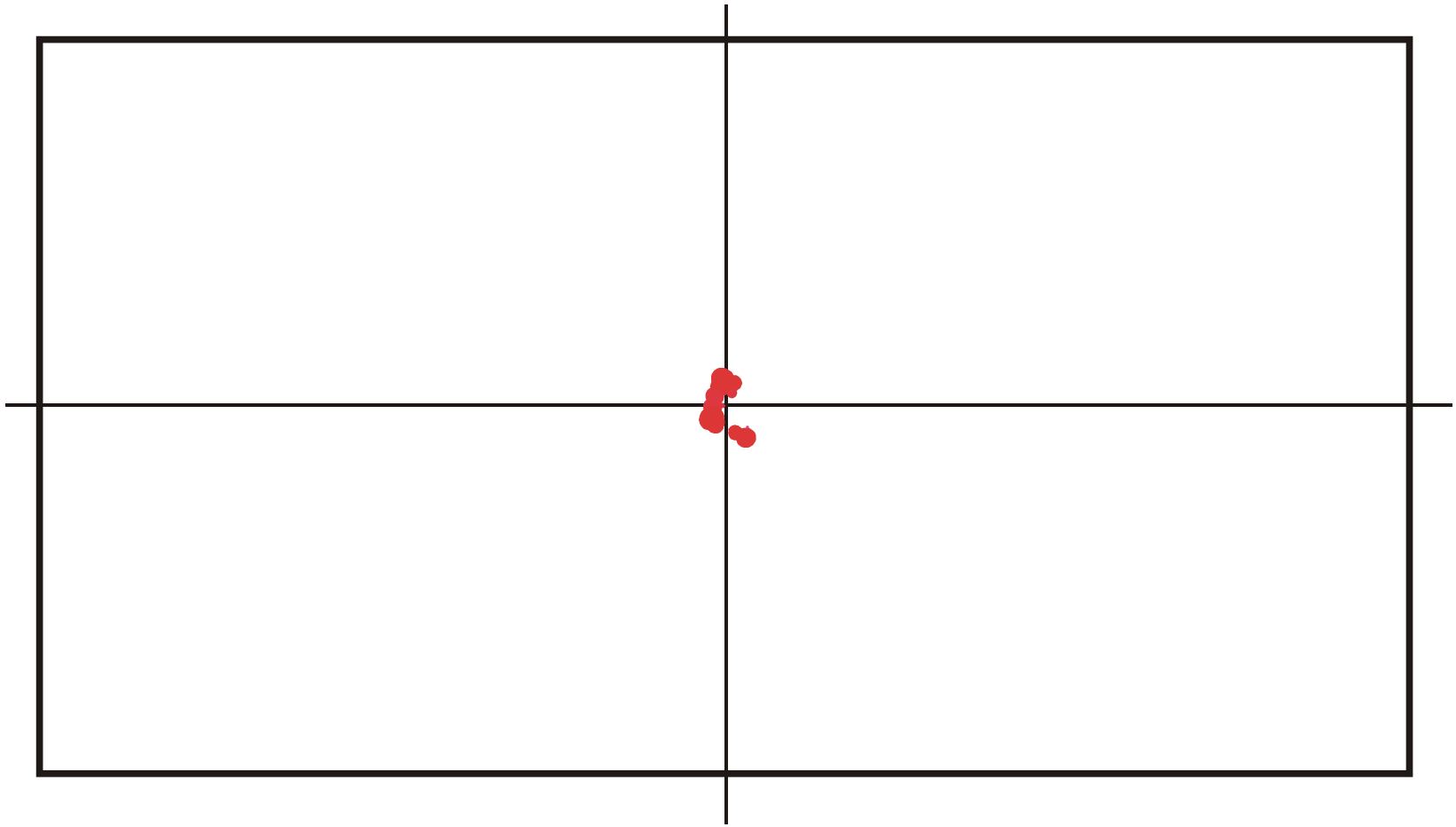
Spreading of the population on neutral networks

Drift of the population center in sequence space

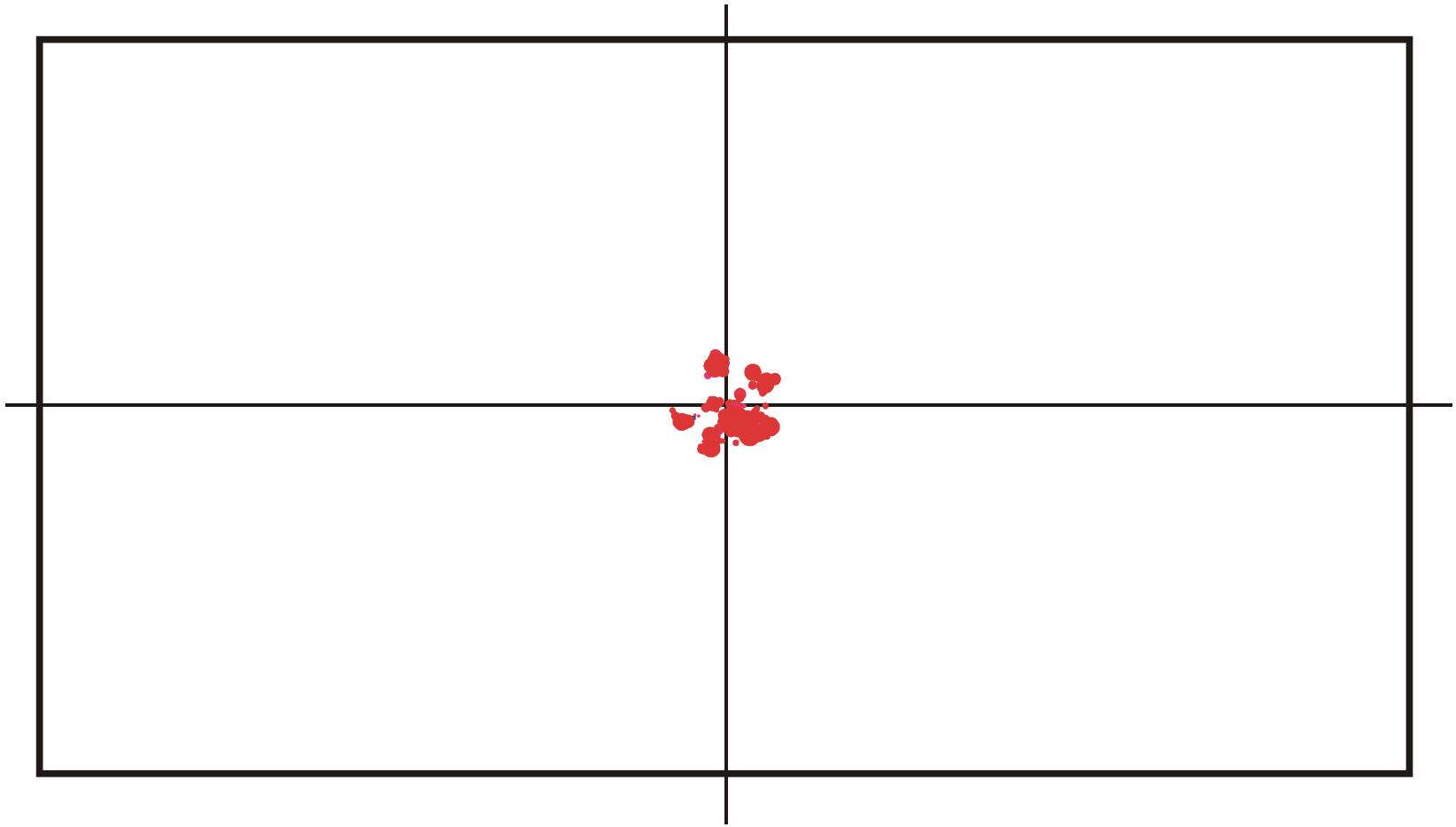




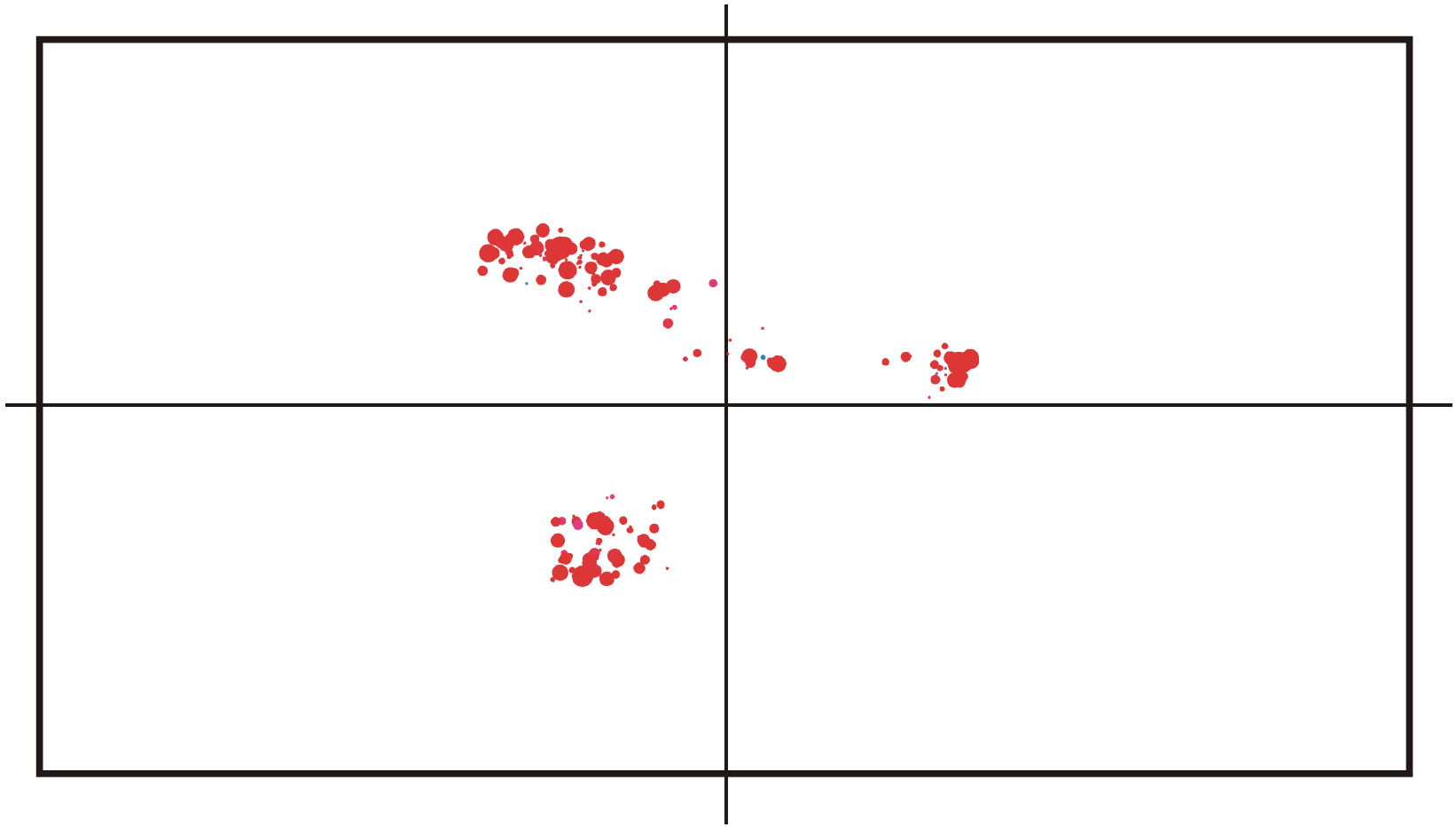
Spreading and evolution of a population on a neutral network: t = 150



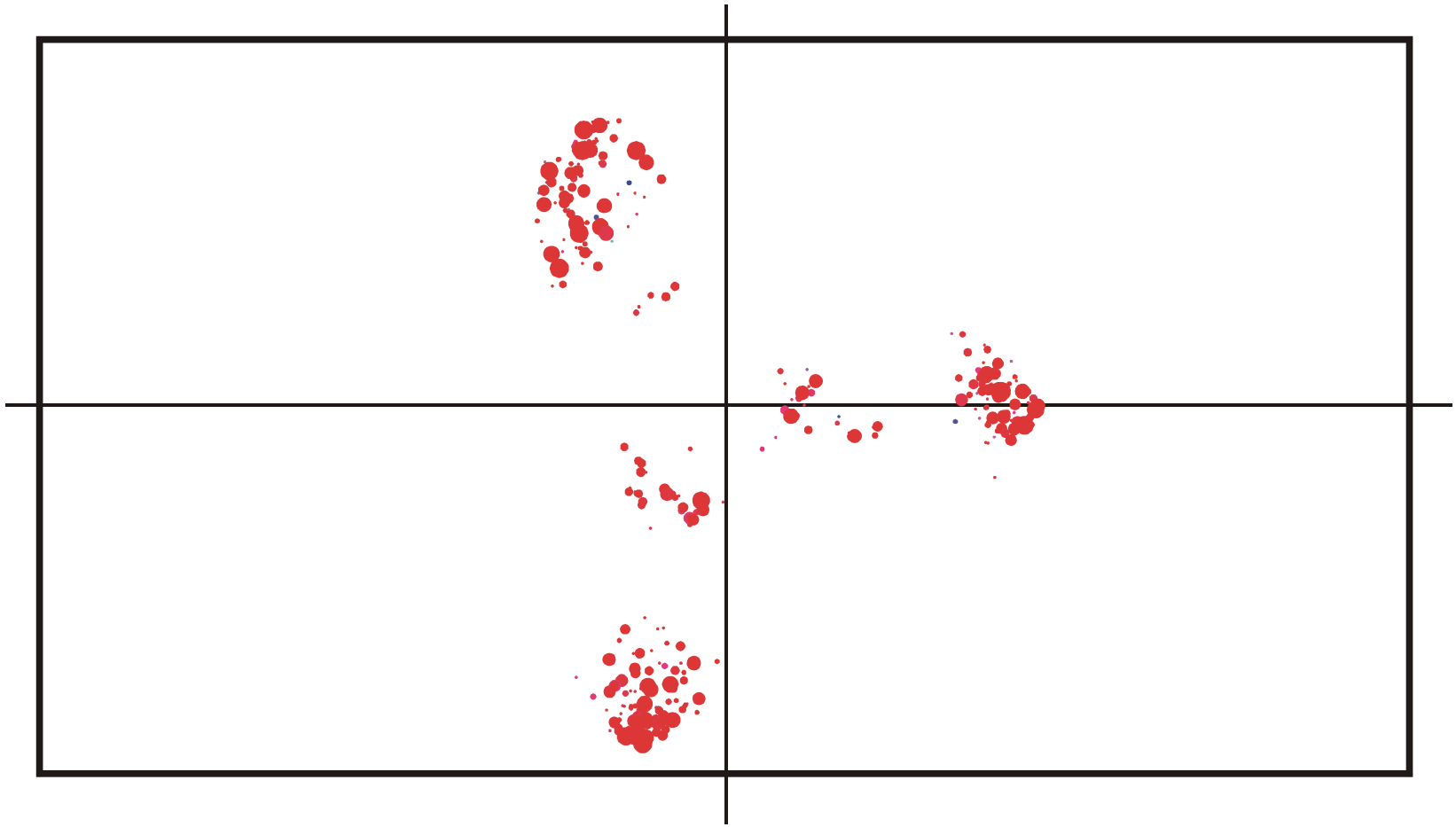
Spreading and evolution of a population on a neutral network : t = 170



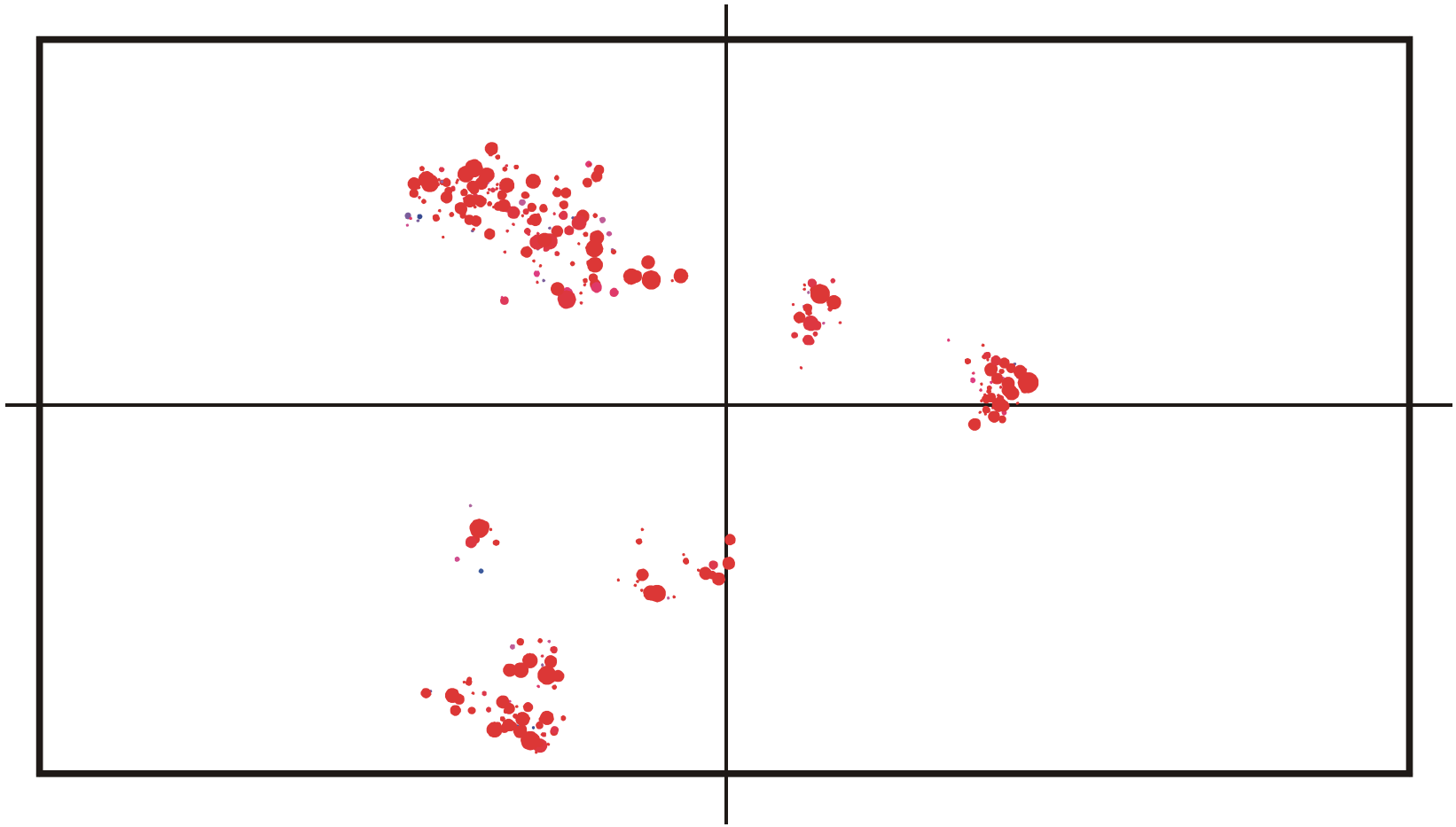
Spreading and evolution of a population on a neutral network : t = 200



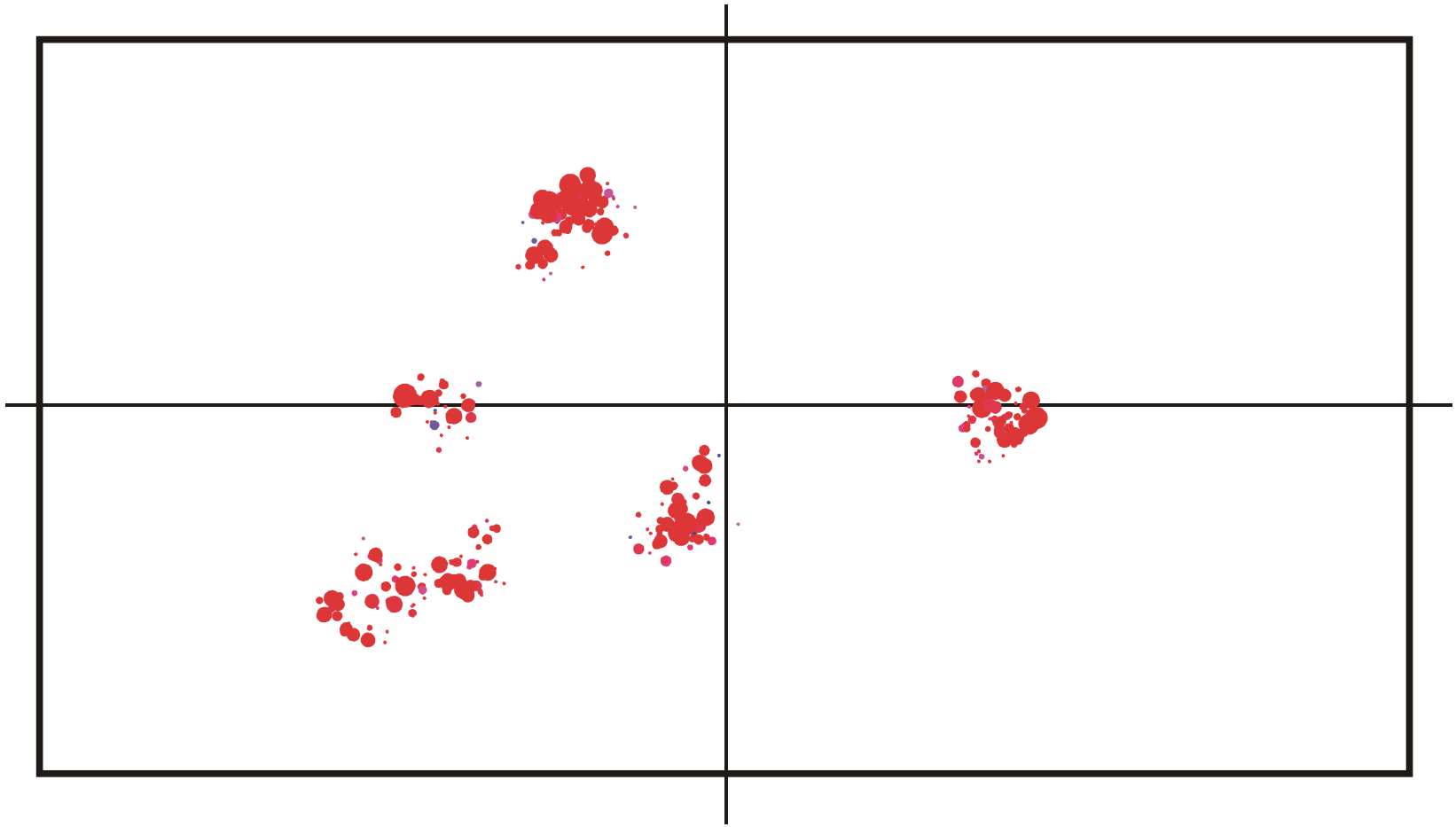
Spreading and evolution of a population on a neutral network : t = 350



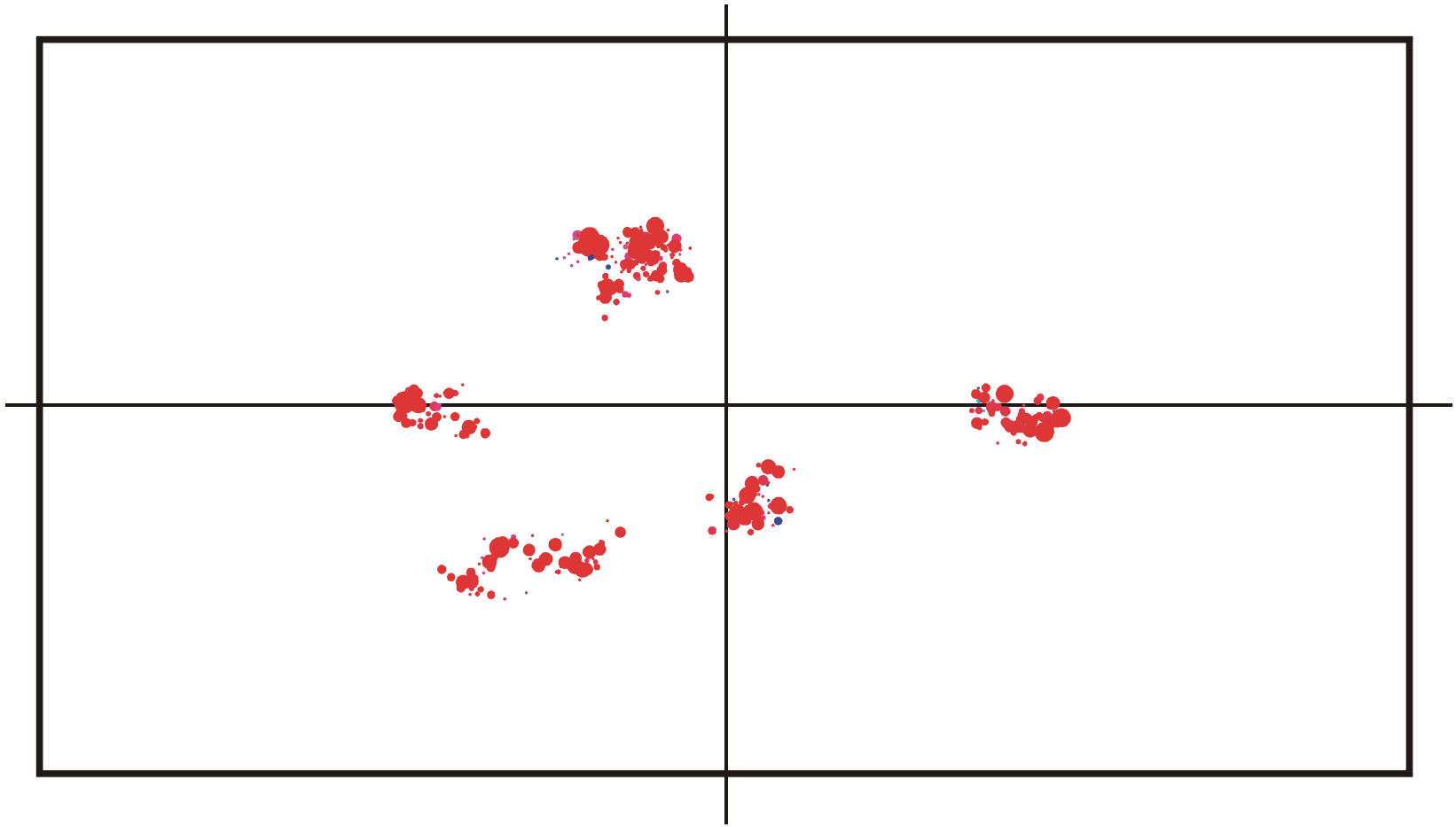
Spreading and evolution of a population on a neutral network : t = 500



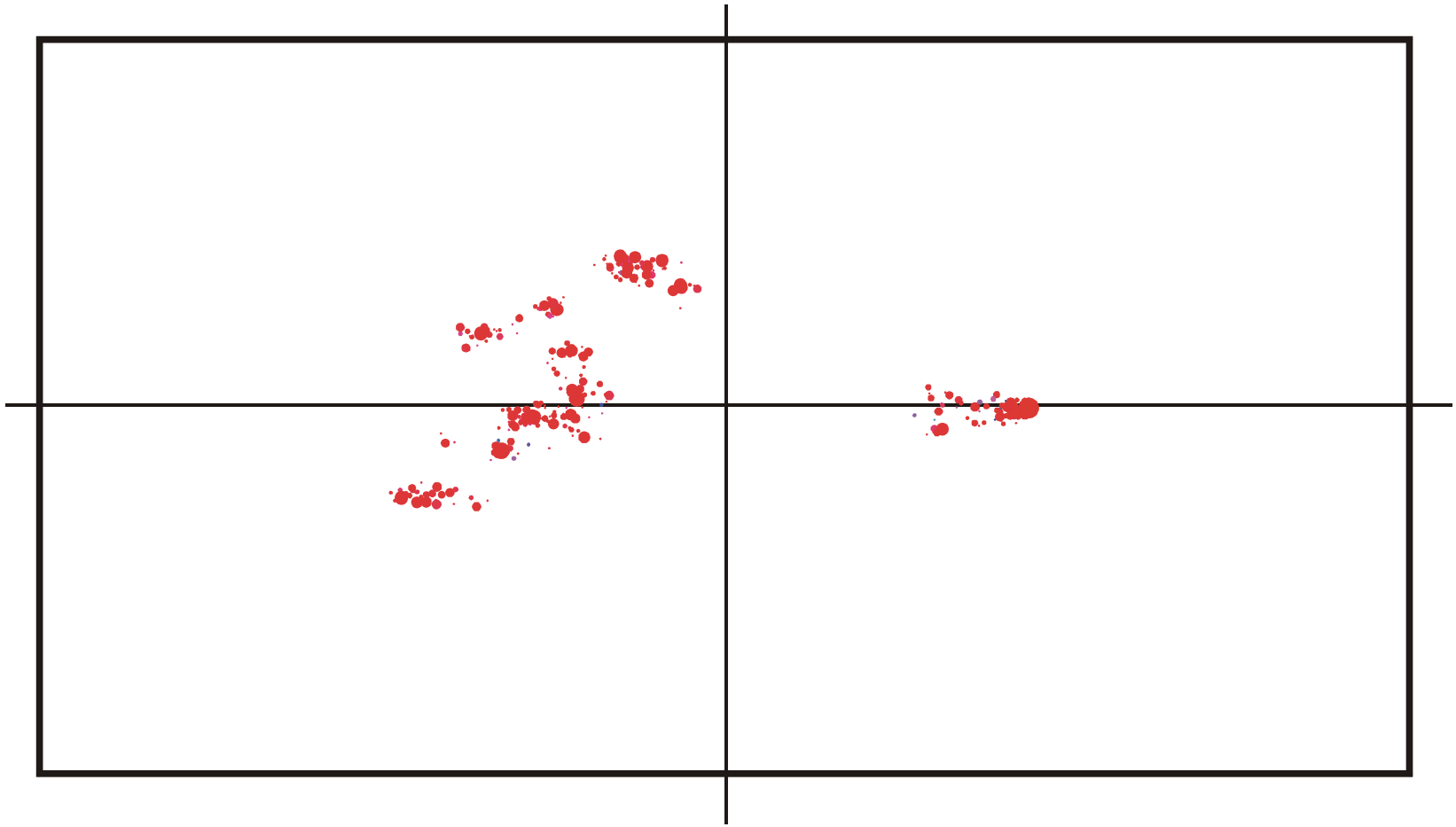
Spreading and evolution of a population on a neutral network : t = 650



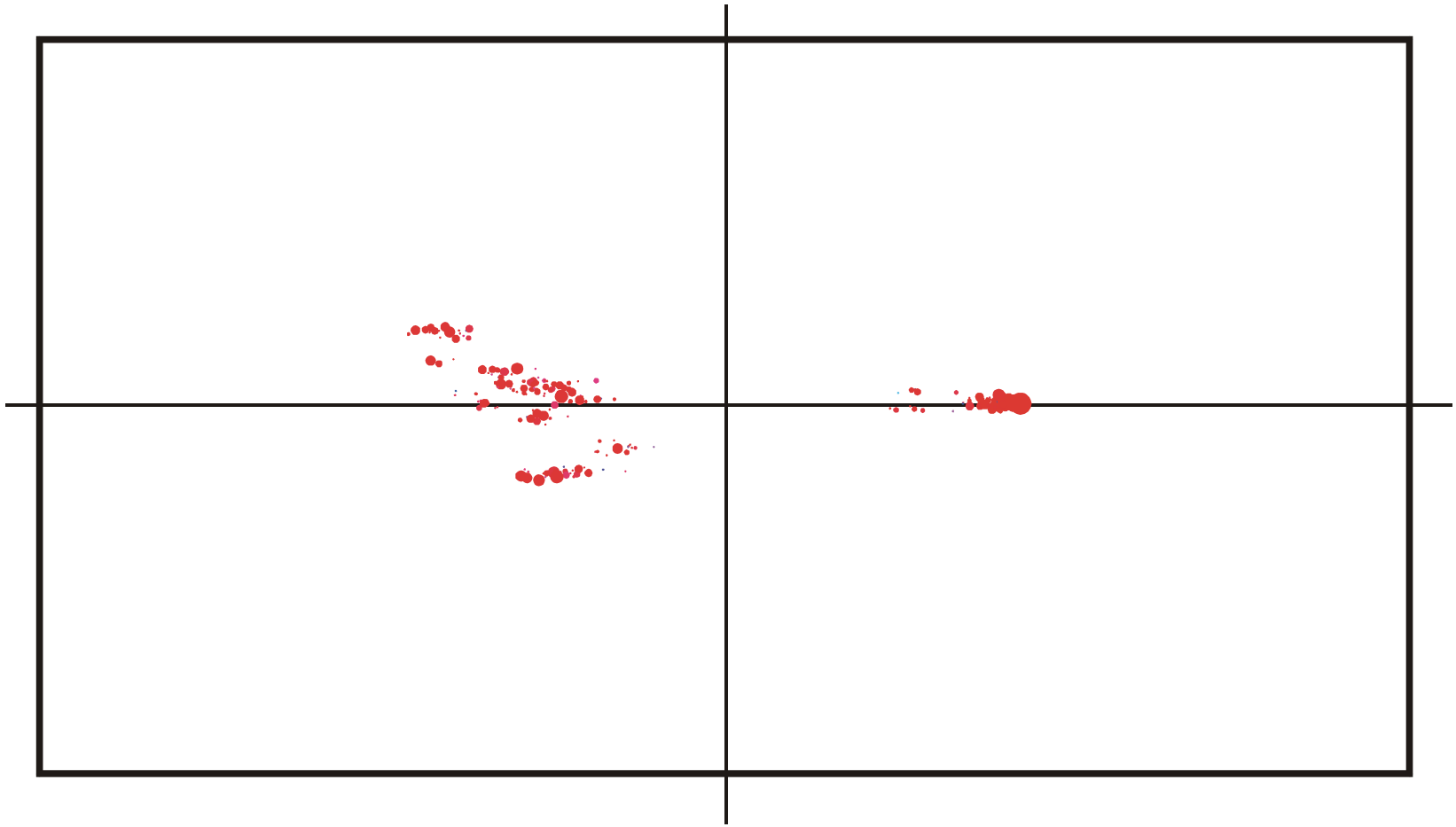
Spreading and evolution of a population on a neutral network : t = 820



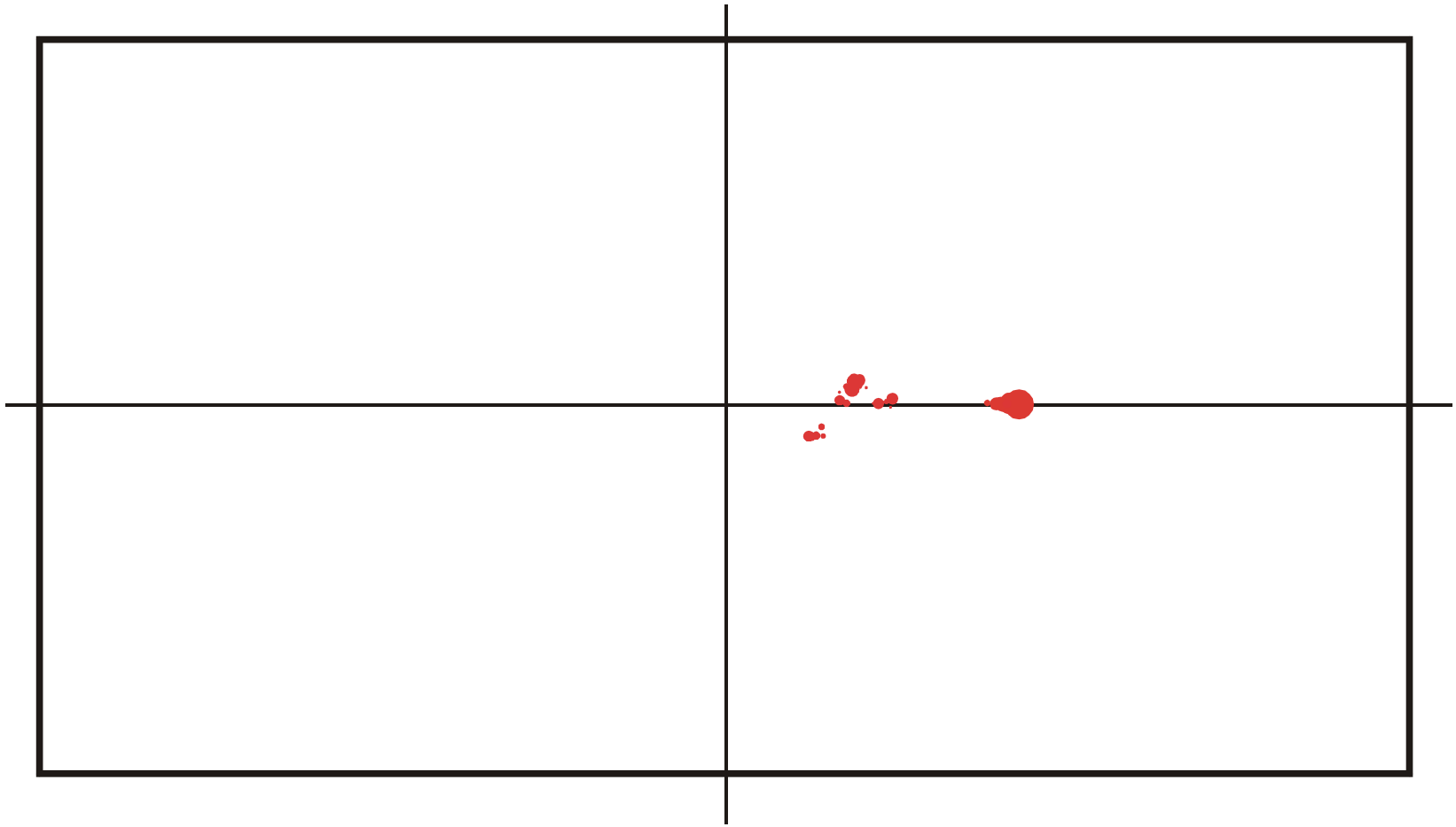
Spreading and evolution of a population on a neutral network : t = 825



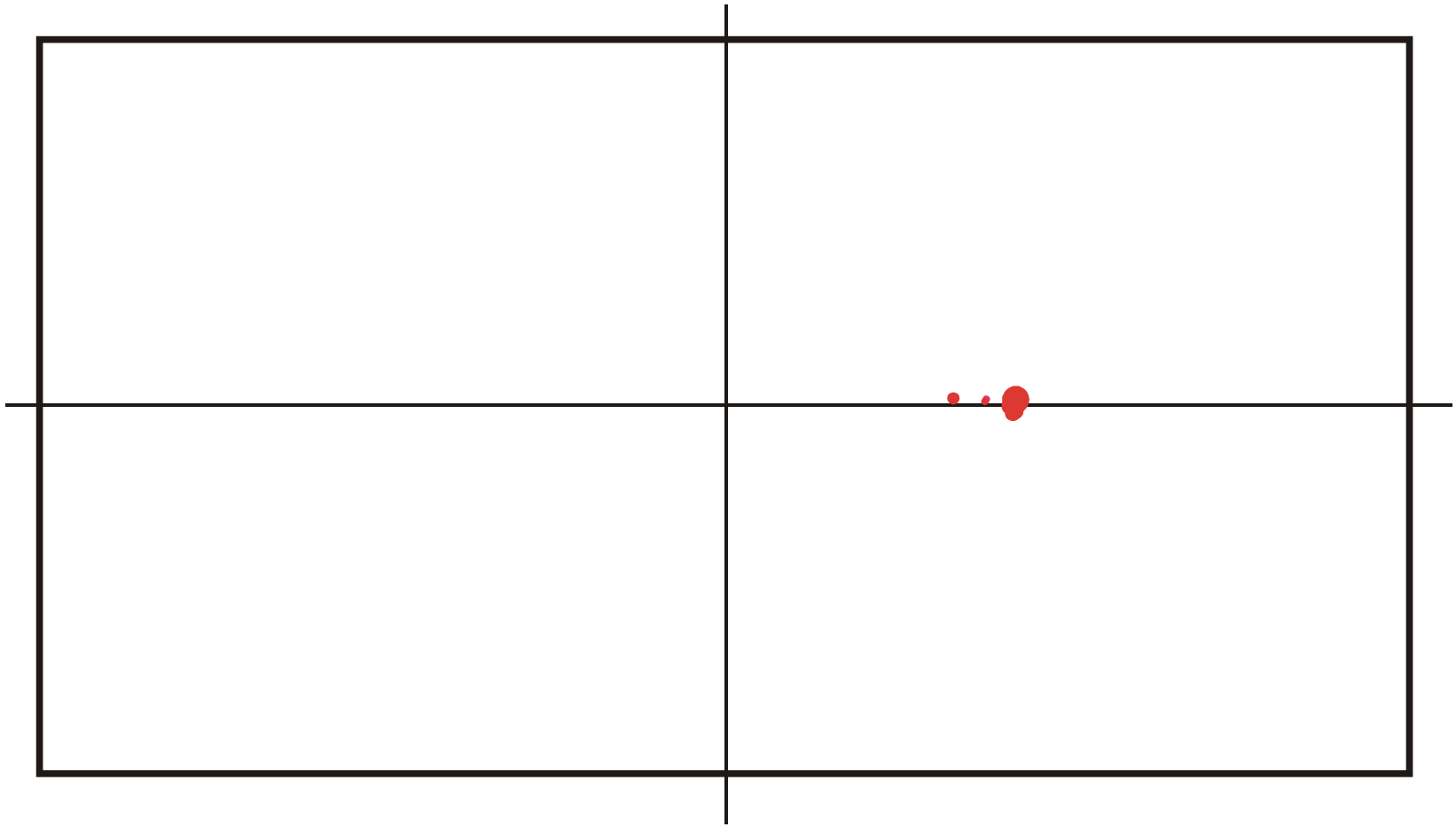
Spreading and evolution of a population on a neutral network : t = 830



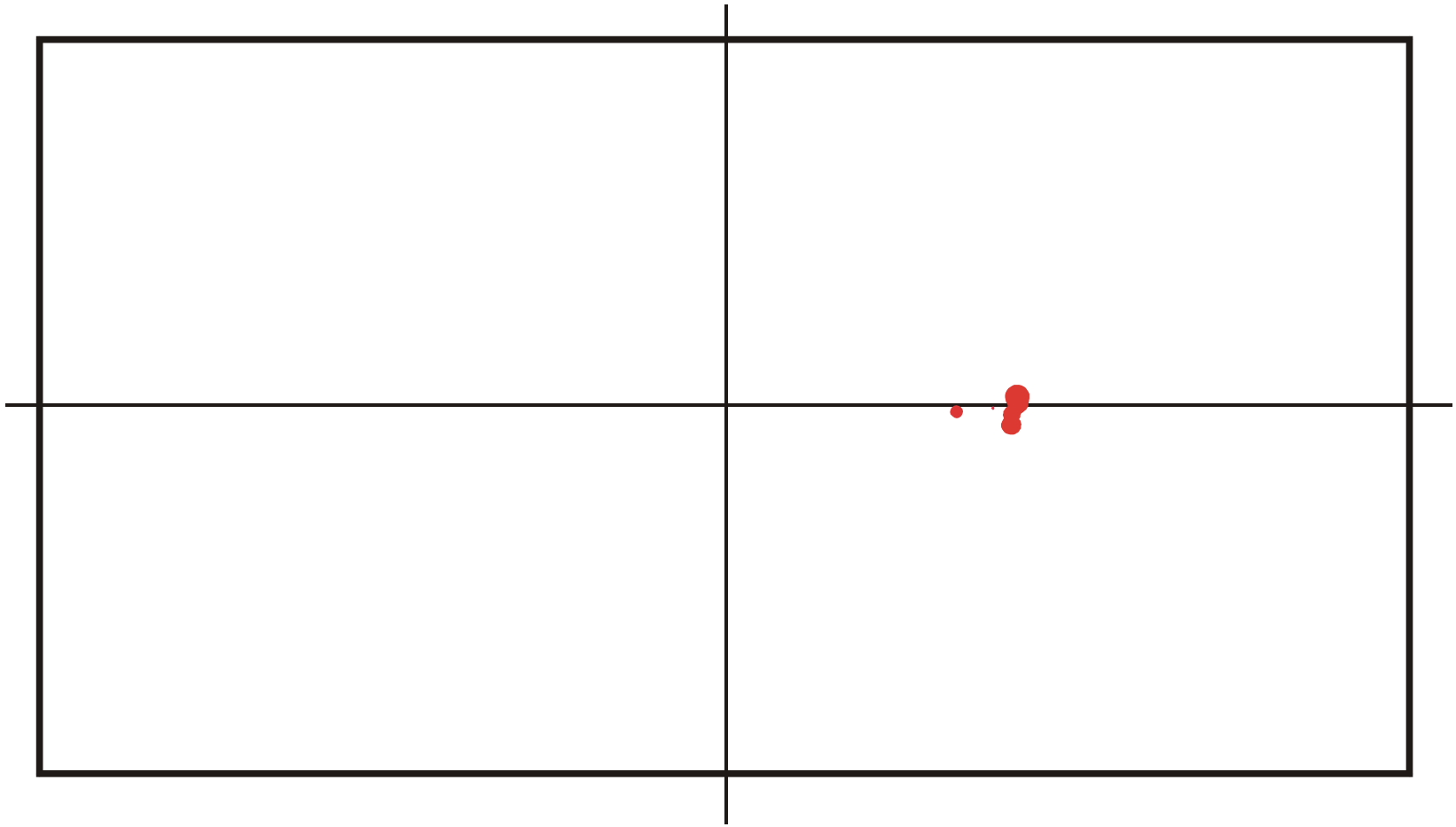
Spreading and evolution of a population on a neutral network : t = 835



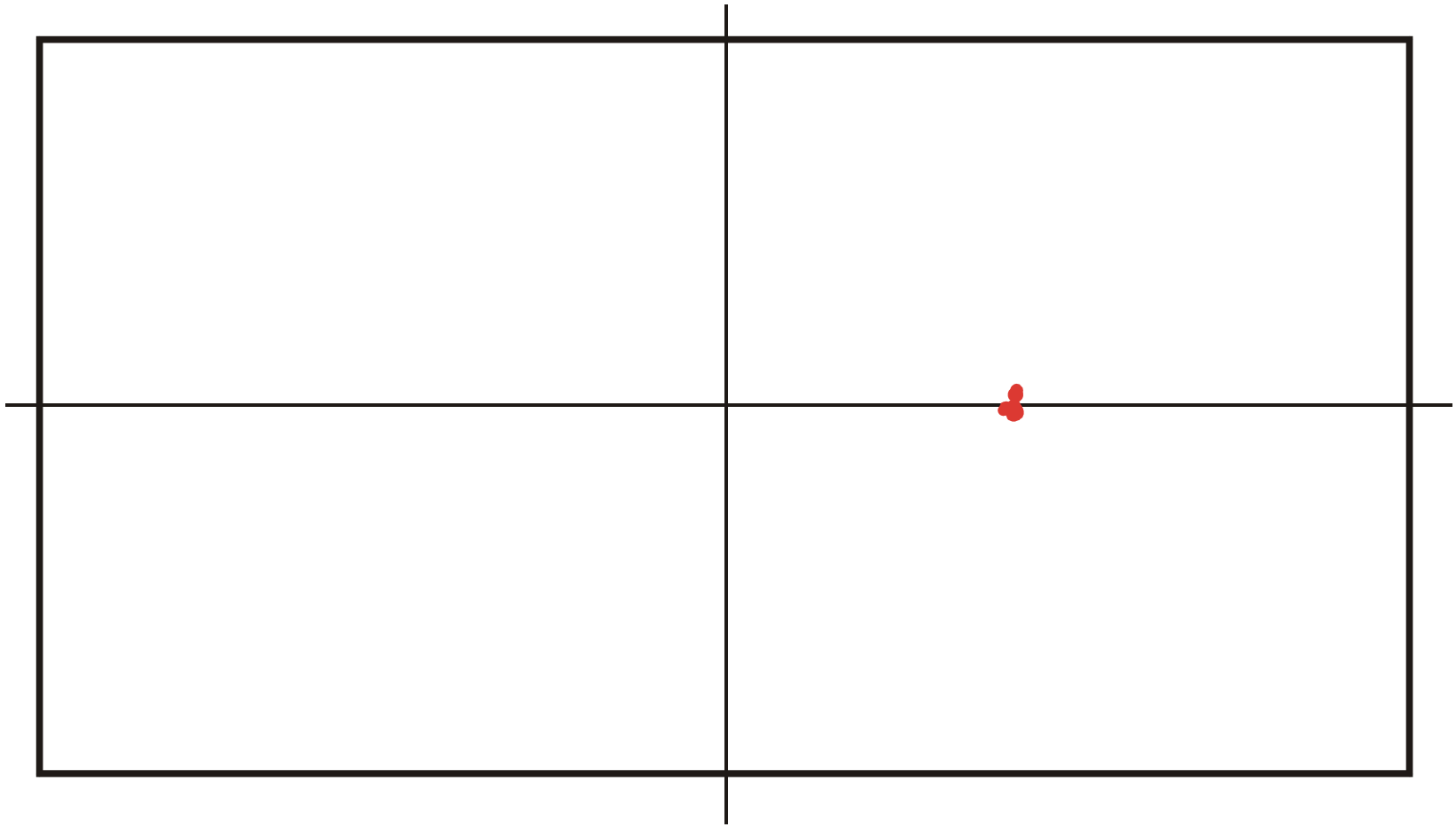
Spreading and evolution of a population on a neutral network : t = 840



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



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



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

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

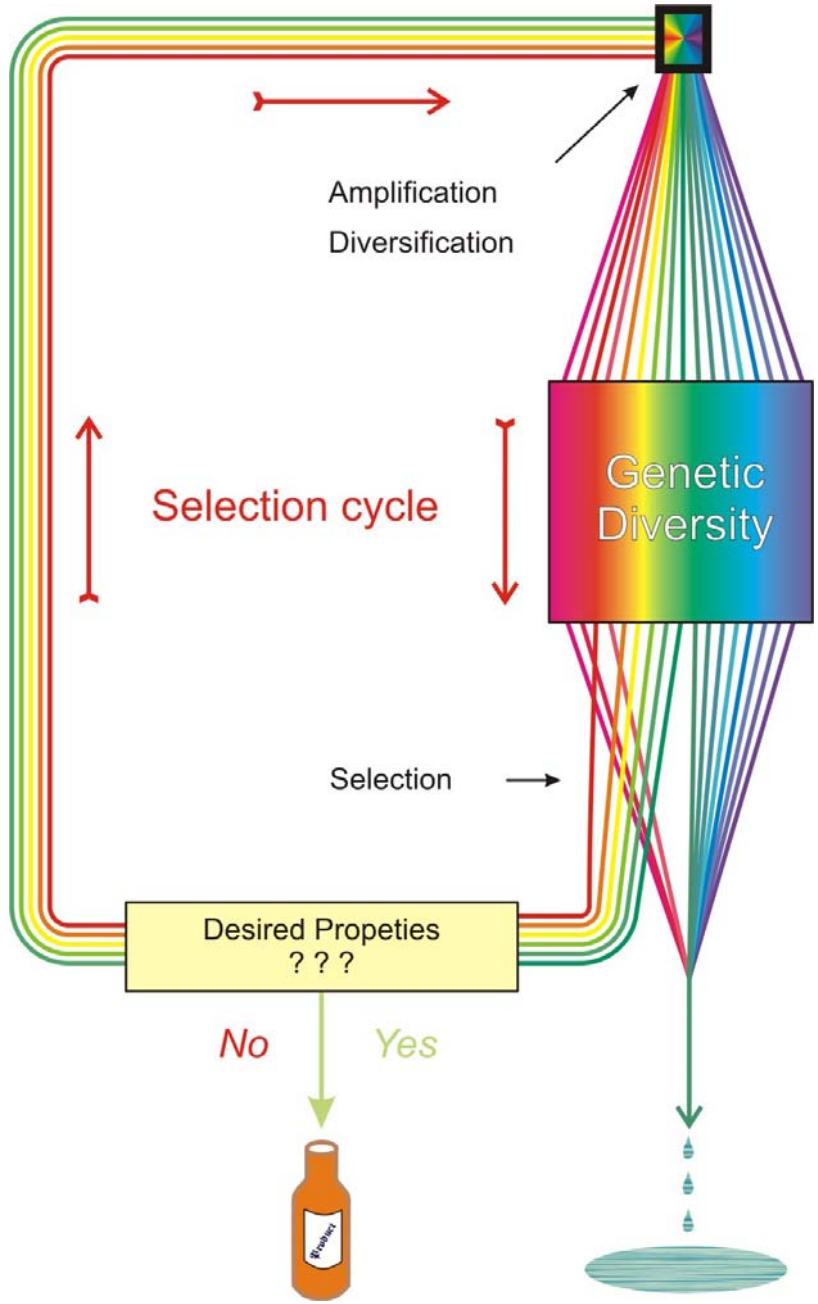
D.P.Bartel, J.W.Szostak, ***Isolation of new ribozymes from a large pool of random sequences***. Science **261** (1993), 1411-1418

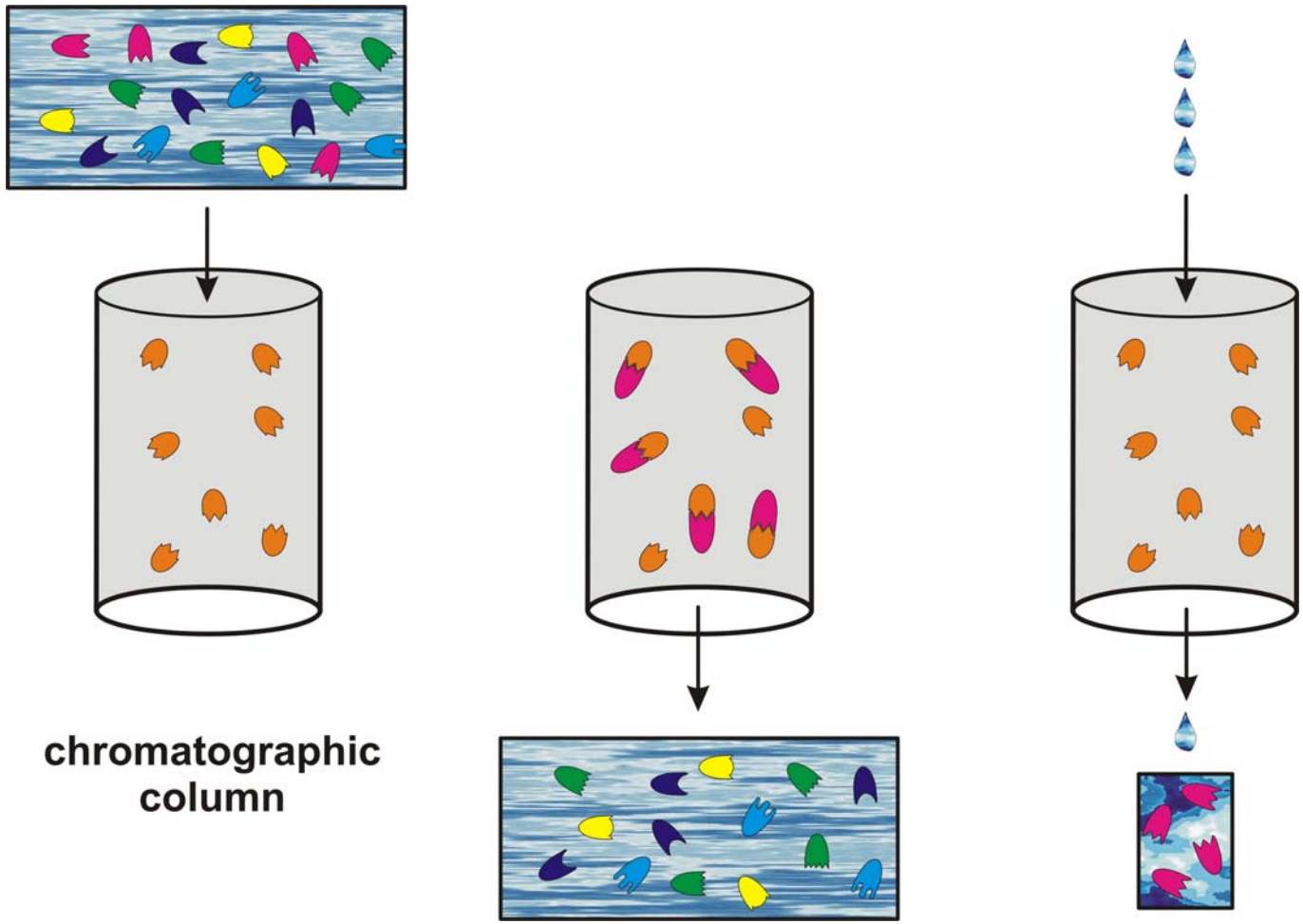
R.D.Jenison, S.C.Gill, A.Pardi, B.Poliski, ***High-resolution molecular discrimination by RNA***. Science **263** (1994), 1425-1429

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

An example of ‘artificial selection’ with RNA molecules or ‘breeding’ of biomolecules





The SELEX technique for the evolutionary preparation of aptamers

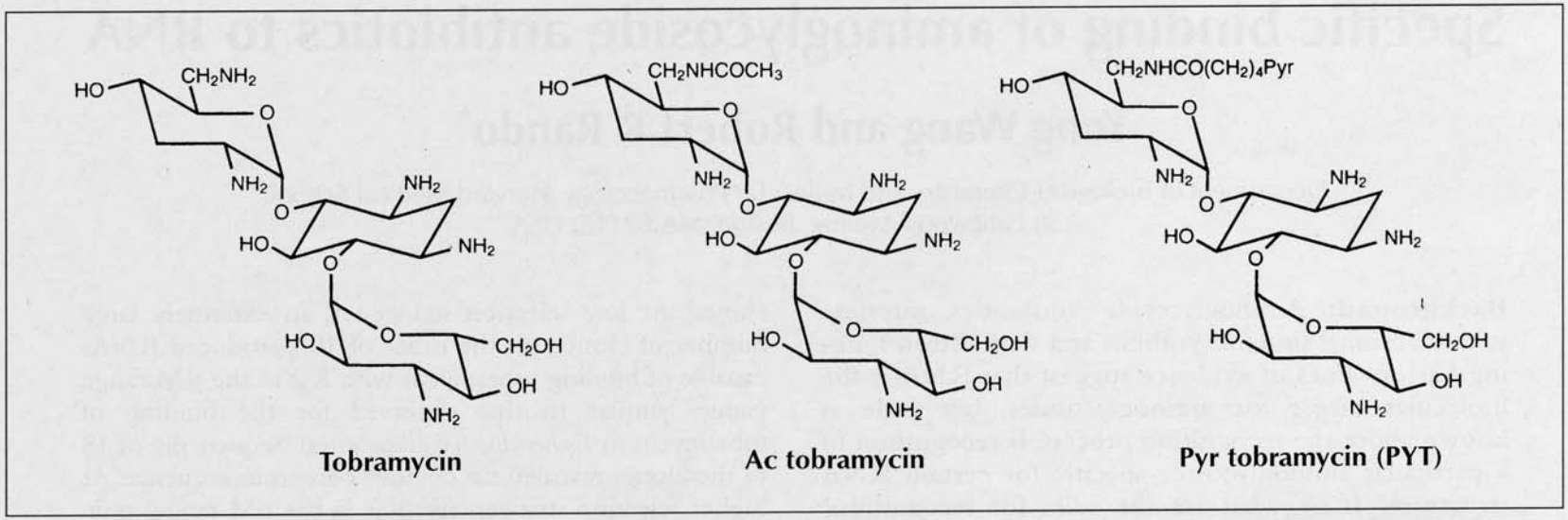
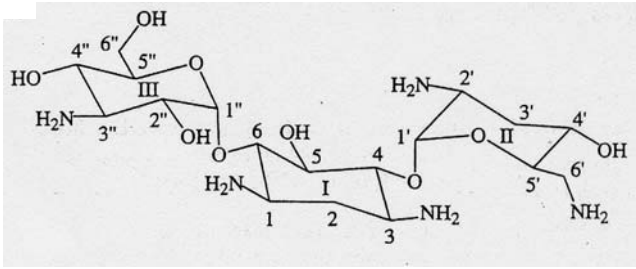


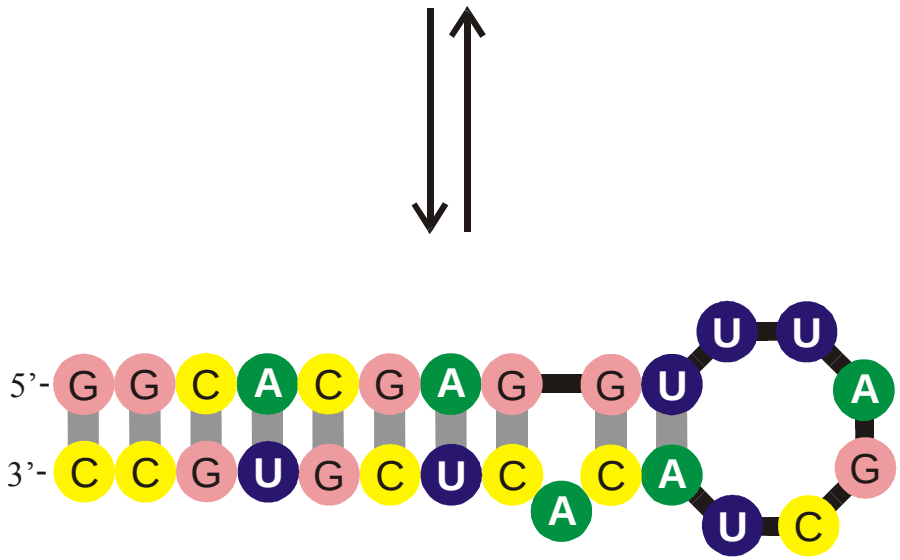
Fig. 1. Structures of tobramycin and analogs used in these studies.

Aptamer binding to aminoglycosid antibiotics: Structure of ligands

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



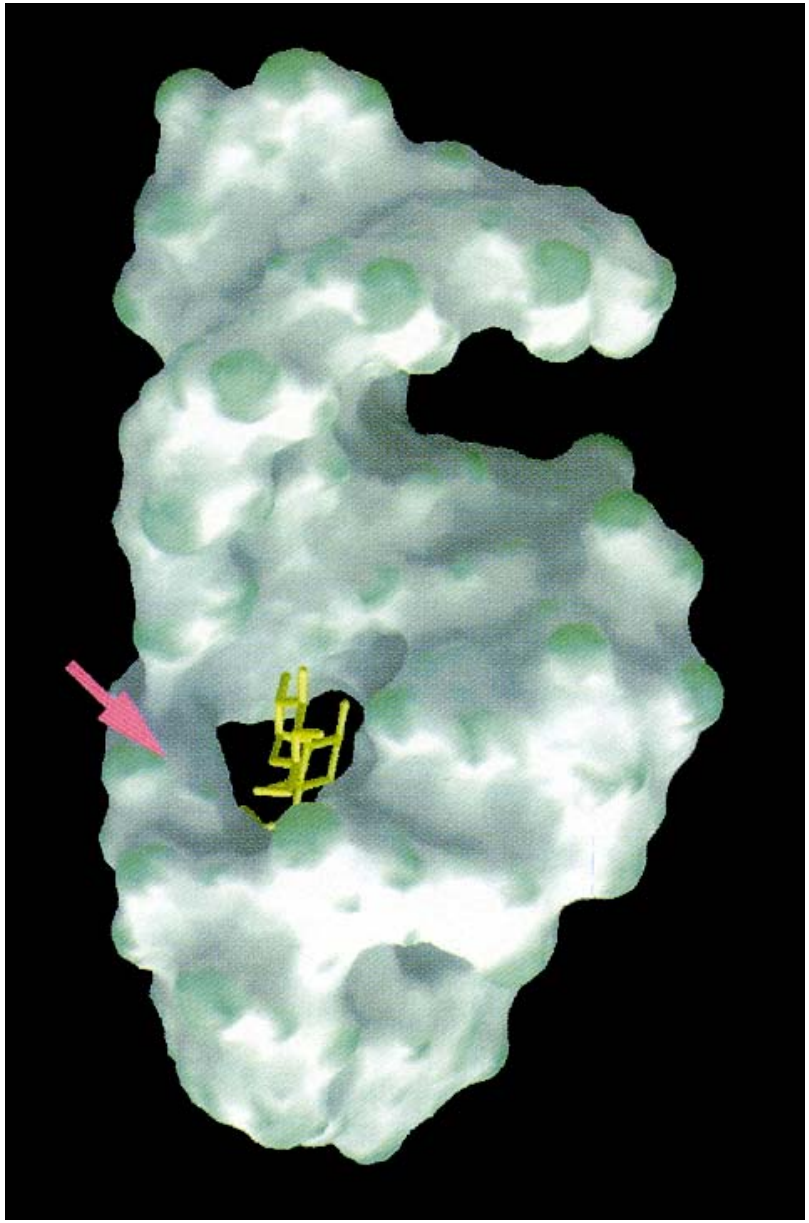
tobramycin



RNA aptamer

Formation of secondary structure of the tobramycin binding RNA aptamer with $K_D = 9$ nM

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)

Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)
Projects No. 09942, 10578, 11065, 13093
13887, and 14898

Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF)
Project No. Mat05

Jubiläumsfonds der Österreichischen Nationalbank
Project No. Nat-7813

European Commission: Contracts No. 98-0189, 12835 (NEST)

Austrian Genome Research Program – GEN-AU: Bioinformatics Network (BIN)

Österreichische Akademie der Wissenschaften

Siemens AG, Austria

Universität Wien and the Santa Fe Institute

Universität Wien

Coworkers

Universität Wien

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

Paul E. Phillipson, University of Colorado at Boulder, CO

Heinz Engl, Philipp Kügler, **James Lu**, **Stefan Müller**, RICAM Linz, AT

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

Walter Fontana, Harvard Medical School, MA

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

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

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

Kurt Grünberger, Michael Kospach , **Andreas Wernitznig**, **Stefanie Widder, Stefan Wuchty**, Universität Wien, AT

Jan Cupal, **Stefan Bernhart**, **Lukas Endler, Ulrike Langhammer**, **Rainer Machne, Ulrike Mückstein**, **Hakim Tafer, Thomas Taylor,** Universität Wien, AT

Web-Page for further information:

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