Ursprung des Lebens und Evolution der Moleküle

Peter Schuster

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

Vortragsreihe: „An den Grenzen des Wissens“

Zürich, 07.12.2006

Web-Page for further information:

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

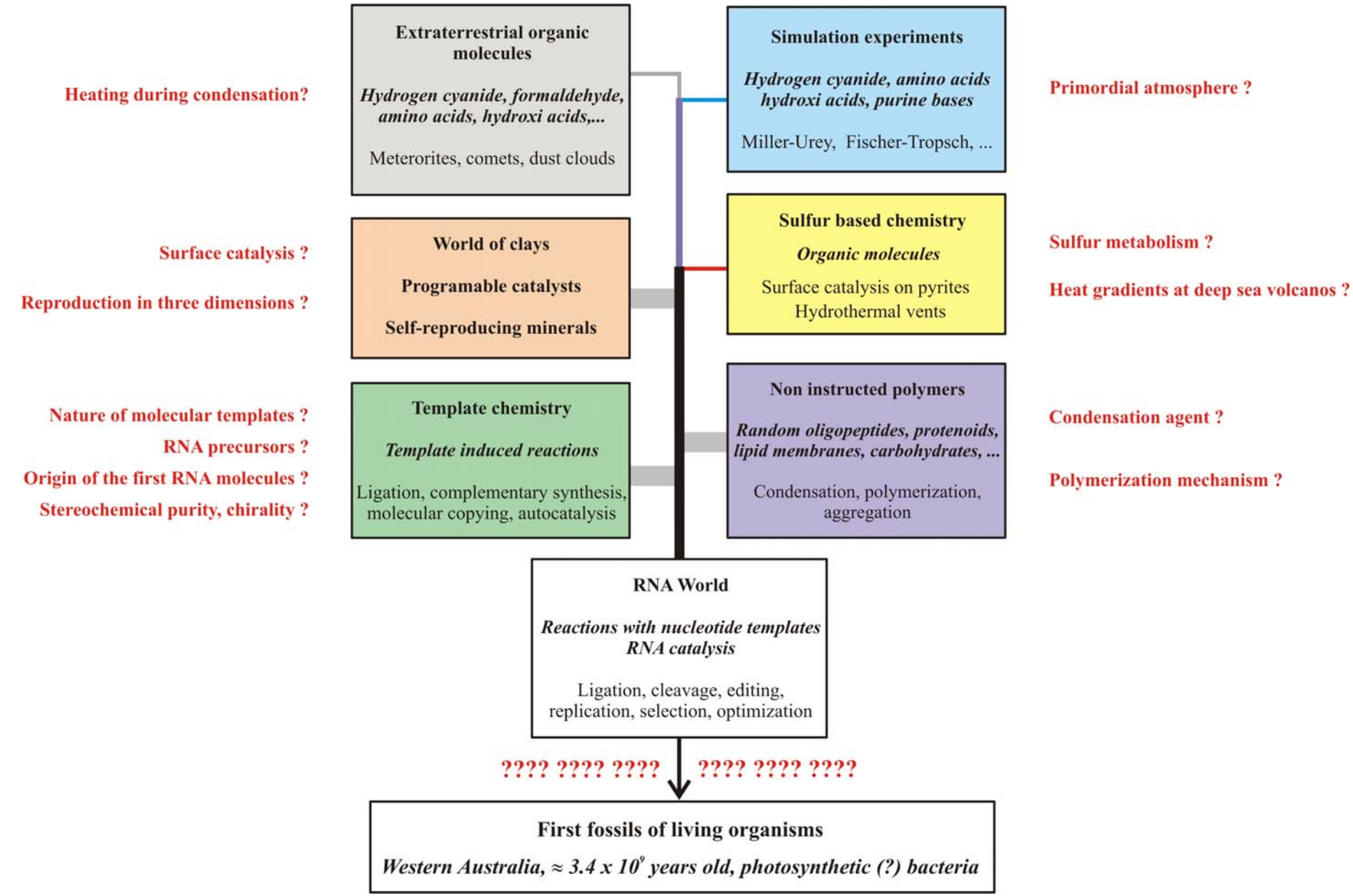
1. Was ist Leben?
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. Evolutionsmechanismen

1. **Was ist Leben?**
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. Evolutionsmechanismen

Kriterien zur Unterscheidung von belebter und unbelebter Materie

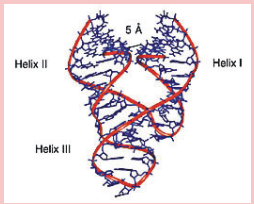
1. Befähigung zur **Vermehrung und Variation** durch Mutation
2. Befähigung zum **Lernen in Populationen** durch Replikation, Variation und Selektion
3. Molekulare Trennung zwischen **Genotyp und Phänotyp**
4. **Abgrenzung** gegenüber der Umwelt durch Membranen, Zellwände, Häute oder Verhalten
5. **Autopoiese** als Selbsterhalt mit Hilfe des **Stoffwechsels**
6. Arbeitsteilung durch **Zelldifferenzierung**
7. Befähigung zum **individuellen Lernen**
8. Übertagung **erworbener Eigenschaften** auf zukünftige Generationen durch Erziehung
9. Sprache, Schrift und Kultur

1. Was ist Leben?
2. **Chemische Evolution**
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. Evolutionsmechanismen


Extraterrestrial organic molecules
Hydrogen cyanide, formaldehyde, amino acids, hydroxi acids,...
Meterorites, comets, dust clouds
Heating during condensation?
Simulation experiments
Hydrogen cyanide, amino acids hydroxi acids, purine bases
Miller-Urey, Fischer-Tropsch, ...
Primordial atmosphere ?
World of clays
Programable catalysts
Self-reproducing minerals
Surface catalysis ?
Reproduction in three dimensions ?
Sulfur based chemistry
Organic molecules
Surface catalysis on pyrites
Hydrothermal vents
Sulfur metabolism ?
Heat gradients at deep sea volcanos ?
Template chemistry
Template induced reactions
Ligation, complementary synthesis, molecular copying, autocatalysis
Nature of molecular templates ?
RNA precursors ?
Origin of the first RNA molecules ?
Stereochemical purity, chirality ?
Non instructed polymers
Random oligopeptides, protenoids, lipid membranes, carbohydrates, ...
Condensation, polymerization, aggregation
Condensation agent ?
Polymerization mechanism ?
RNA World
Reactions with nucleotide templates
RNA catalysis
Ligation, cleavage, editing, replication, selection, optimization
???? ???? ????
???? ???? ????
First fossils of living organisms
Western Australia, ≈ 3.4 x 10⁹ years old, photosynthetic (?) bacteria

RNA

RNA *as catalyst*



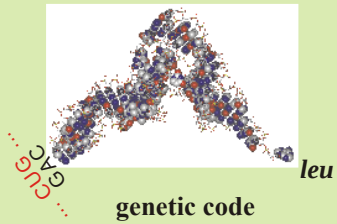
Ribozyme

RNA *as transmitter of genetic information*

DNA
↓ transcription
...AGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUC...
messenger-**RNA**
↓ translation
protein

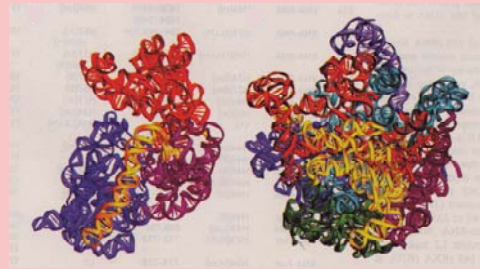
RNA as **working copy** of genetic information

RNA *as adapter molecule*



genetic code

RNA *is the catalytic subunit in supramolecular complexes*



The **ribosome** is a **ribozyme !**

The* RNA *world as a precursor of the current* DNA + protein *biology

RNA *is modified by epigenetic control*

RNA editing

Alternative splicing of messenger **RNA**

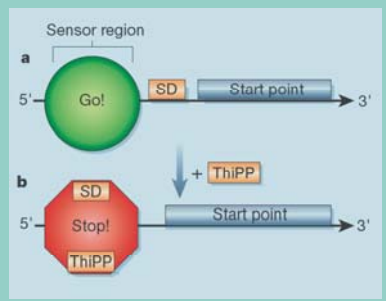
RNA *as carrier of genetic information*

RNA viruses and retroviruses

RNA evolution *in vitro*

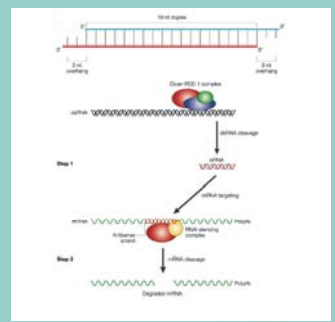
Evolutionary biotechnology
RNA aptamers, artificial ribozymes, allosteric ribozymes

***Allosteric control of transcribed* RNA**



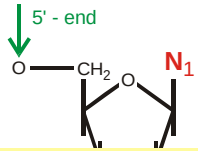
Riboswitches controlling transcription and translation through **metabolites**

RNA *as regulator of gene expression*

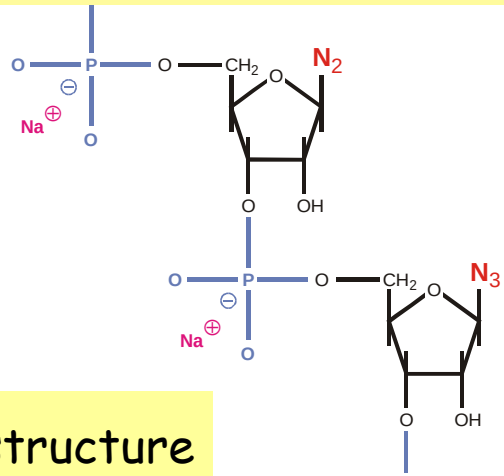


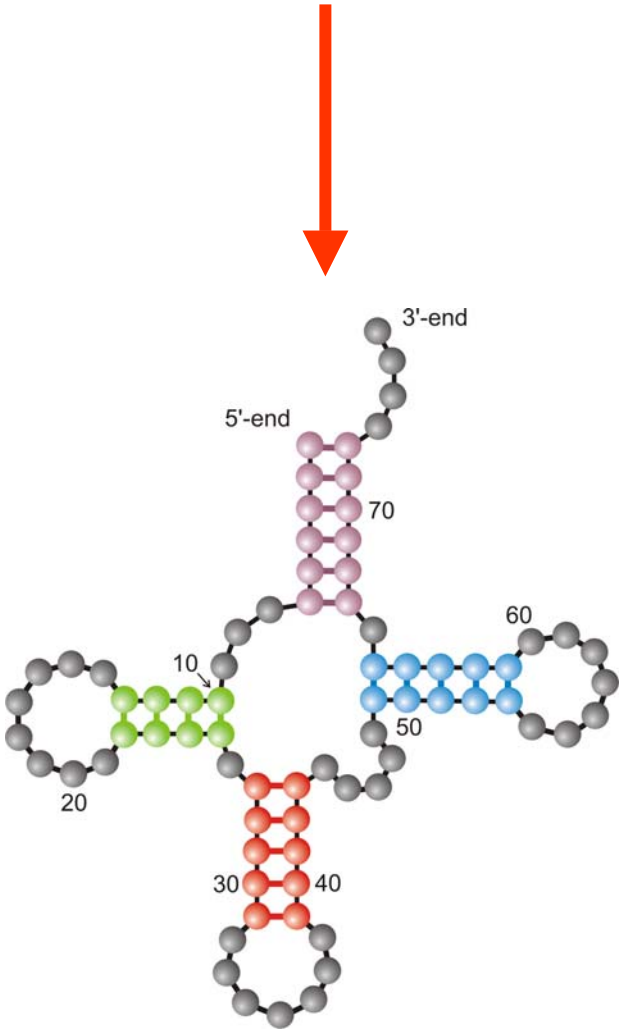
Gene silencing by small interfering RNAs

Functions of RNA molecules

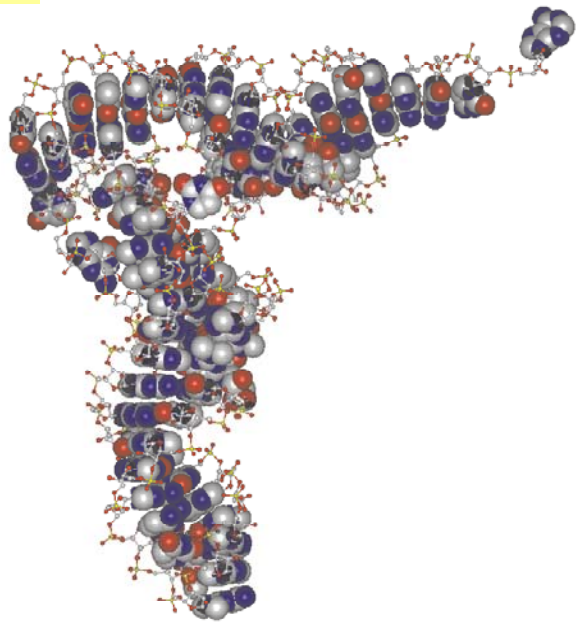


5'-end GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'-end

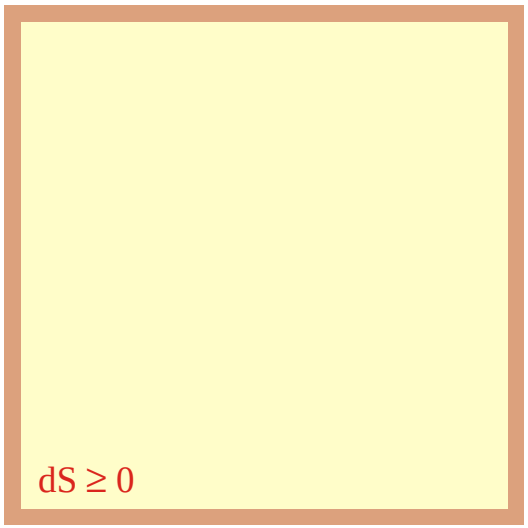


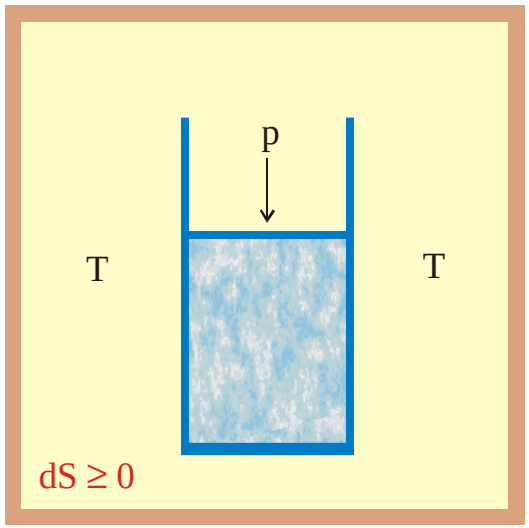


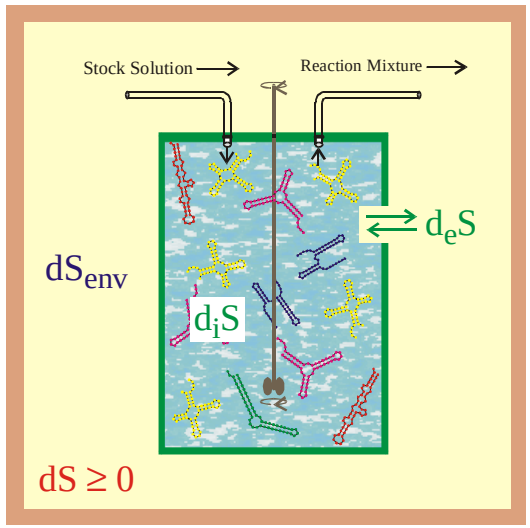
Definition of RNA structure



1. Was ist Leben?
2. Chemische Evolution
3. **Der Ursprung biologischer Information**
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. Evolutionsmechanismen







Isolated system

U = const., V = const.,

$dS \geq 0$

Closed system

T = const., p = const.,

$dG = dU - pdV - TdS \leq 0$

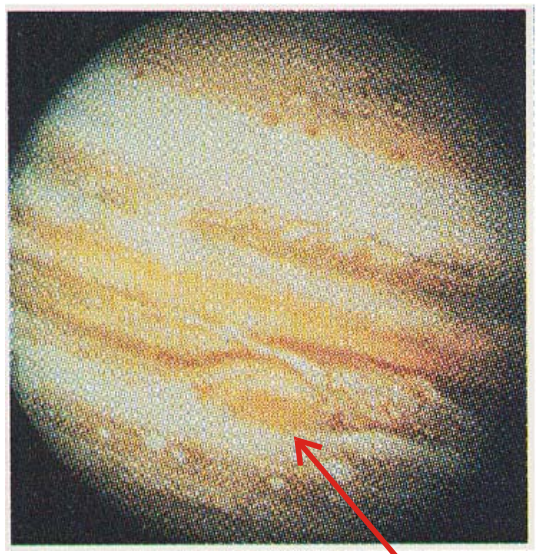
Open system

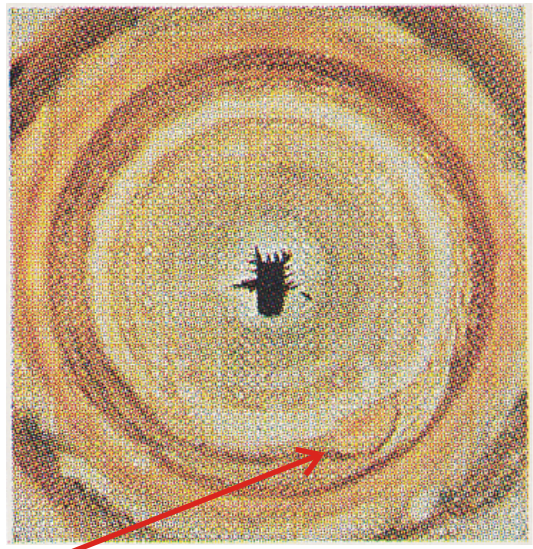
$dS = dS_{env} + dS \geq 0$

$dS = d_iS + d_eS$

$d_iS \geq 0$

Entropy changes in different thermodynamic systems

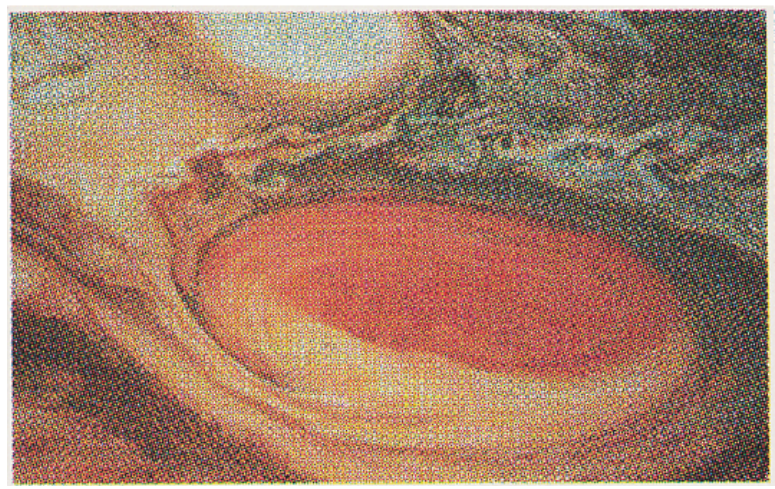




View from south pole

South pole

Red spot



Jupiter: Observation of the gigantic vortex

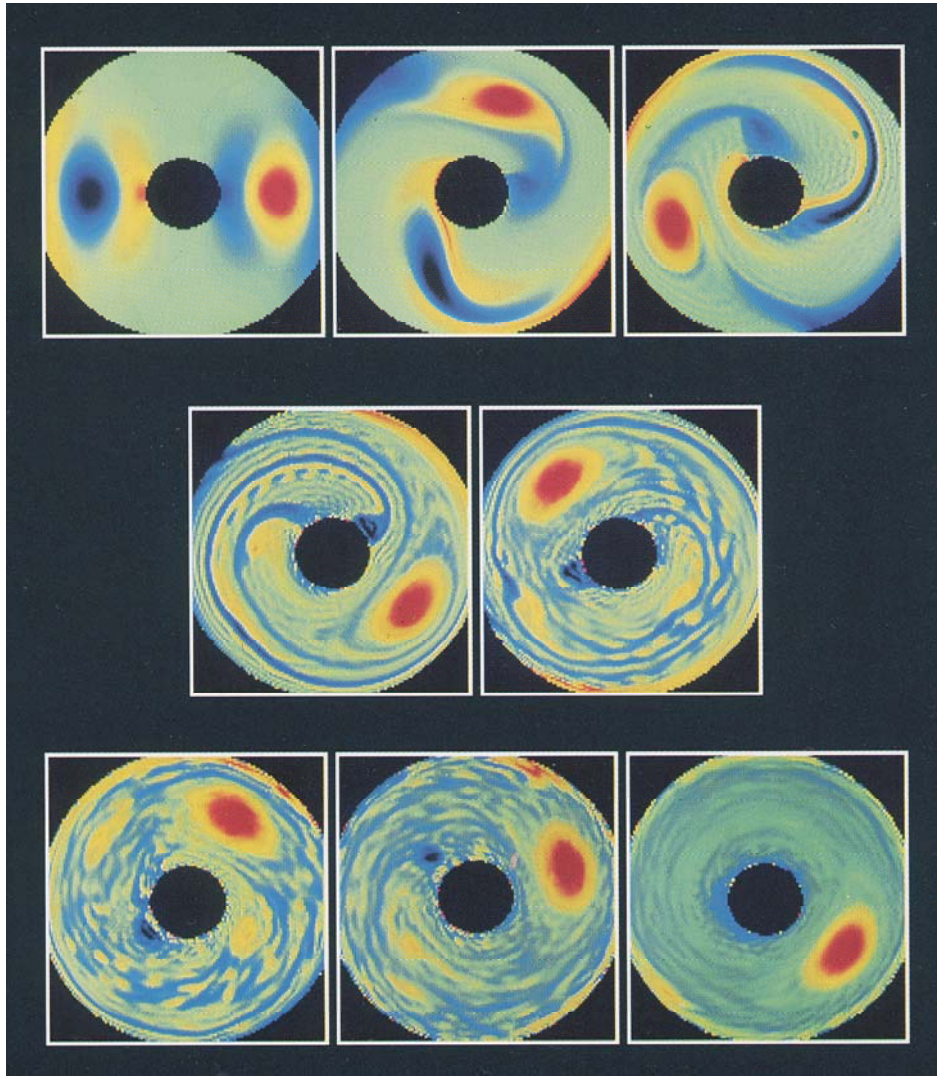
Picture taken from James Gleick, *Chaos*. Penguin Books, New York, 1988

Computer simulation of the gigantic vortex on Jupiter

View from south pole

Particles turning counterclockwise

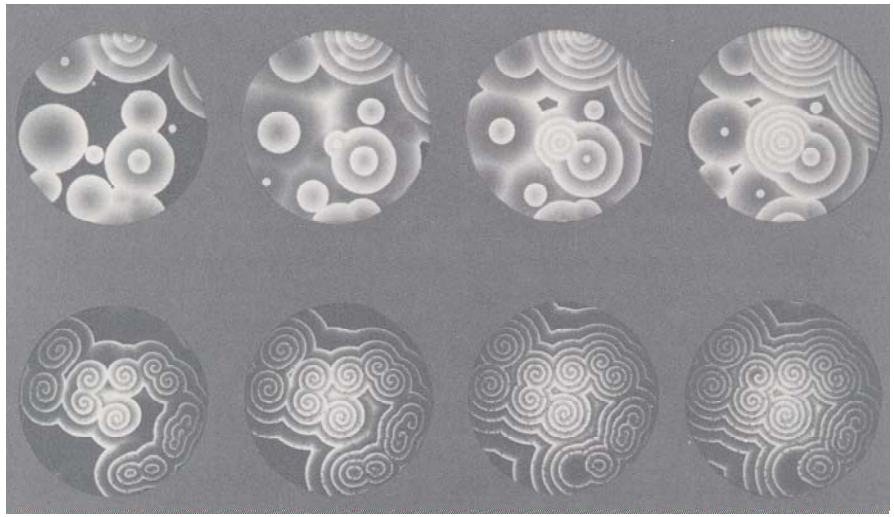
Particles turning clockwise



Jupiter: Computer simulation of the giant vortex

Philip Marcus, 1980. Picture taken from James Gleick, *Chaos*. Penguin Books, New York, 1988

Formation of target waves
and spirals in the
Belousov-Zhabotinskii reaction



Winding number:

number of left-handed spirals
minus
number of right-handed spirals

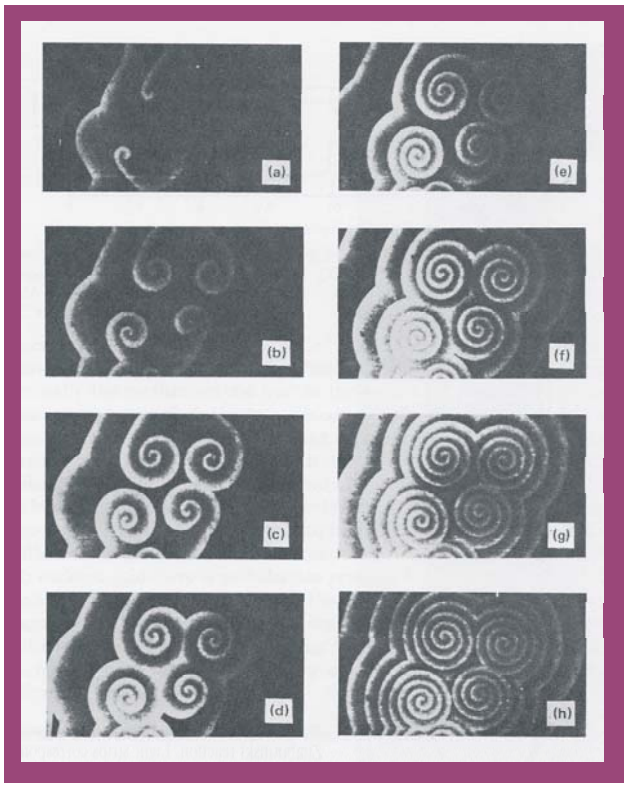
Target waves and spirals in the Belousov-Zhabotinskii reaction

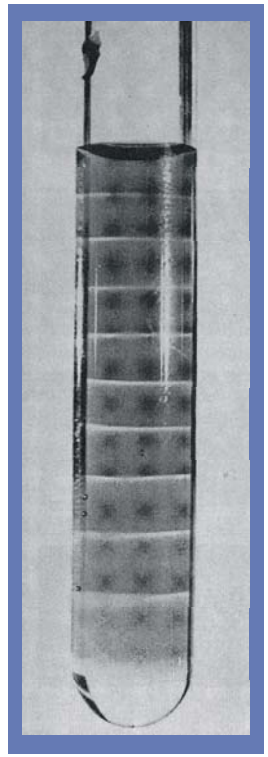
Pictures taken from Arthur T. Winfree, *The geometry of biological time*. Springer-Verlag, New York, 1980.

Autocatalytic third order reactions

Direct, $A + 2X \rightarrow 3X$, or hidden in the reaction mechanism (Belousow-Zhabotinskii reaction).

- Multiple steady states
- Oscillations in homogeneous solution
- Deterministic chaos
- Turing patterns
- Spatiotemporal patterns (spirals)
- Deterministic chaos in space and time



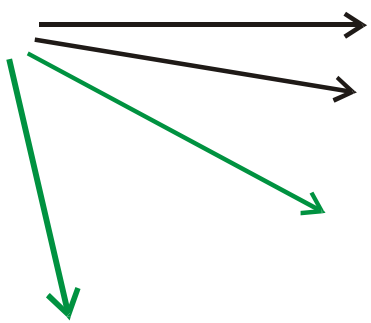


Pattern formation in autocatalytic third order reactions

G.Nicolis, I.Prigogine. *Self-Organization in Nonequilibrium Systems. From Dissipative Structures to Order through Fluctuations*. John Wiley, New York 1977

Autocatalytic second order reactions

Direct, $A + I \rightarrow 2\,I$, or hidden in the reaction mechanism



Chemical self-enhancement

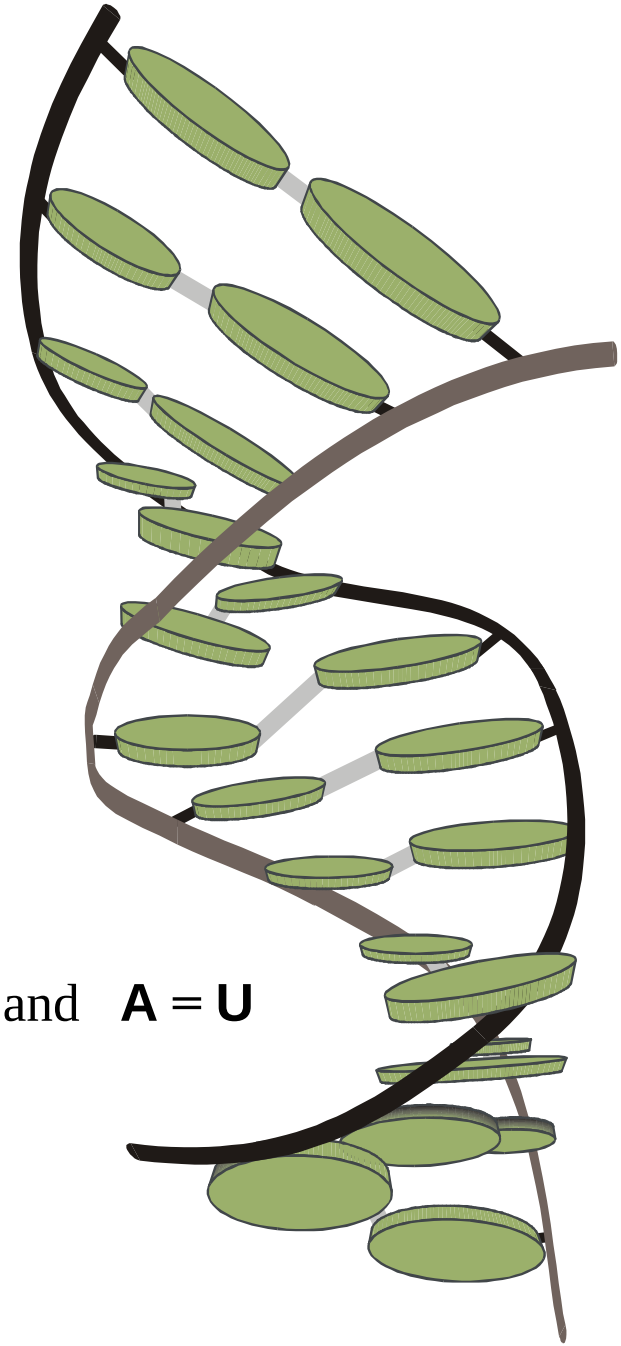
Combustion and chemistry of flames

Selection of laser modes

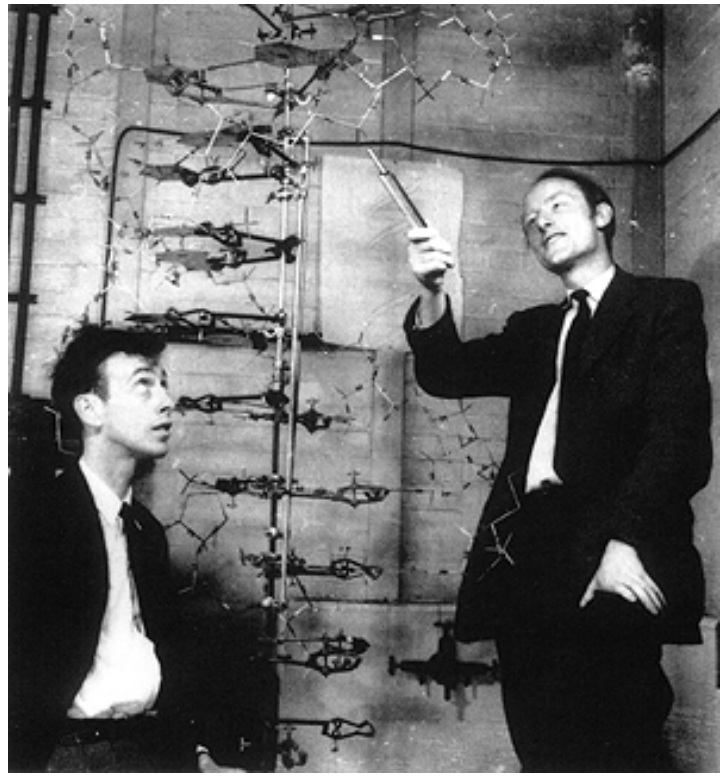
Selection of molecular or organismic species competing for common sources

Autocatalytic second order reaction as basis for selection processes.

The autocatalytic step is formally equivalent to replication or reproduction.

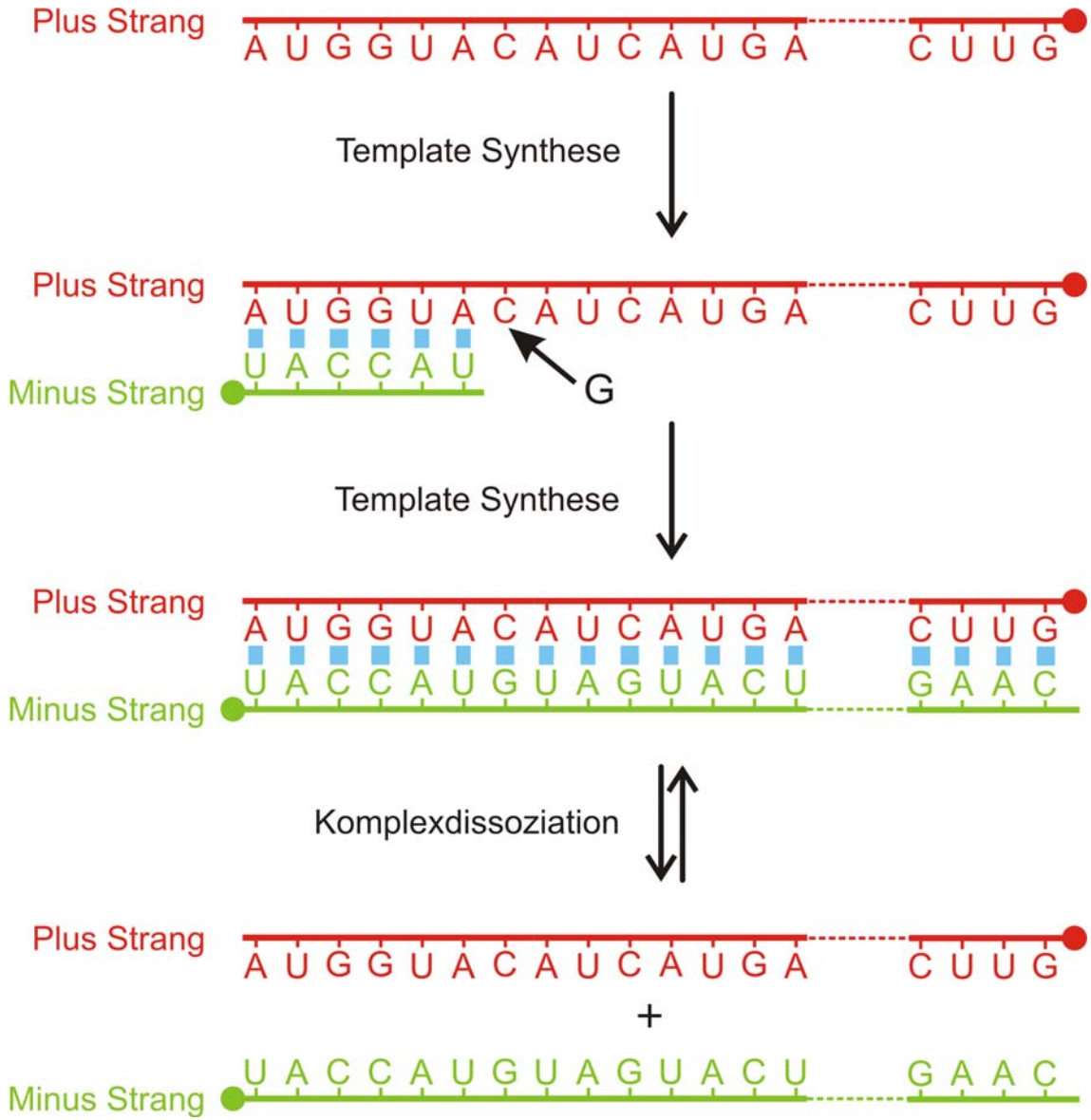


G ≡ C and **A = U**

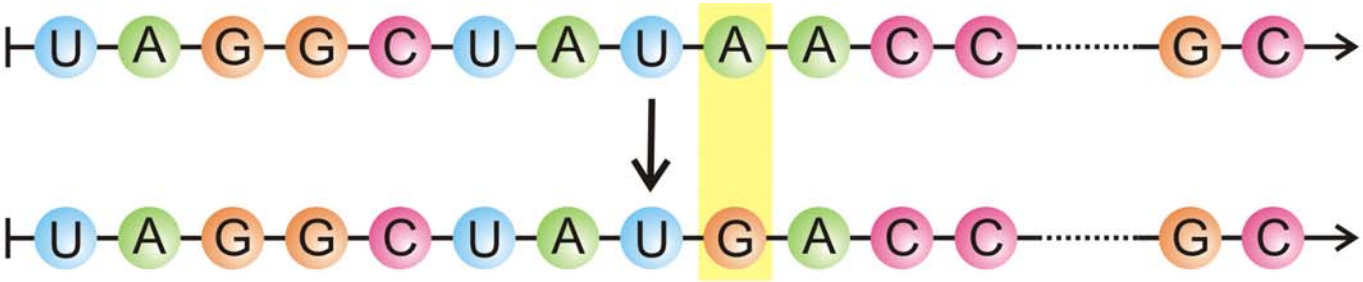


James D. Watson, 1928- , and Francis Crick, 1916-2004,
Nobel Prize 1962

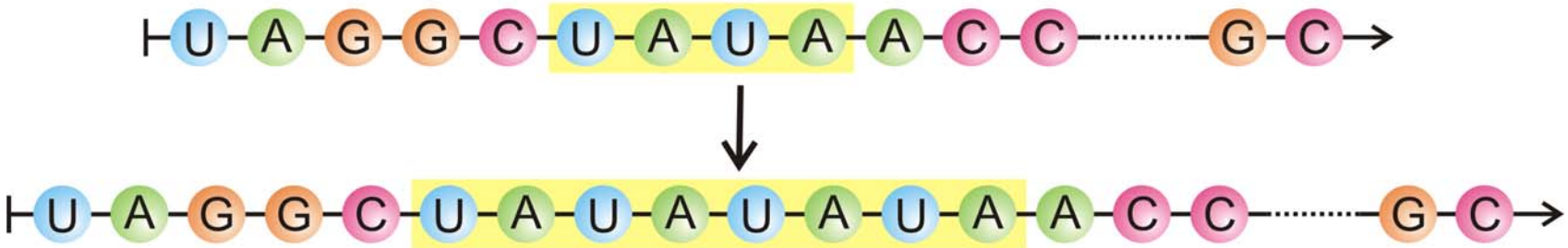
The three-dimensional structure of a short double helical stack of B-DNA



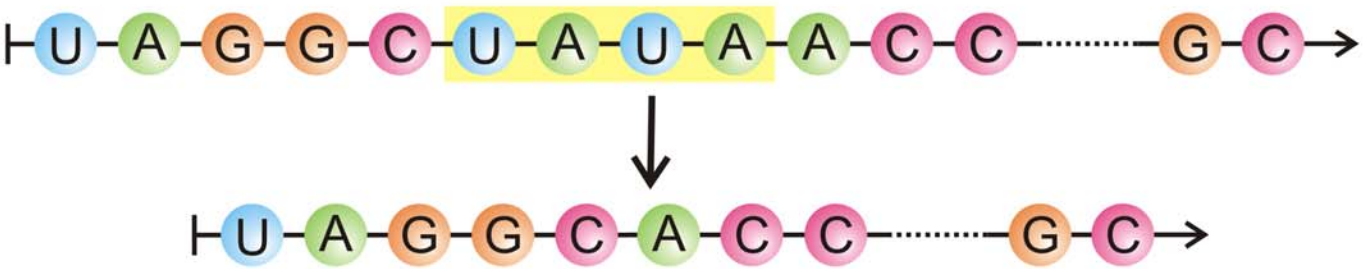
Der Mechanismus der Replikation einsträngiger RNA-Moleküle



Punktmutation



Insertion



Deletion

1. Was ist Leben?
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. **Darwinsche Evolution mit Molekülen**
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. Evolutionsmechanismen

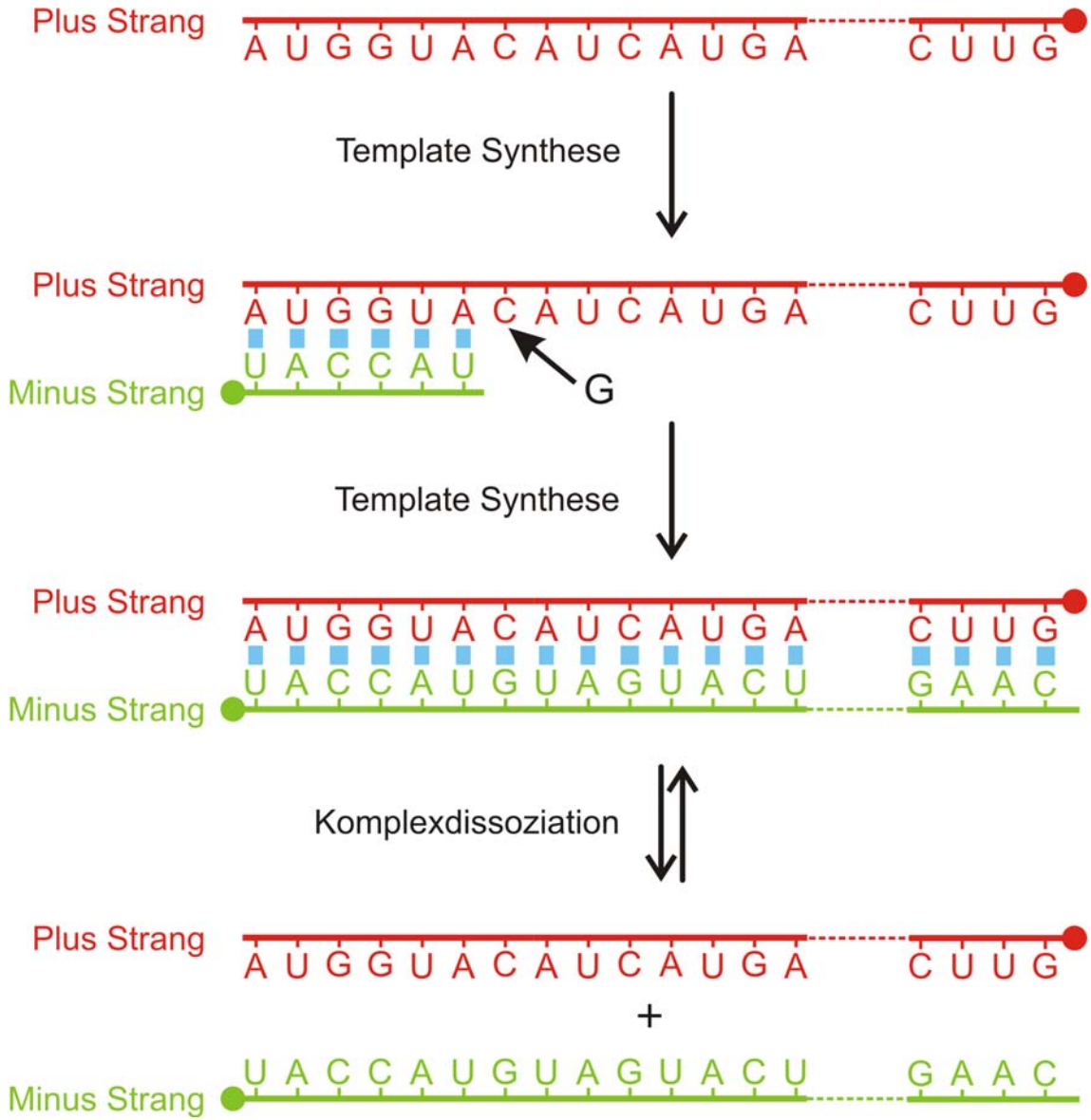
Three necessary conditions for Darwinian evolution are:

1. **Multiplication,**
2. **Variation**, and
3. **Selection**.

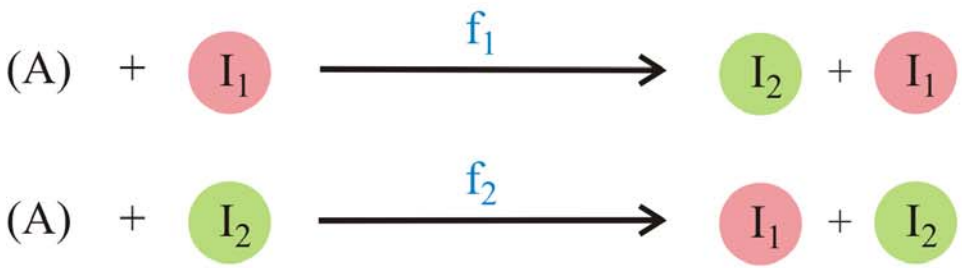
Variation through mutation and recombination operates on the **genotype** whereas the **phenotype** is the target of **selection**.

One important property of the Darwinian scenario is that **variations** in the form of mutations or recombination events occur **uncorrelated** with their **effects on the selection process**.

All conditions can be fulfilled not only by cellular organisms but also by **nucleic acid molecules** in suitable **cell-free experimental assays**.



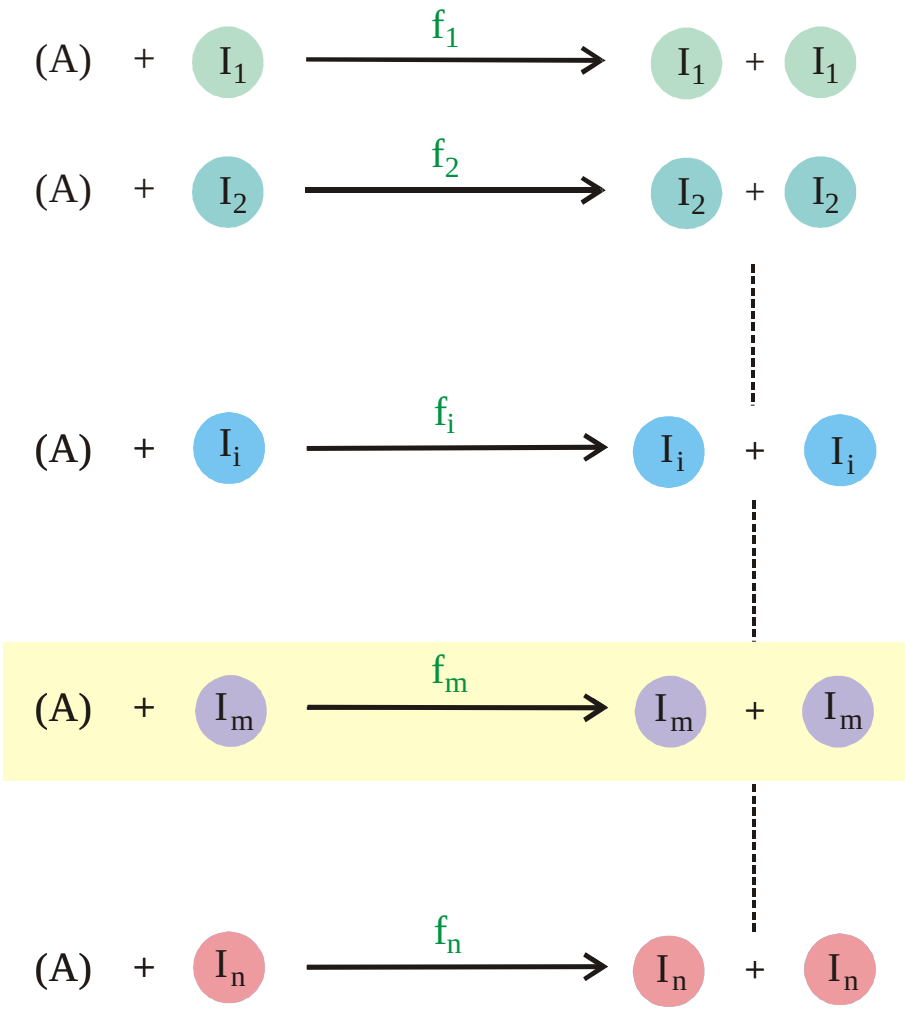
Der Mechanismus der Replikation einsträngiger RNA-Moleküle



$$dx_1 / dt = f_2 x_2 - x_1 \Phi$$
$$dx_2 / dt = f_1 x_1 - x_2 \Phi$$

$$\Phi = \Sigma_i f_i x_i ; \quad \Sigma_i x_i = 1 ; \quad i = 1,2$$

Complementary replication as the simplest molecular mechanism of reproduction



$$dx_i / dt = f_i x_i - x_i \Phi = x_i (f_i - \Phi)$$

$$\Phi = \Sigma_j f_j x_j ; \quad \Sigma_j x_j = 1 ; \quad i,j = 1,2,...,n$$

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

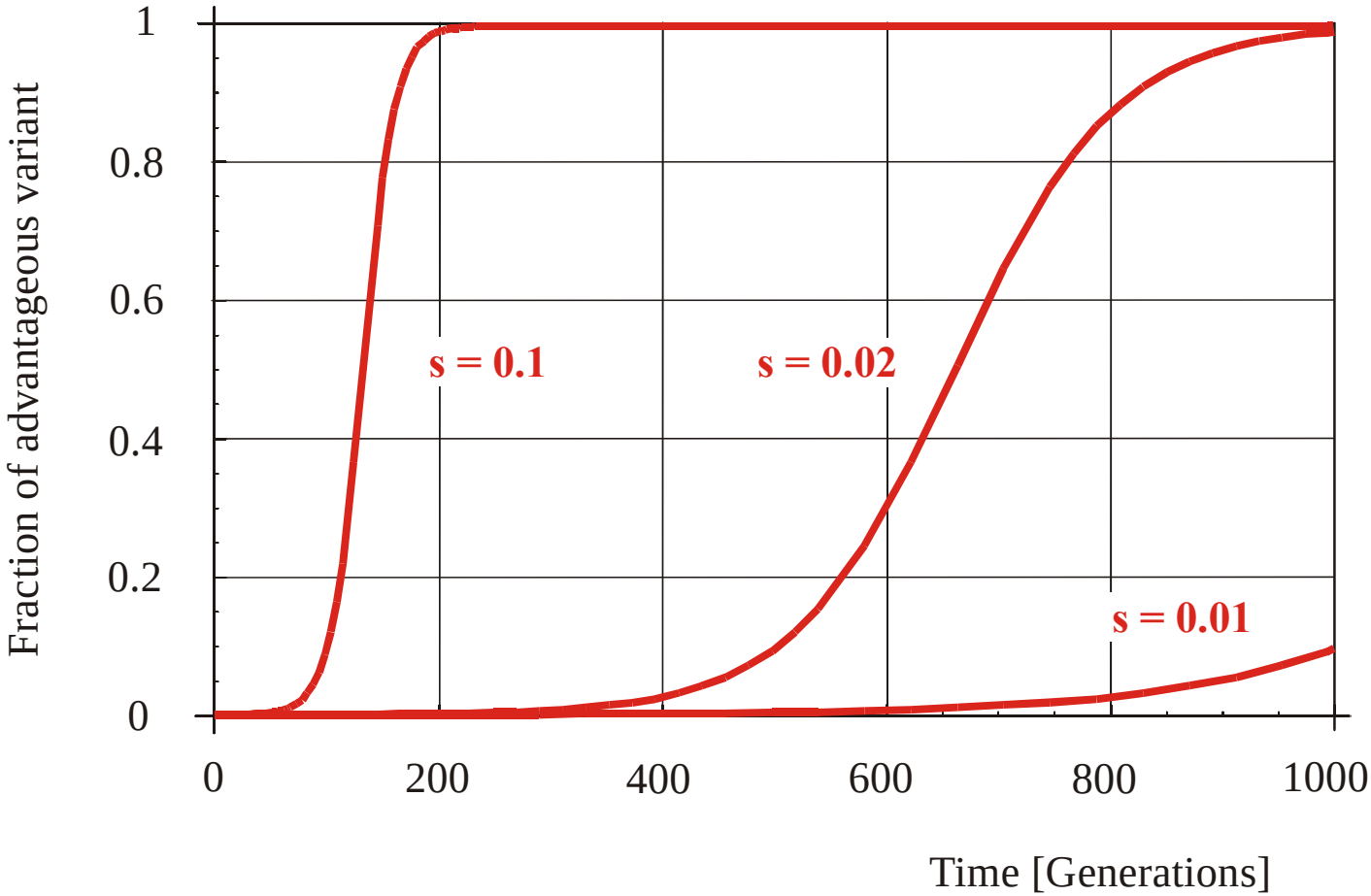
$$[A] = a = \text{constant}$$

$$f_m = \max \{f_j; j=1,2,...,n\}$$

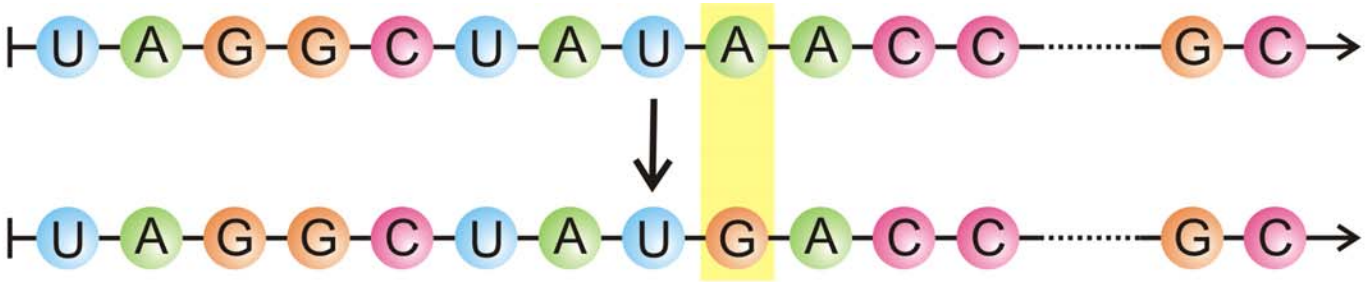
$$x_m(t) \rightarrow 1 \quad \text{for} \quad t \rightarrow \infty$$

Reproduction of organisms **or replication** of molecules as the basis of selection

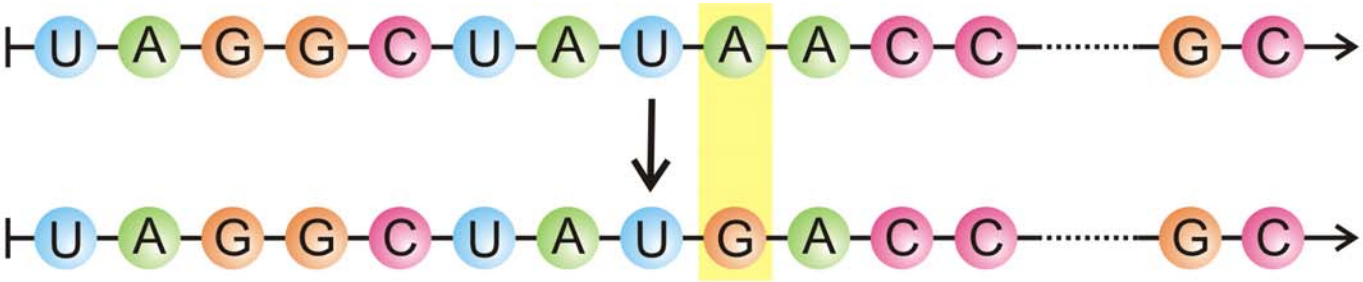
$s = (f_2 - f_1) / f_1;\ f_2 > f_1\ ;\ x_1(0) = 1 - 1/N\ ;\ x_2(0) = 1/N$



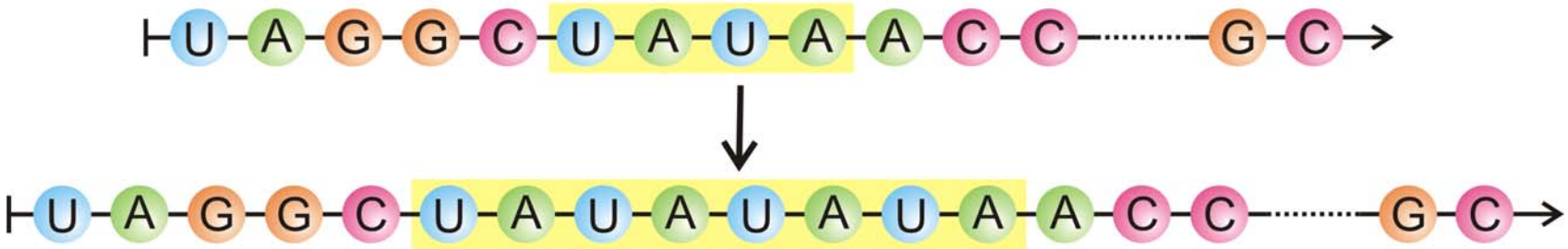
Selection of advantageous mutants in populations of N = 10 000 individuals



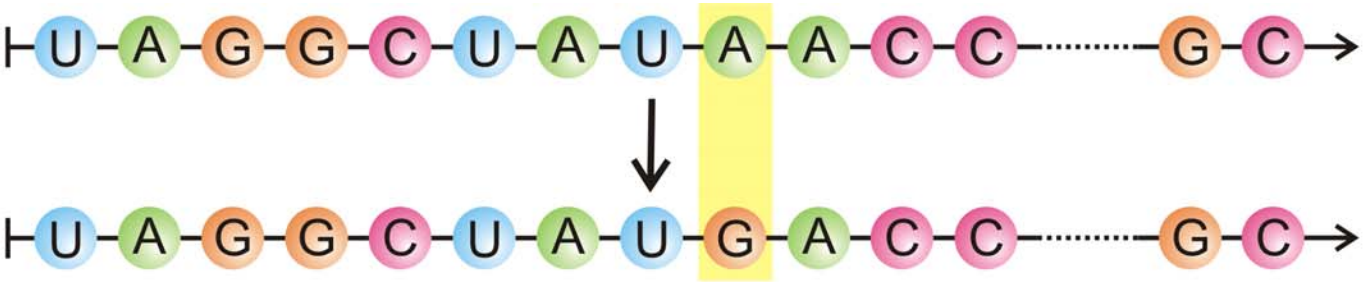
Punktmutation



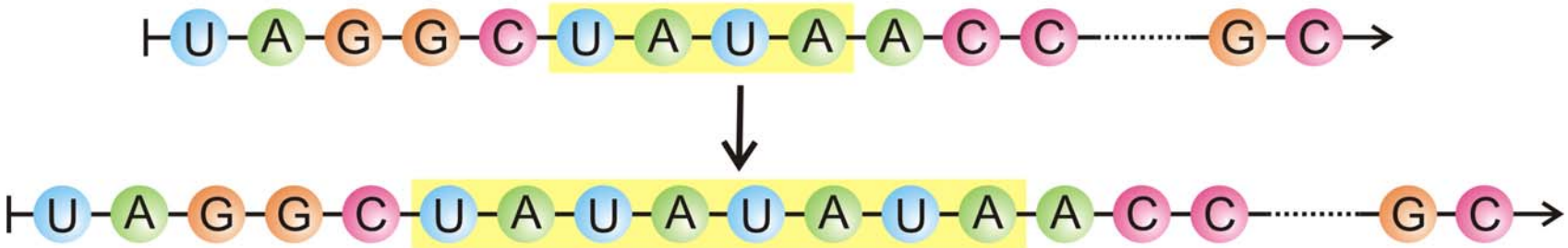
Punktmutation



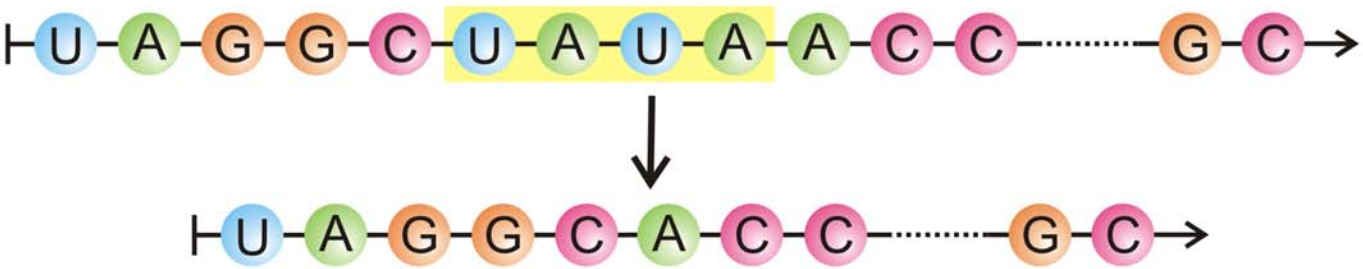
Insertion



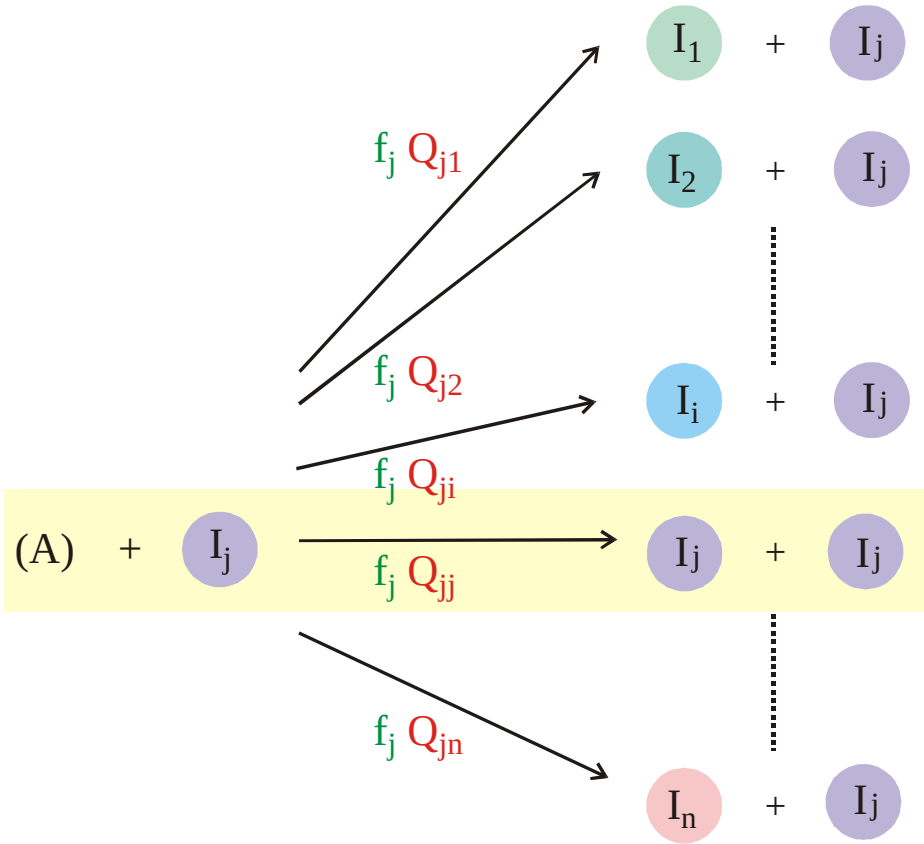
Punktmutation



Insertion



Deletion



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

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

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

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

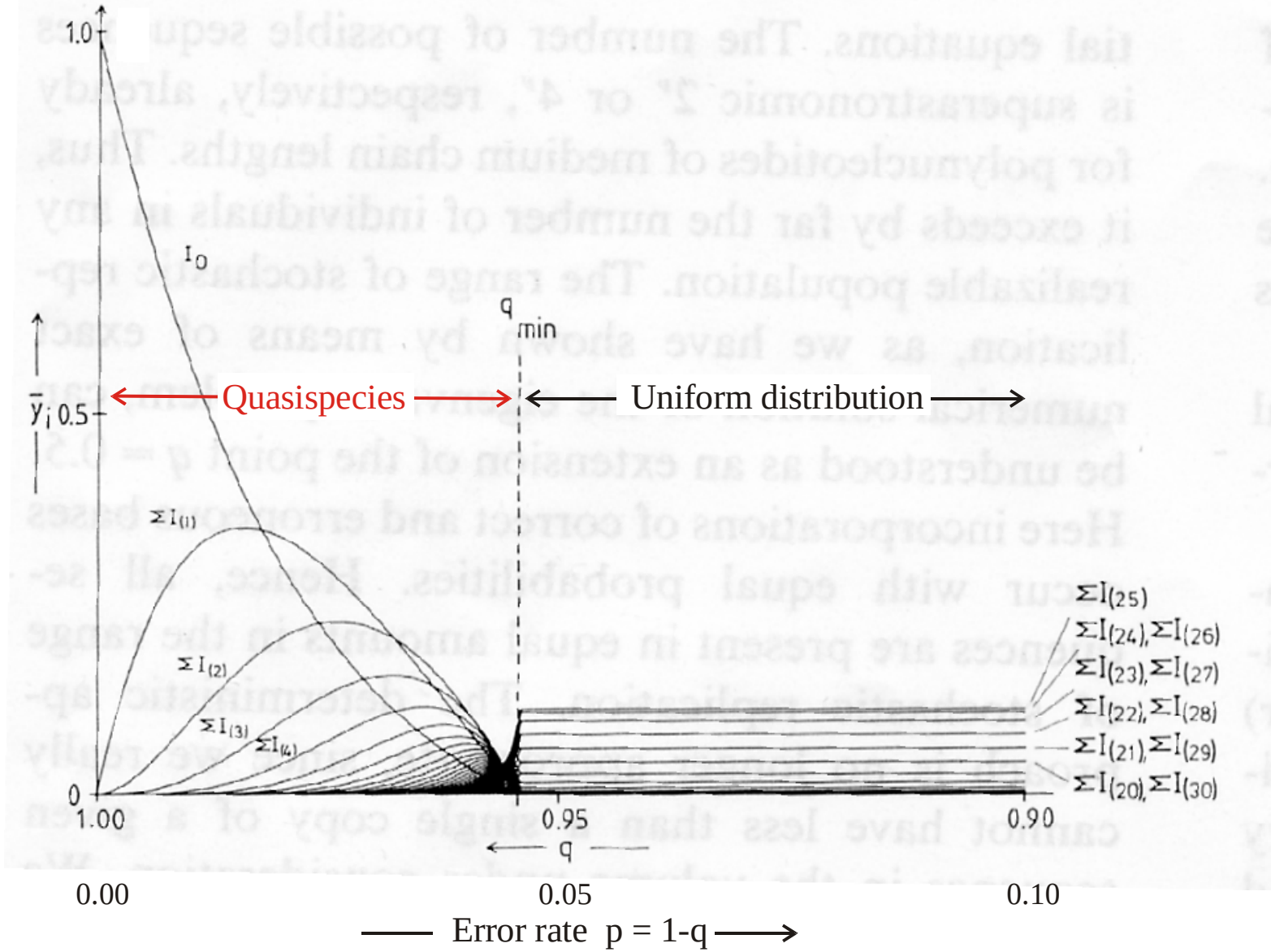
$$Q_{ij} = (1-p)^{\ell - d(i,j)} \, p^{d(i,j)}$$

p ..... Error rate per digit

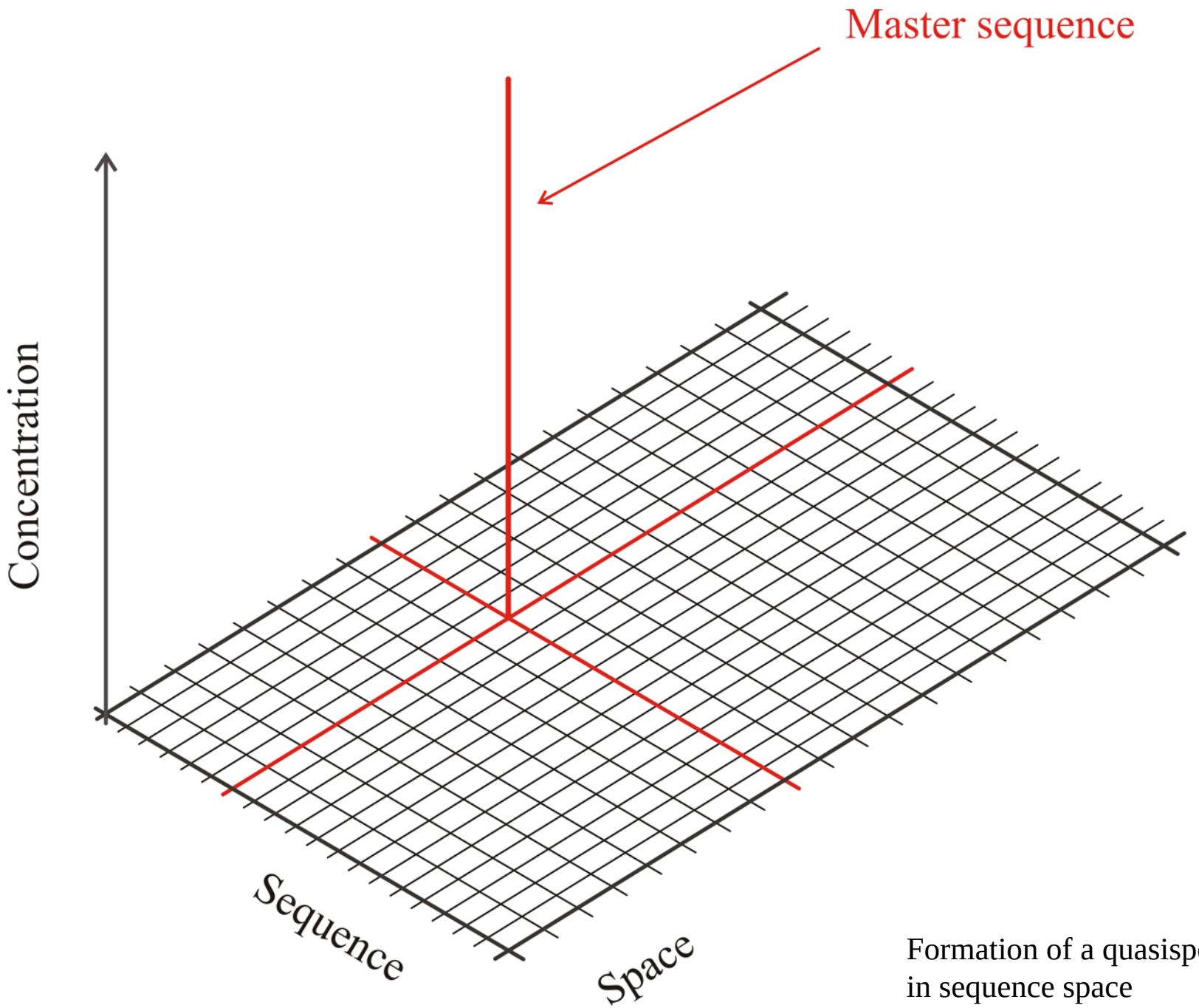
ℓ ..... Chain length of the polynucleotide

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

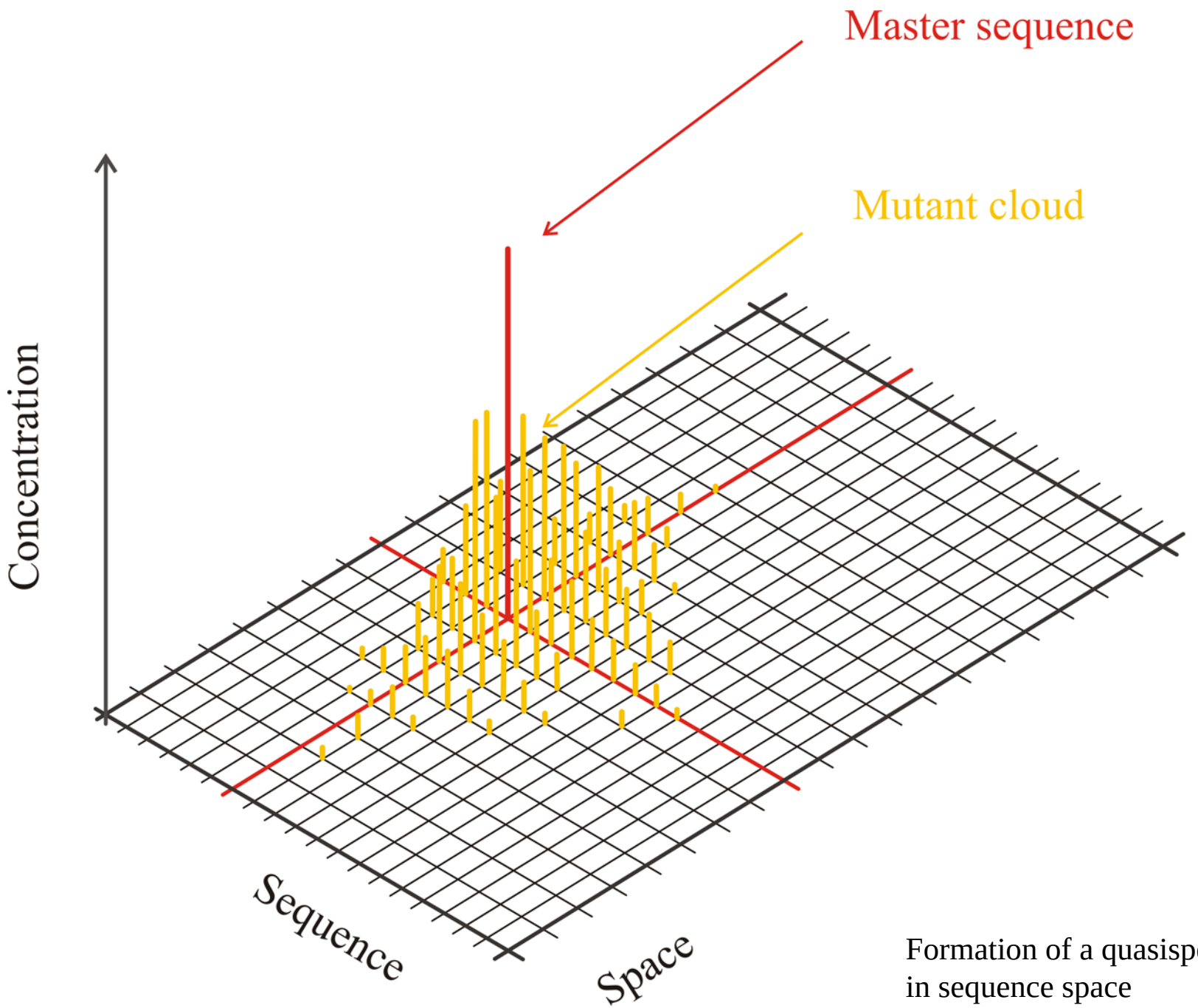
Chemical kinetics of replication and mutation as parallel reactions



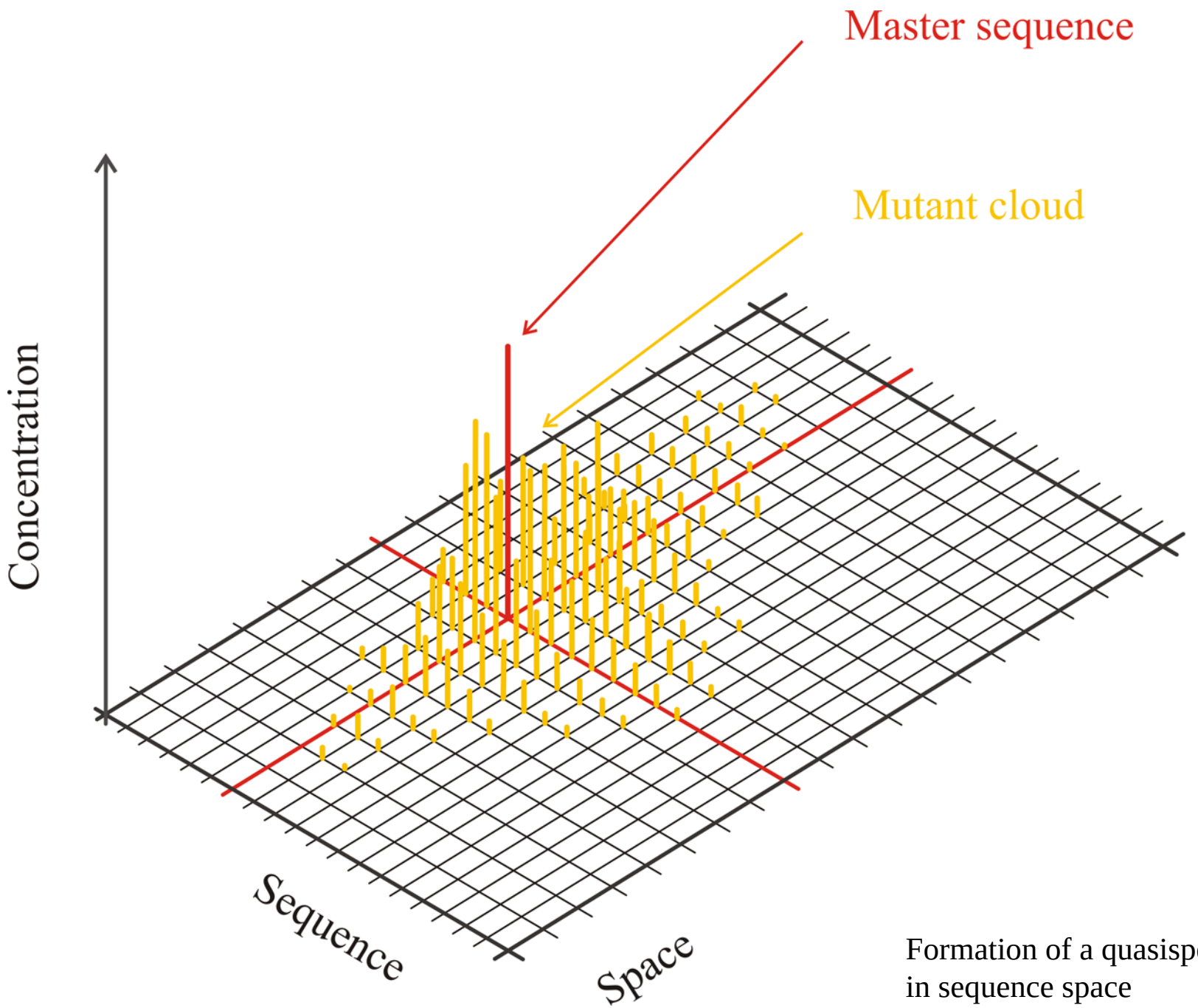
Quasispecies as a function of the replication accuracy q



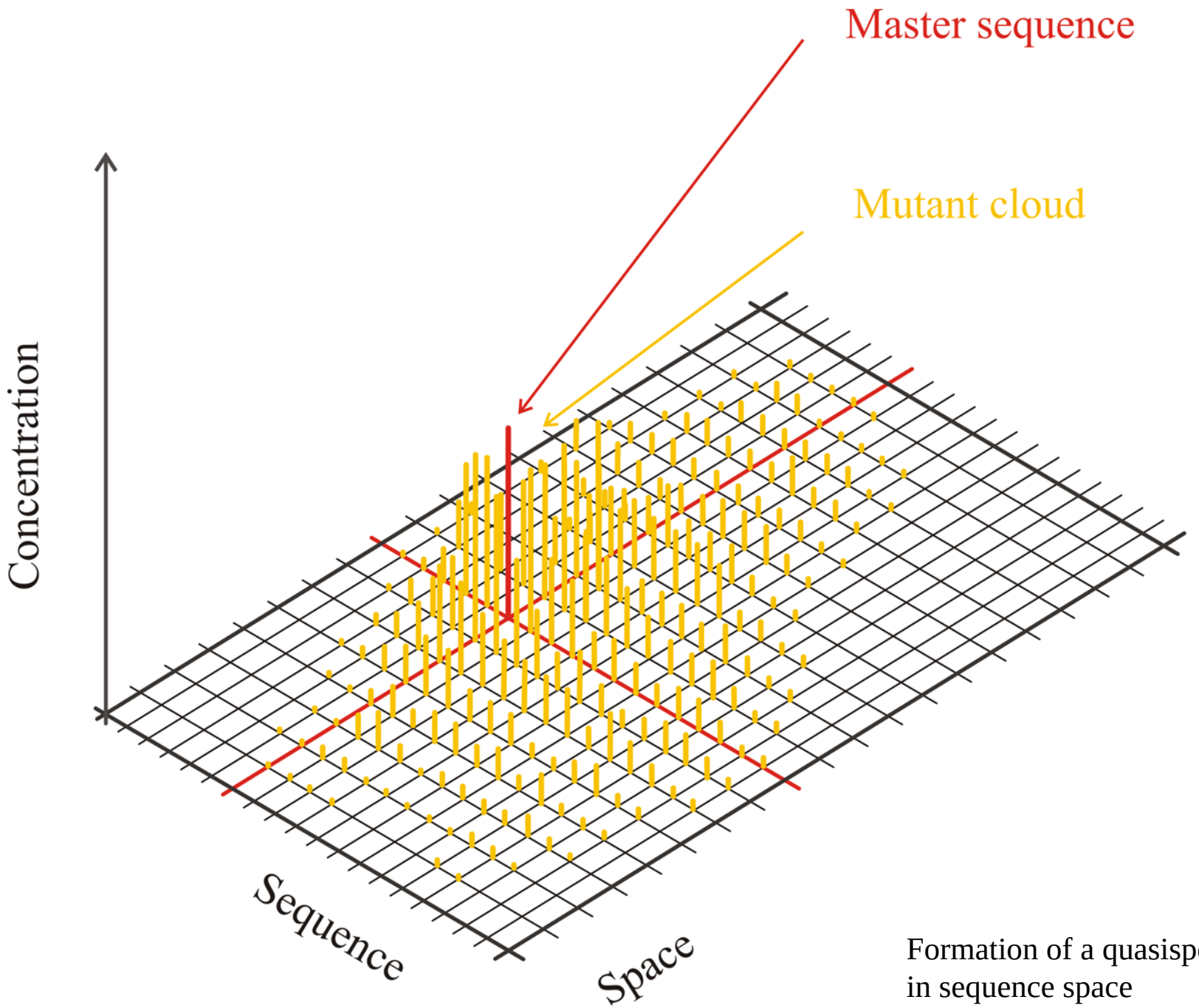
Formation of a quasispecies in sequence space



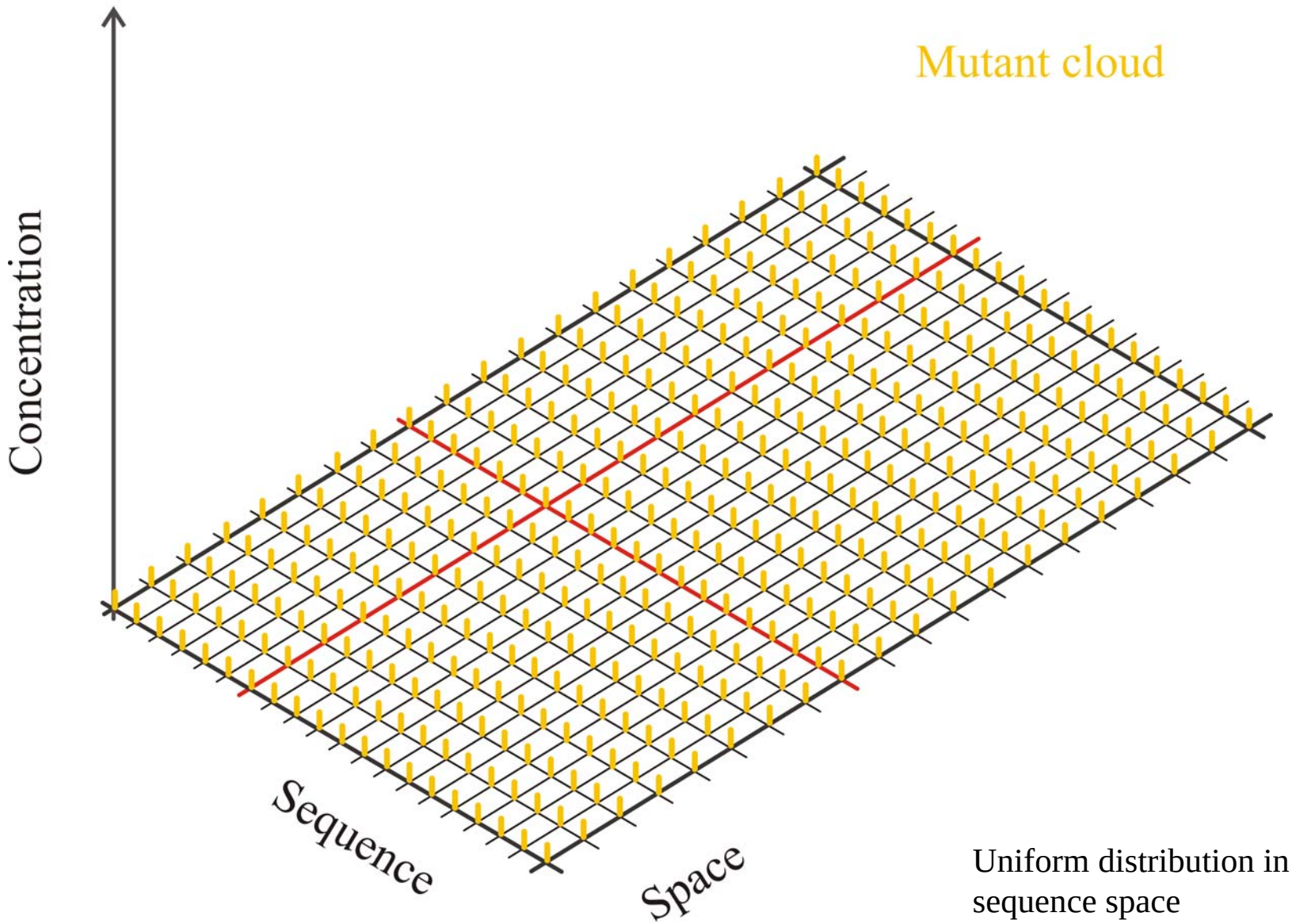
Formation of a quasispecies in sequence space



Formation of a quasispecies in sequence space



Formation of a quasispecies in sequence space



Uniform distribution in sequence space

Evolution *in silico*

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

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

32. A nonsense mutation may affect mRNA stability and result in degradation of the transcript [L. Maquat, *Am. J. Hum. Genet.* **59**, 279 (1996)].
33. Data not shown; a dot blot with poly (A)$^+$ RNA from 50 human tissues (The Human RNA Master Blot, 7770-1, Clontech Laboratories) was hybridized with a probe from exons 29 to 47 of *MYO15* using the same condition as Northern blot analysis (*13*).
34. Smith–Magenis syndrome (SMS) is due to deletions of 17p11.2 of various sizes, the smallest of which includes *MYO15* and perhaps 20 other genes [(*6*); K-S Chen, L. Potocki, J. R. Lupski, *MRDD Res. Rev.* **2**, 122 (1996)]. *MYO15* expression is easily detected in the pituitary gland (data not shown). Haploinsufficiency for *MYO15* may explain a portion of the SMS phenotype such as short stature. Moreover, a few SMS patients have sensorineural hearing loss, possibly because of a point mutation in *MYO15* in trans to the SMS 17p11.2 deletion.
35. R. A. Fridell, data not shown.
36. K. B. Avraham *et al.*, *Nature Genet.* **11**, 369 (1995); X-Z. Liu *et al.*, *ibid.* **17**, 268 (1997); F. Gibson *et al.*, *Nature* **374**, 62 (1995); D. Weil *et al.*, *ibid.*, p. 60.
37. RNA was extracted from cochlea (membranous labyrinths) obtained from human fetuses at 18 to 22 weeks of development in accordance with guidelines established by the Human Research Committee at the Brigham and Women's Hospital. Only samples without evidence of degradation were pooled for poly (A)$^+$ selection over oligo(dT) columns. First-strand cDNA was prepared using an Advantage RT-for-PCR kit (Clontech Laboratories). A portion of the first-strand cDNA (4%) was amplified by PCR with Advantage cDNA polymerase mix (Clontech Laboratories) using human *MYO15*-specific oligonucleotide primers (forward, 5′-GCATGACCTGCCGGCTAAT-GGG-3′; reverse, 5′-CTCACGGCTTCTGCATGGT-GCTCGGCTGGC-3′). Cycling conditions were 40 s at 94°C; 40 s at 66°C (3 cycles), 60°C (5 cycles), and 55°C (29 cycles); and 45 s at 68°C. PCR products were visualized by ethidium bromide staining after fractionation in a 1% agarose gel. A 688-bp PCR product is expected from amplification of the human *MYO15* cDNA. Amplification of human genomic DNA with this primer pair would result in a 2903-bp fragment.

38. We are grateful to the people of Bengkala, Bali, and the two families from India. We thank J. R. Lupski and K.-S. Chen for providing the human chromosome 17 cosmid library. For technical and computational assistance, we thank N. Dietrich, M. Fergusson, A. Gupta, E. Sorbello, R. Torkzadeh, C. Varner, M. Walker, G. Bouffard, and S. Beckstrom-Sternberg (National Institutes of Health Intramural Sequencing Center). We thank J. T. Hinnant, I. N. Arhya, and S. Winata for assistance in Bali, and T. Barber, S. Sullivan, E. Green, D. Drayna, and J. Battey for helpful comments on this manuscript. Supported by the National Institute on Deafness and Other Communication Disorders (NIDCD) (Z01 DC 00035-01 and Z01 DC 00038-01 to T.B.F. and E.R.W. and R01 DC 03402 to C.C.M.), the National Institute of Child Health and Human Development (R01 HD30428 to S.A.C.) and a National Science Foundation Graduate Research Fellowship to F.J.P. This paper is dedicated to J. B. Snow Jr. on his retirement as the Director of the NIDCD.

9 March 1998; accepted 17 April 1998

Continuity in Evolution: On the Nature of Transitions

Walter Fontana and Peter Schuster

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

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

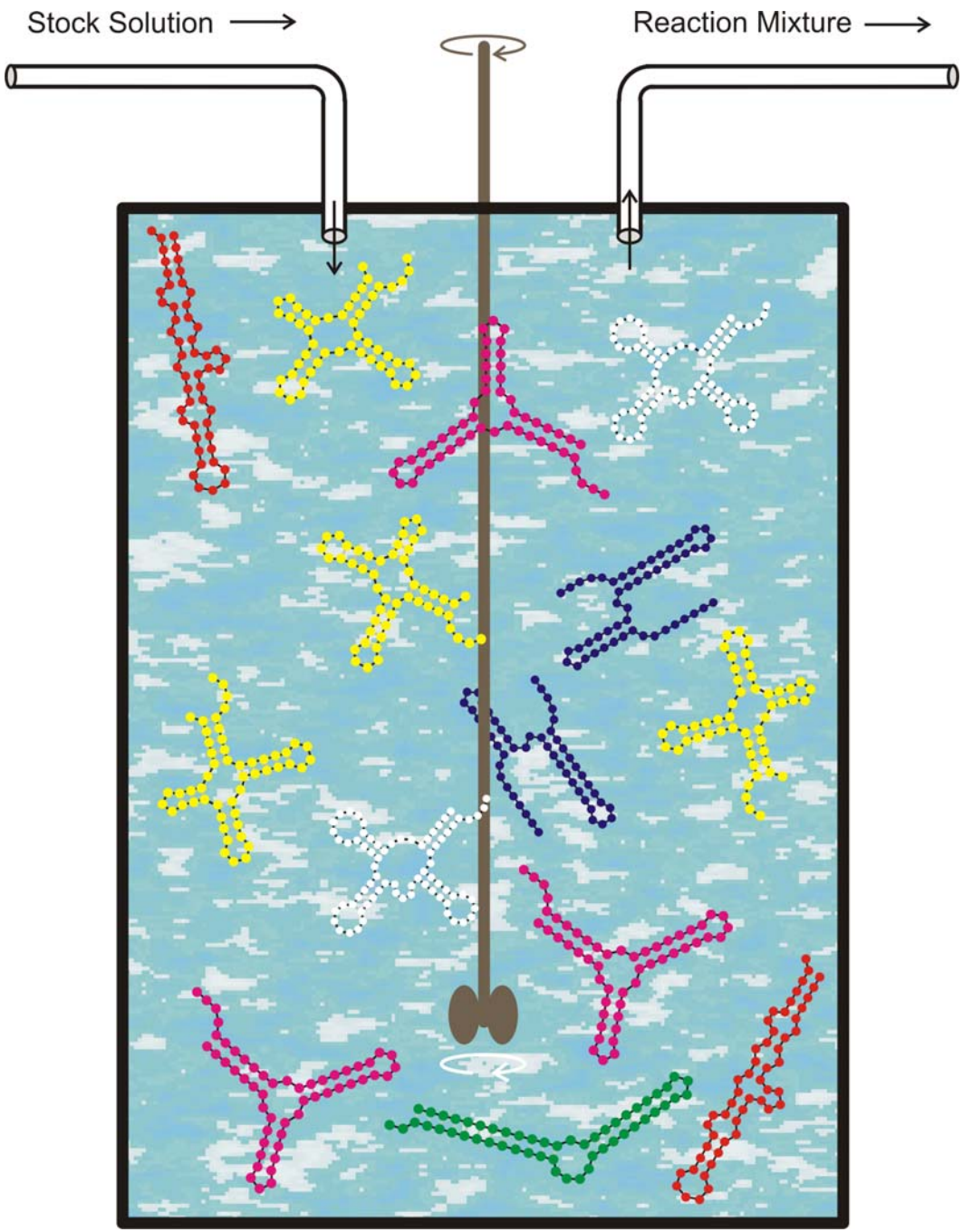
We focus on the narrow domain of RNA secondary structure, which is currently the simplest computationally tractable, yet realistic phenotype (*2*). This choice enables the definition and exploration of concepts that may prove useful in a wider context. RNA secondary structures represent a coarse level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

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

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

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

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



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

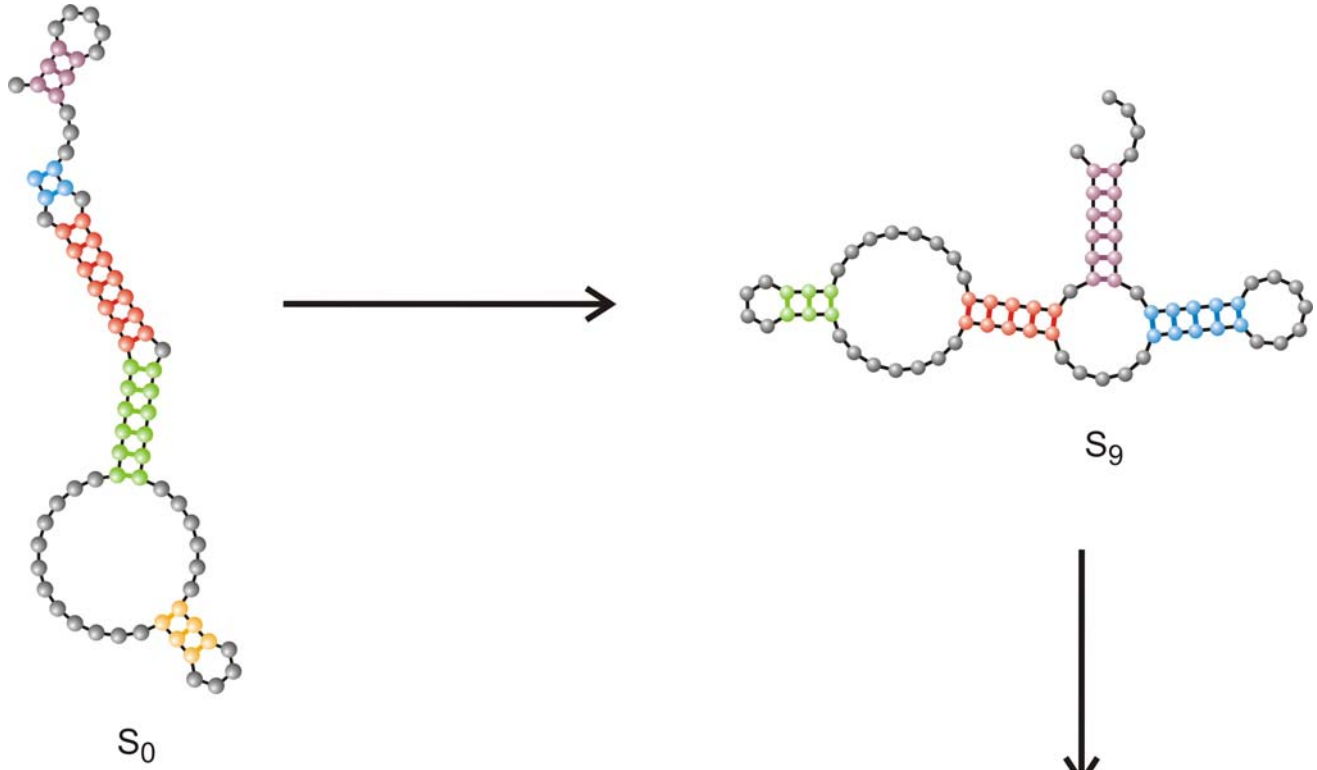
$$N(t) \approx \overline{N} \pm \sqrt{\overline{N}}$$

Mutation rate:

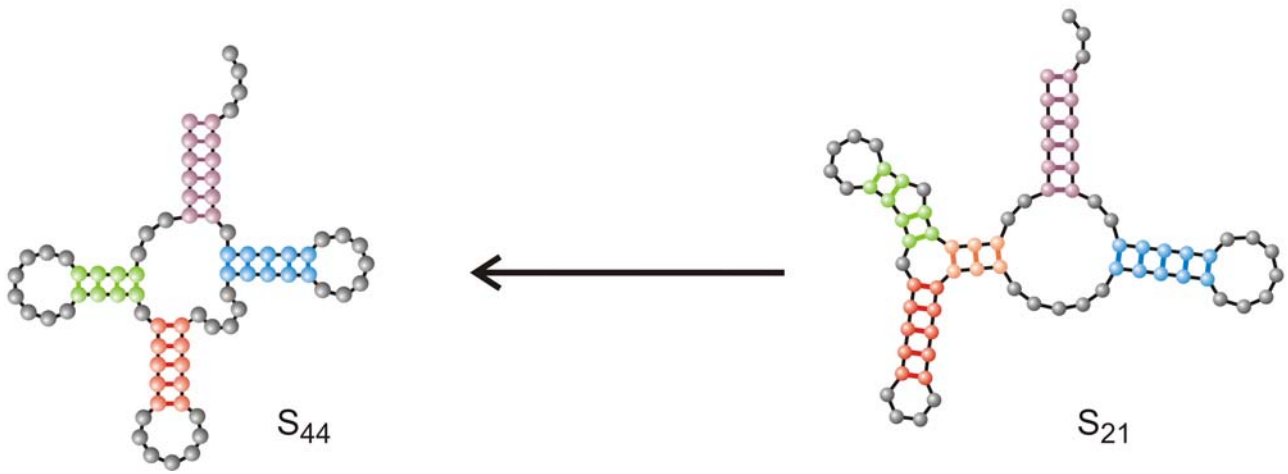
p = 0.001 / site × replication

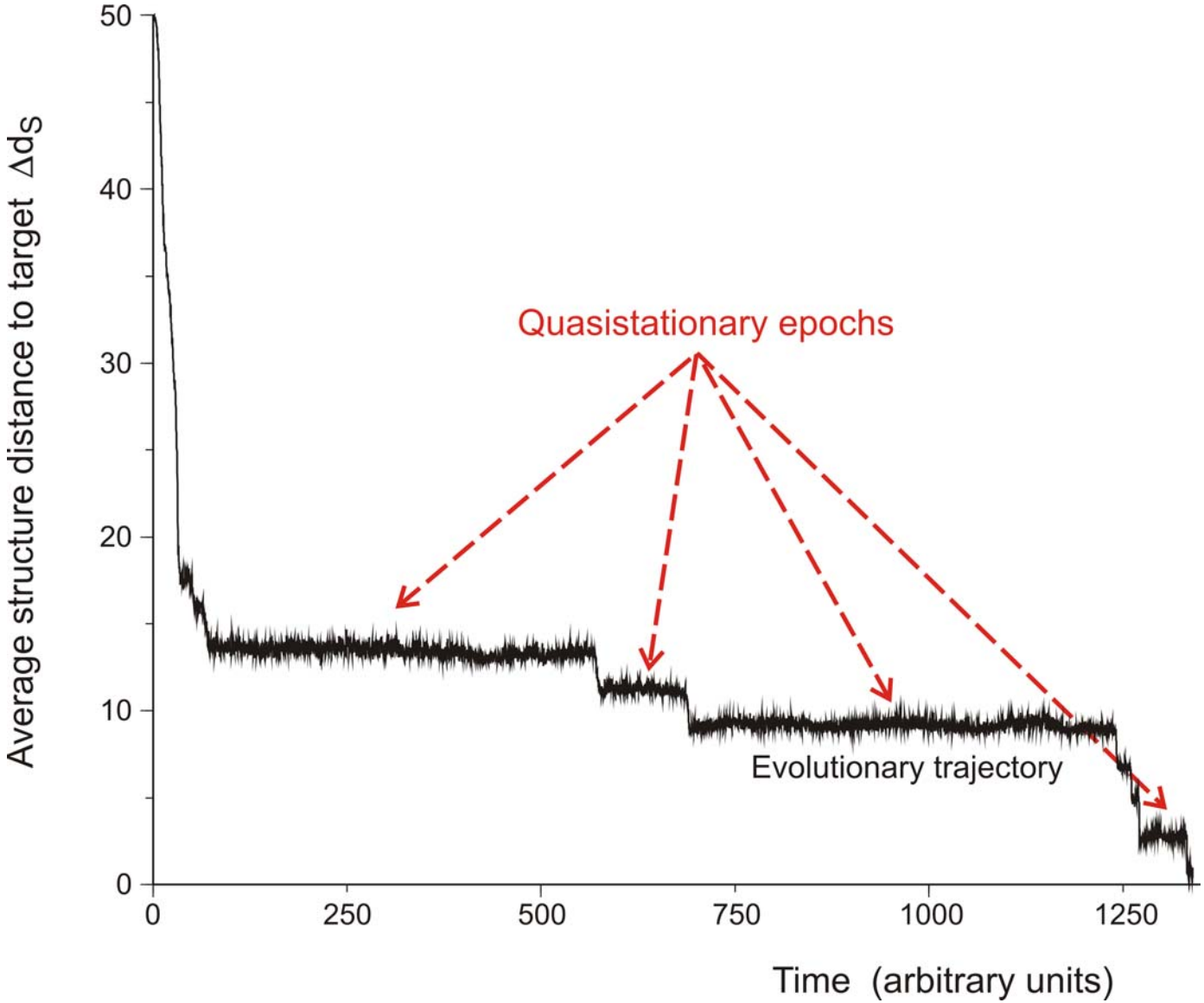
The flowreactor as a device for **studies** of evolution ***in vitro*** and ***in silico***

Randomly chosen
initial structure



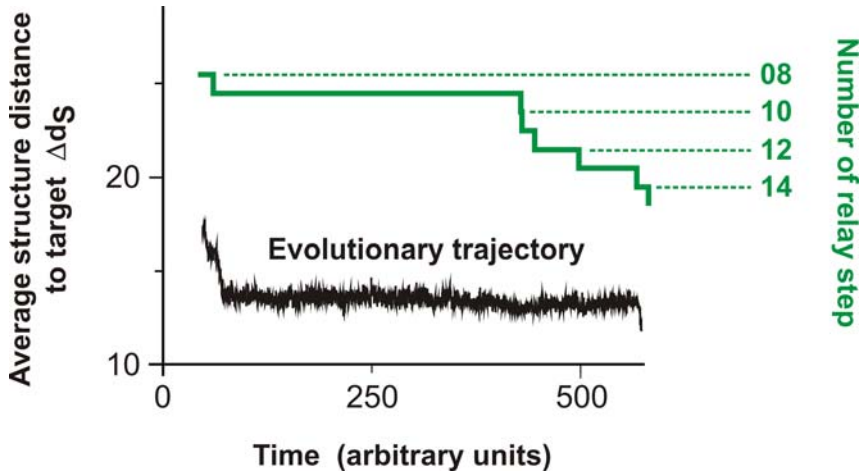
Phenylalanyl-tRNA
as target structure





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

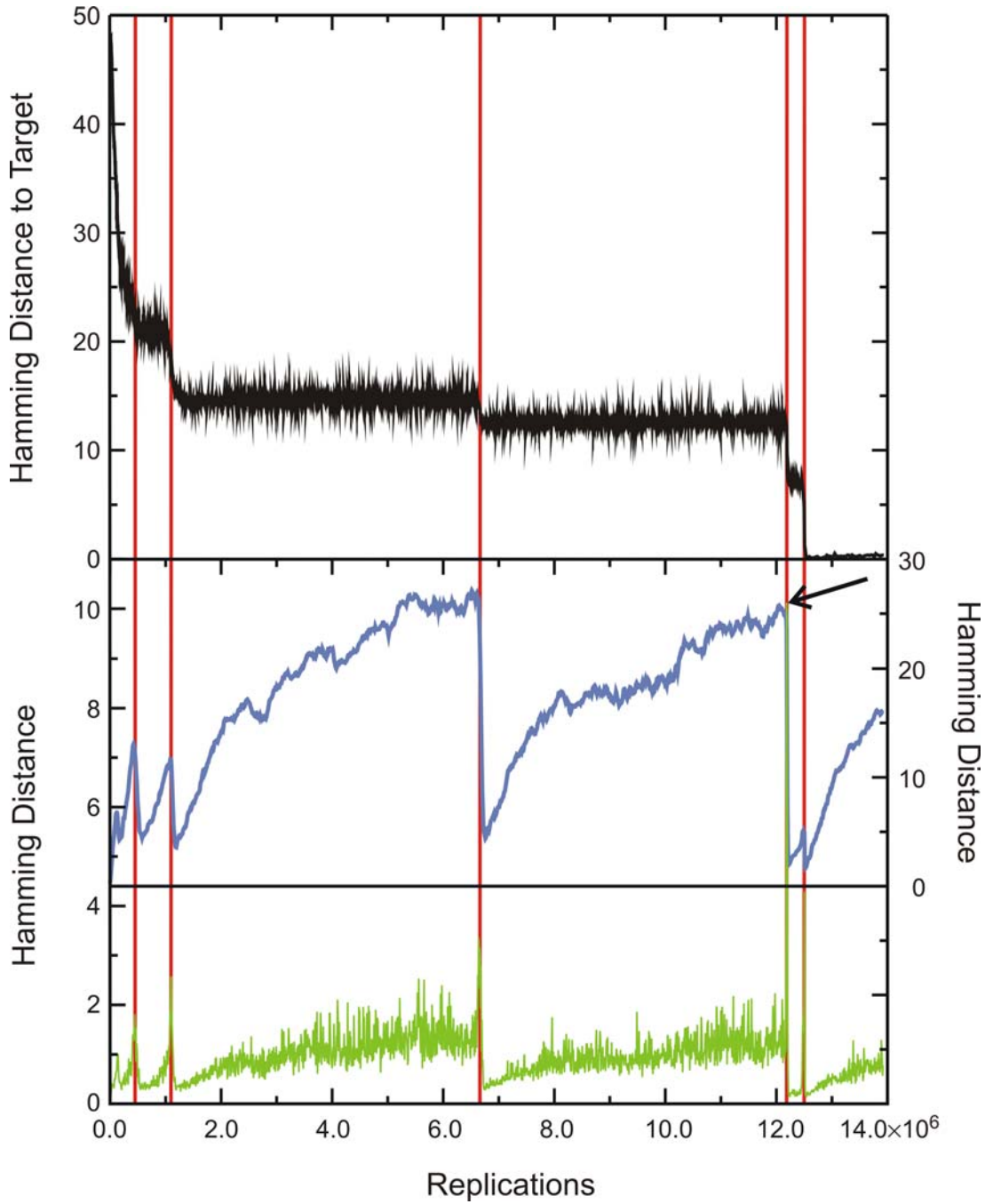
Neutral point mutations leave the molecular structure unchanged

Neutral genotype evolution during phenotypic stasis

Evolutionary trajectory

Spreading of the population on neutral networks

Drift of the population center in sequence space

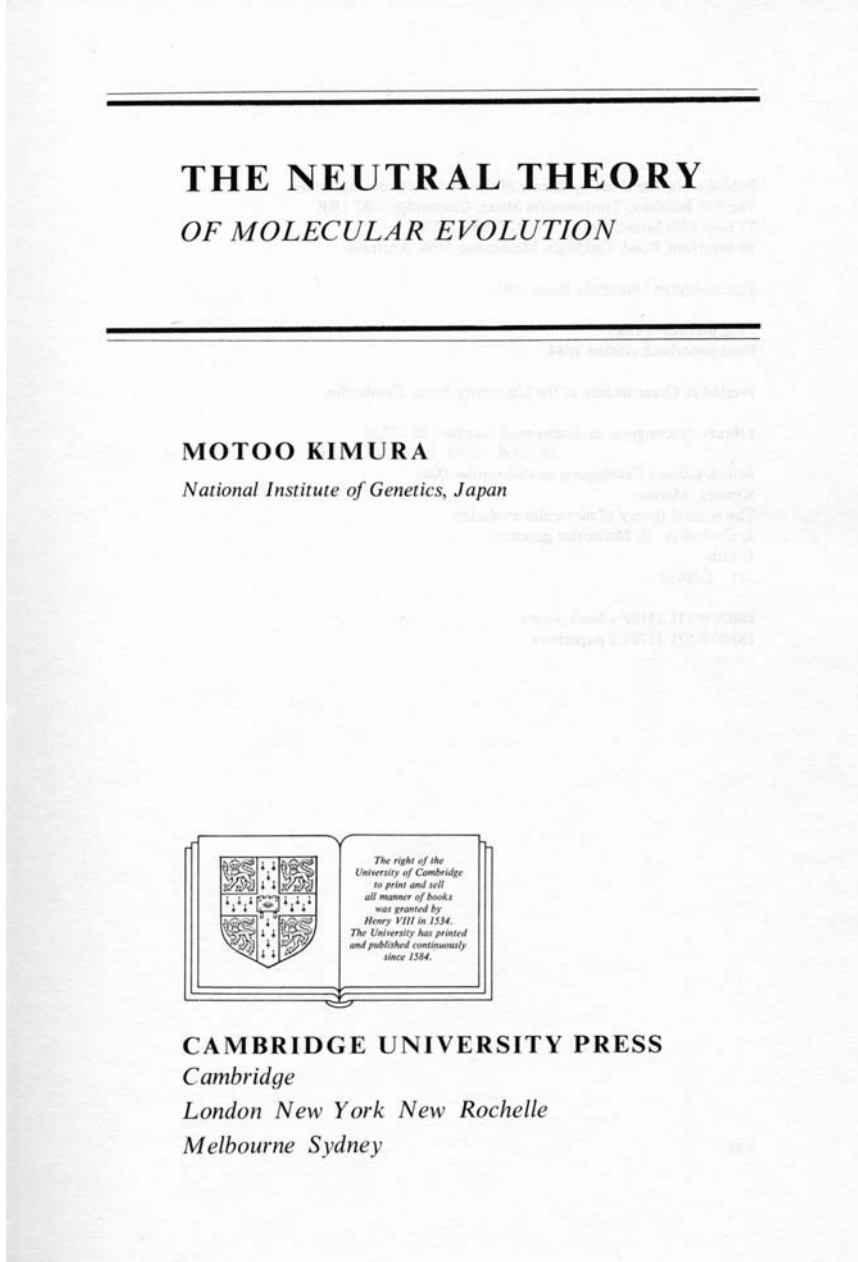


MOTOO KIMURA

Motoo Kimuras Populationsgenetik der neutralen Evolution.

Evolutionary rate at the molecular level. *Nature* **217**: 624-626, 1955.

The Neutral Theory of Molecular Evolution. Cambridge University Press. Cambridge, UK, 1983.



THE NEUTRAL THEORY

OF MOLECULAR EVOLUTION

MOTOO KIMURA

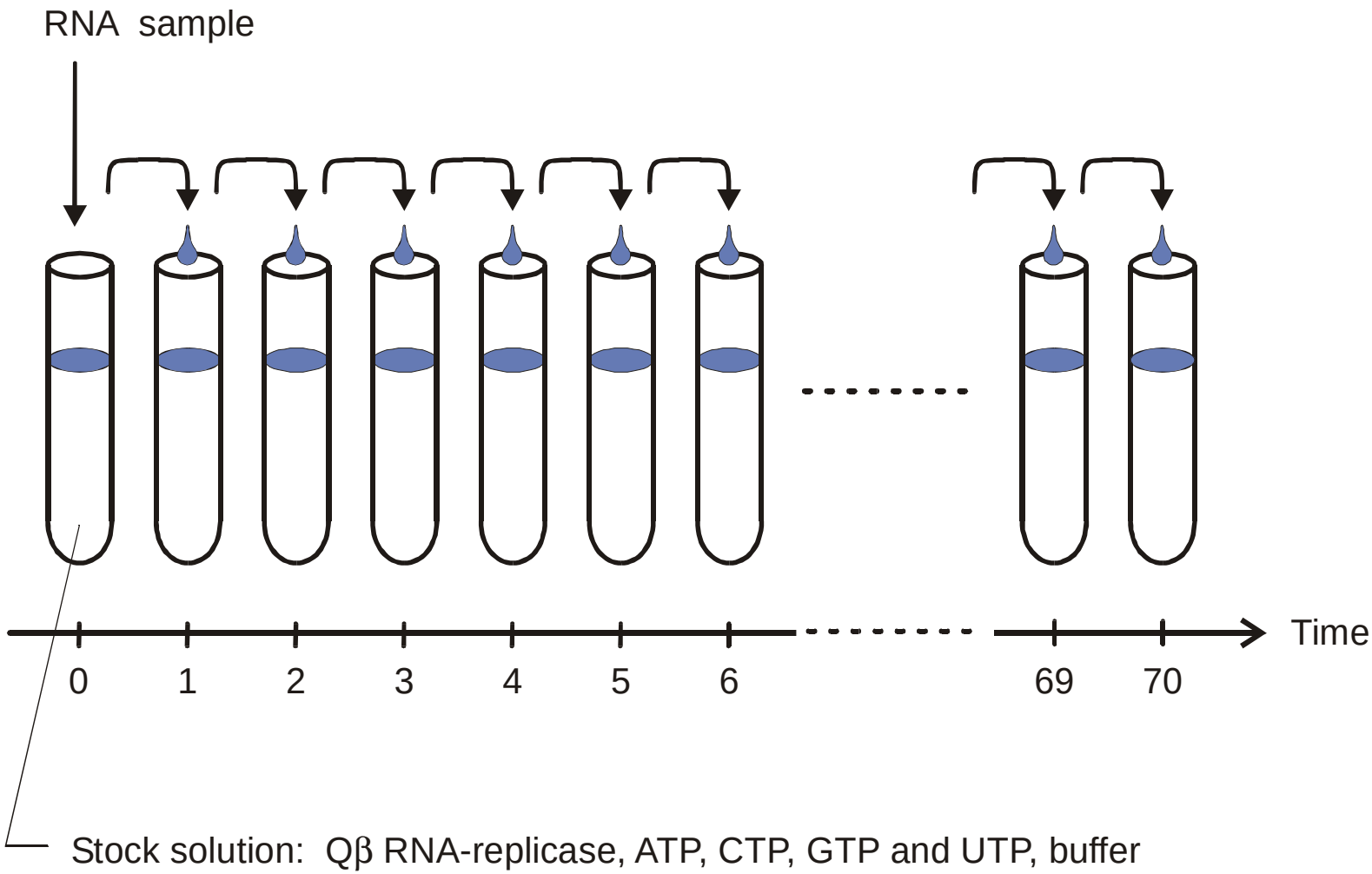
National Institute of Genetics, Japan

CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle
Melbourne Sydney

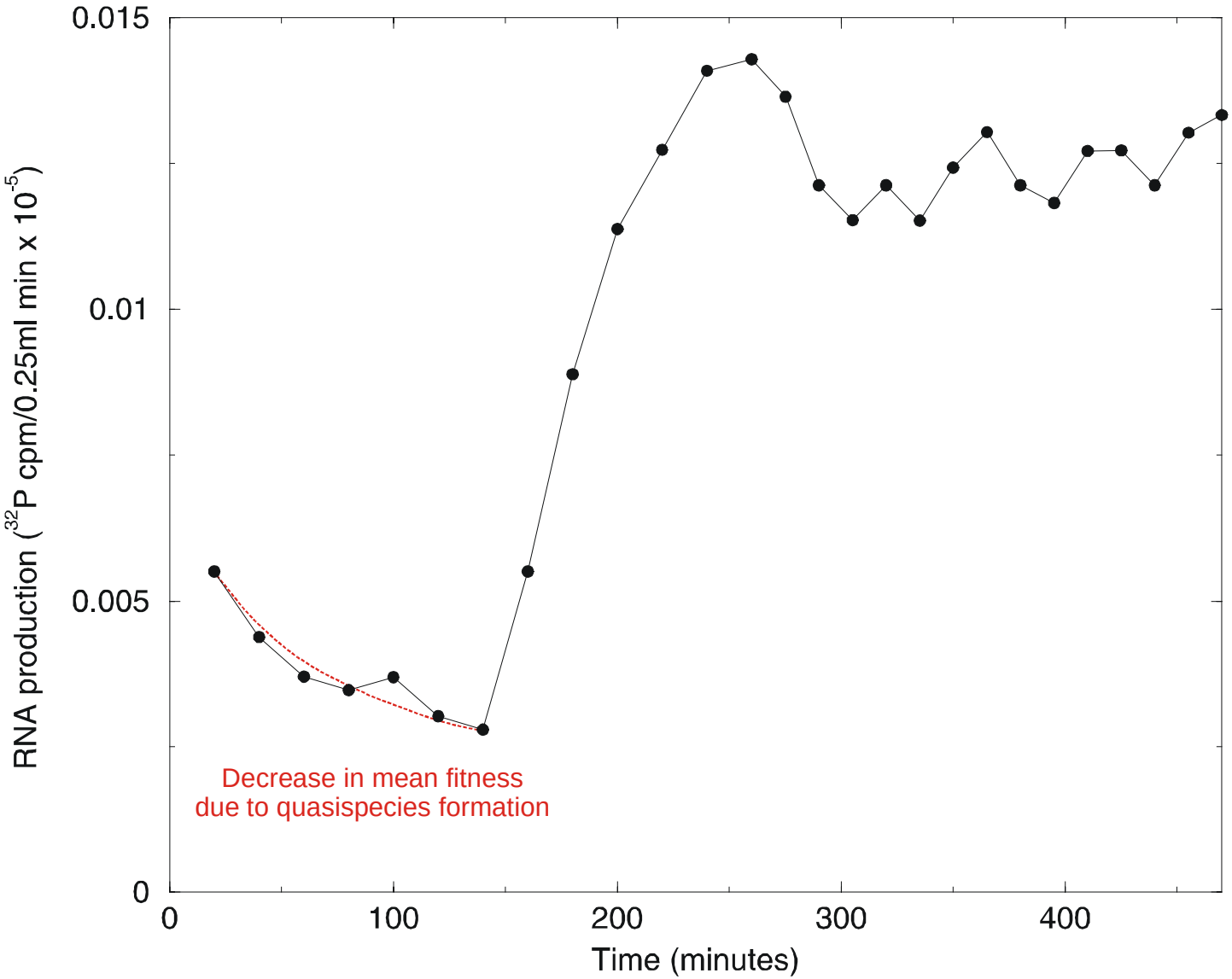
1. Was ist Leben?
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. **Evolutionsexperimente**
6. Die DNA + Protein Welt
7. Evolutionsmechanismen

	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 20 000 a	27 380 a 2×10^7 a	273 800 a 2×10^8 a

Time scales of evolutionary change

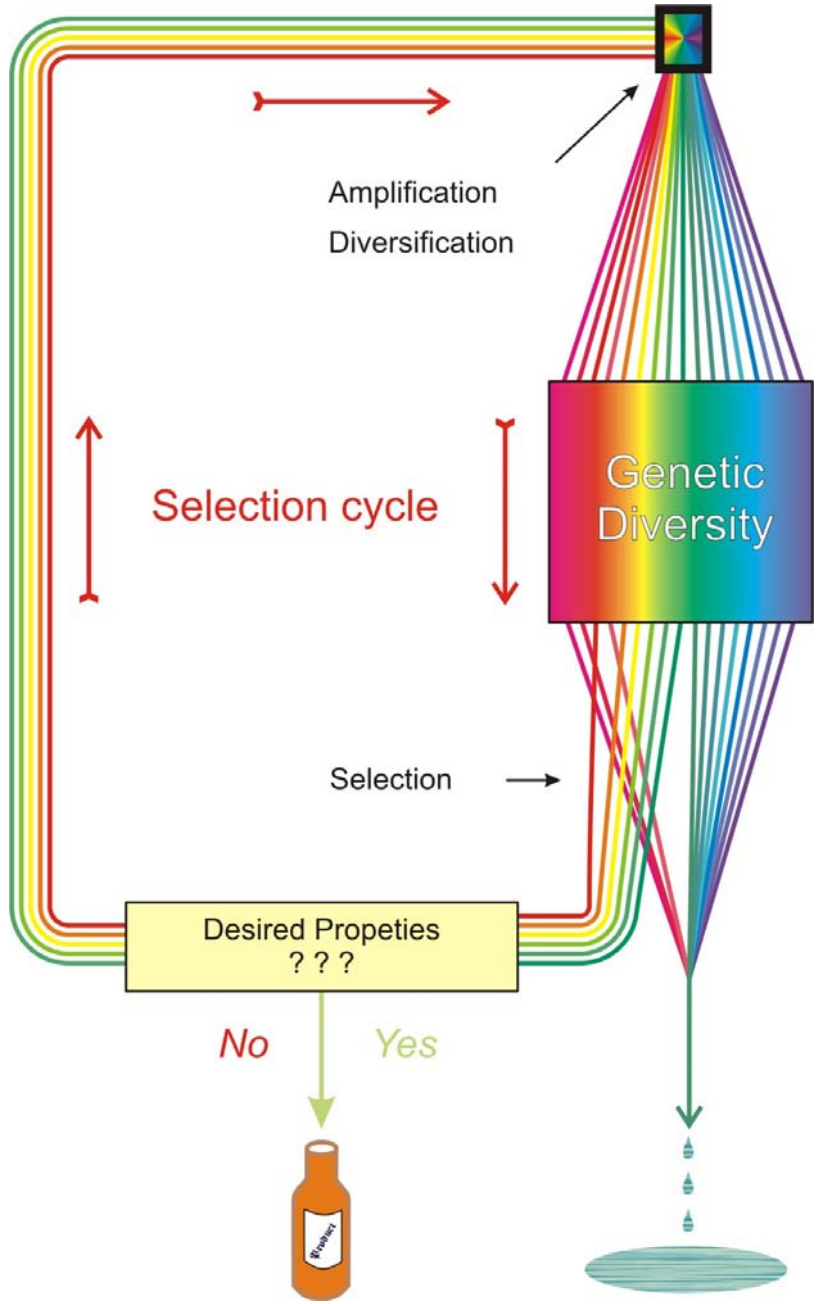


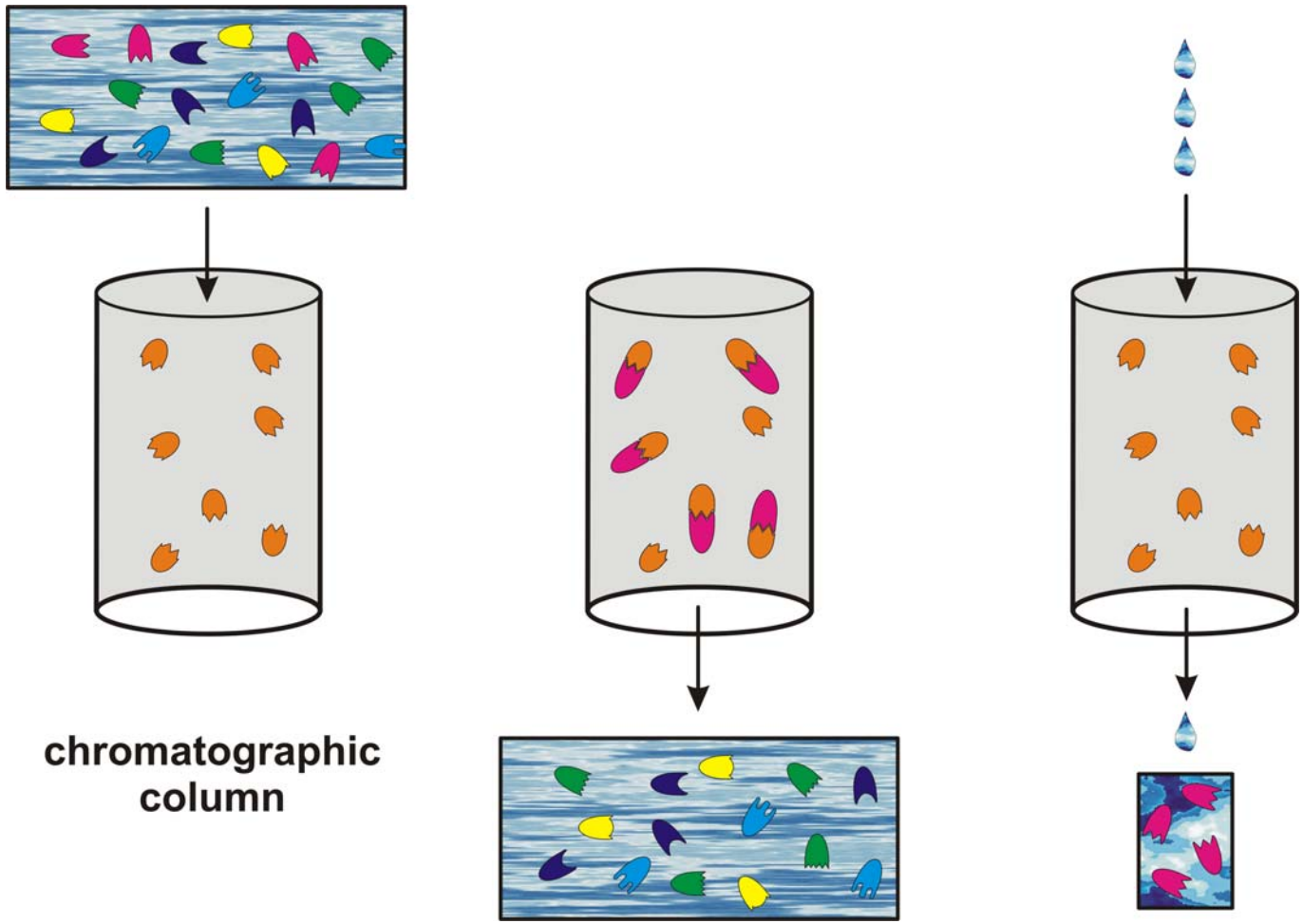
Anwendung der seriellen Überimpfungstechnik auf RNA-Evolution in Reagenzglas



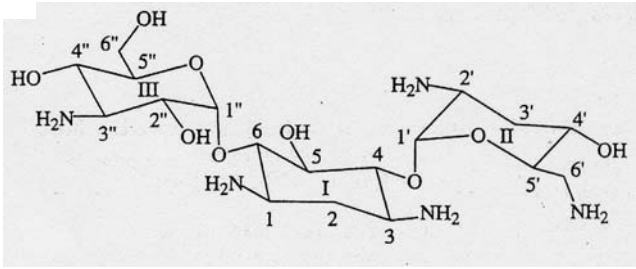
The increase in RNA production rate during a serial transfer experiment

Ein Beispiel für Selektion von Molekülen mit vorbestimmbaren Eigenschaften im Laborexpriment



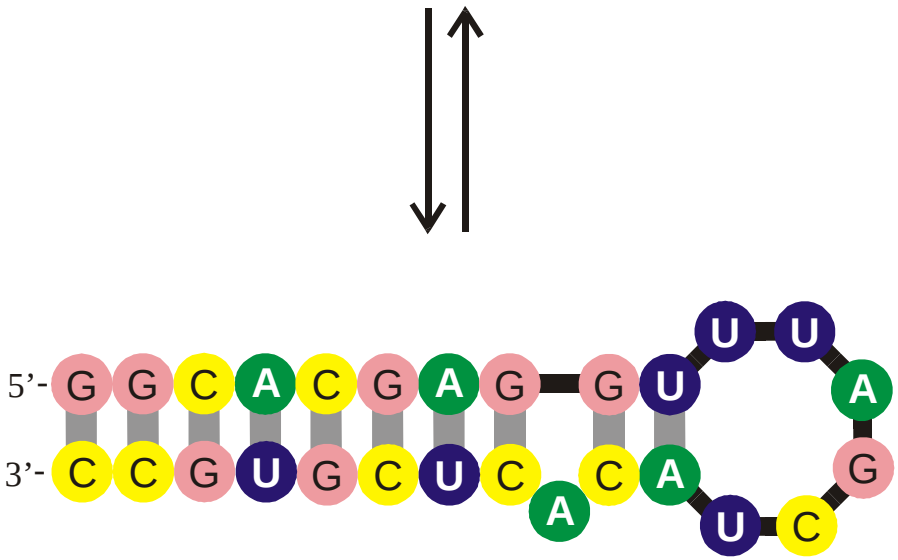


Die SELEX-Technik zur evolutionären Erzeugung von stark bindenden Molekülen



tobramycin

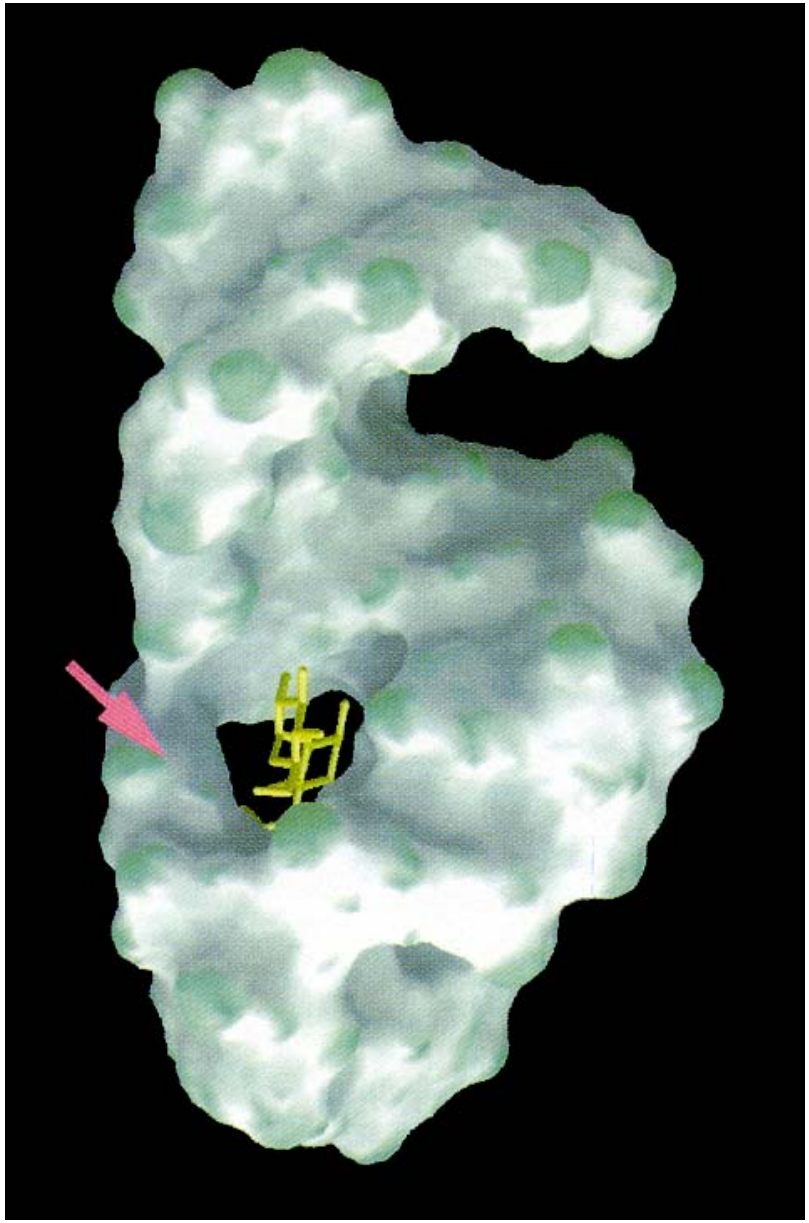
5’-GGCACGAGGUUUAGCUACACUCGUGCC-3’



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)

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*).

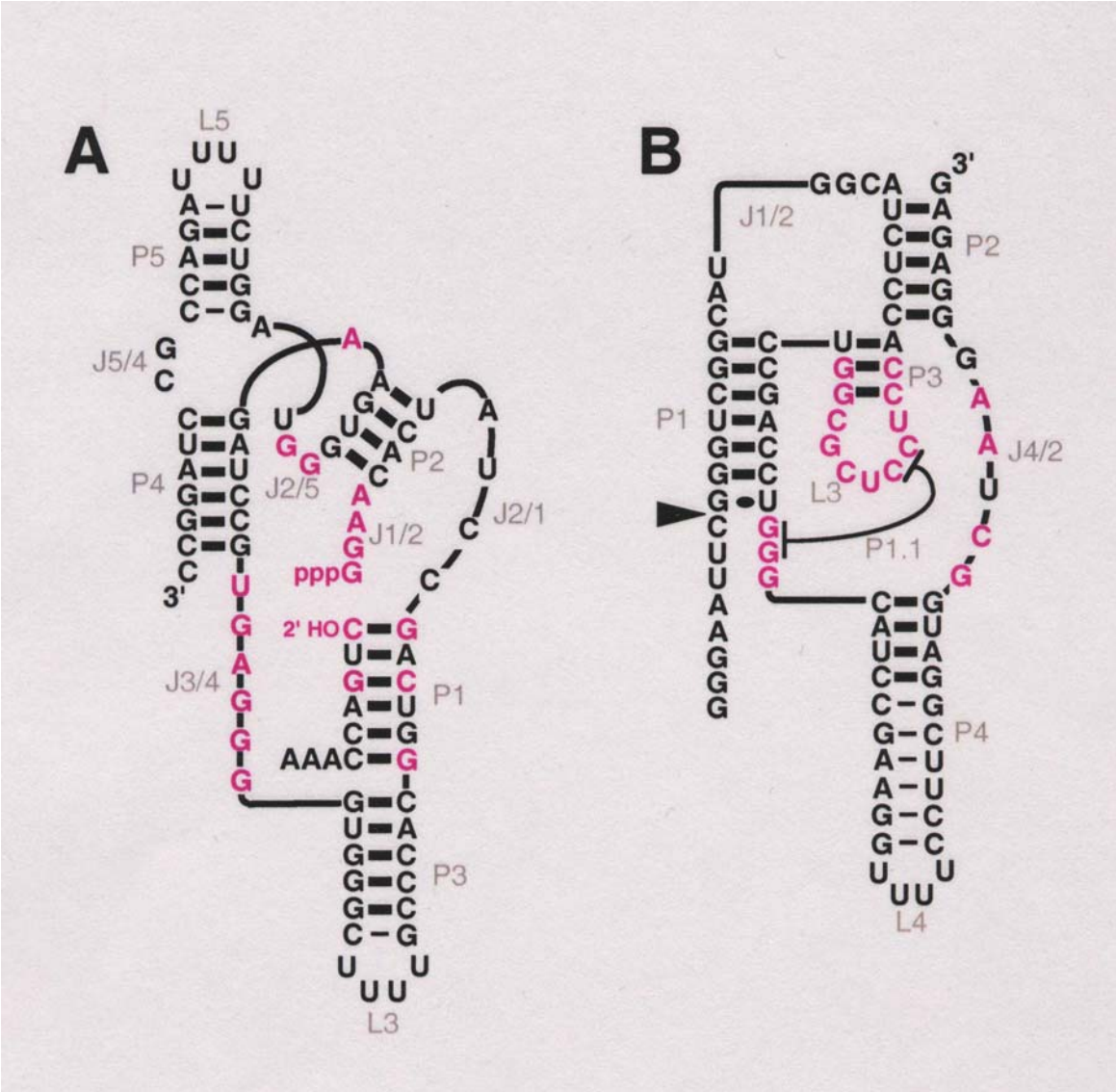
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect.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

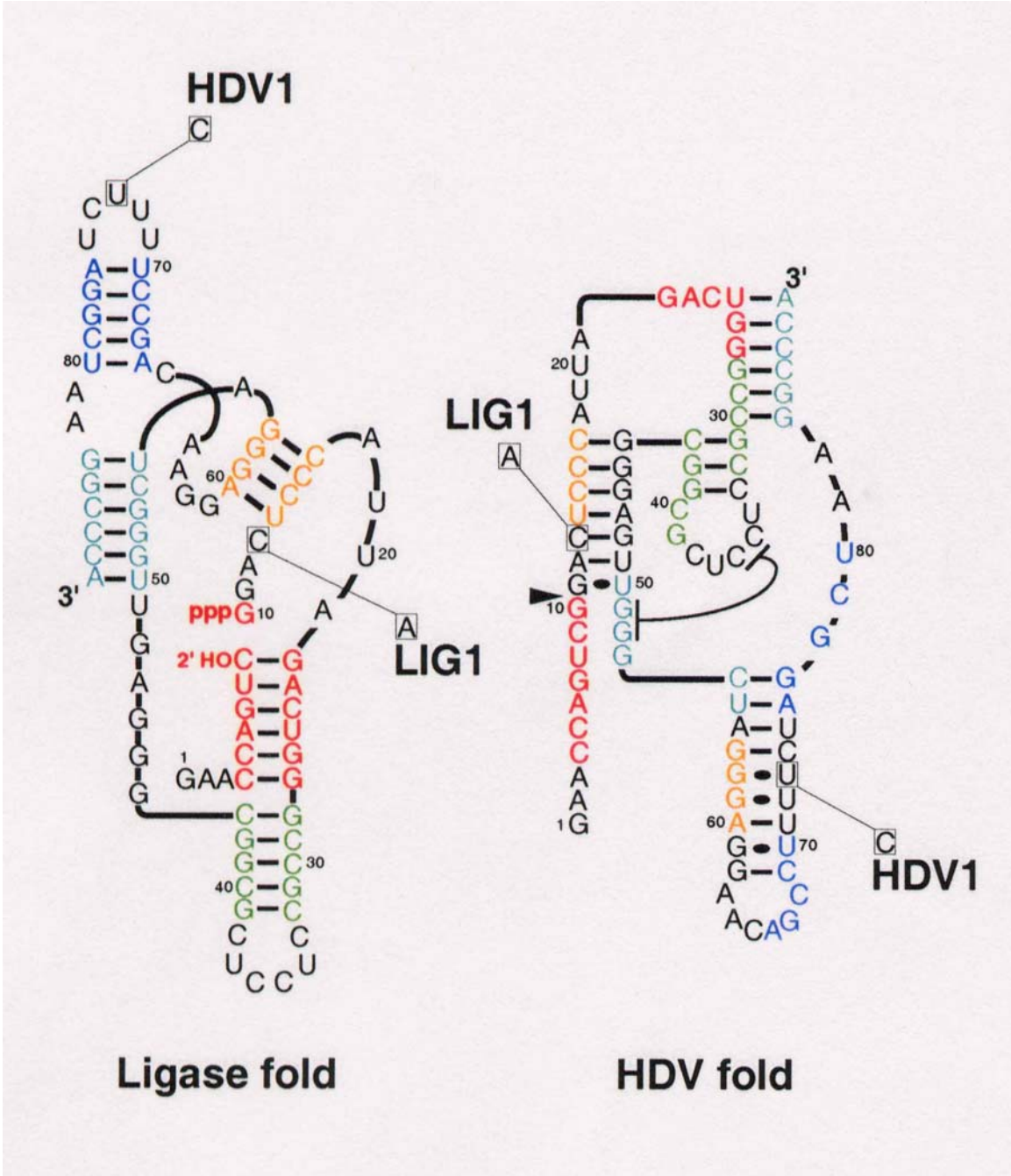
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

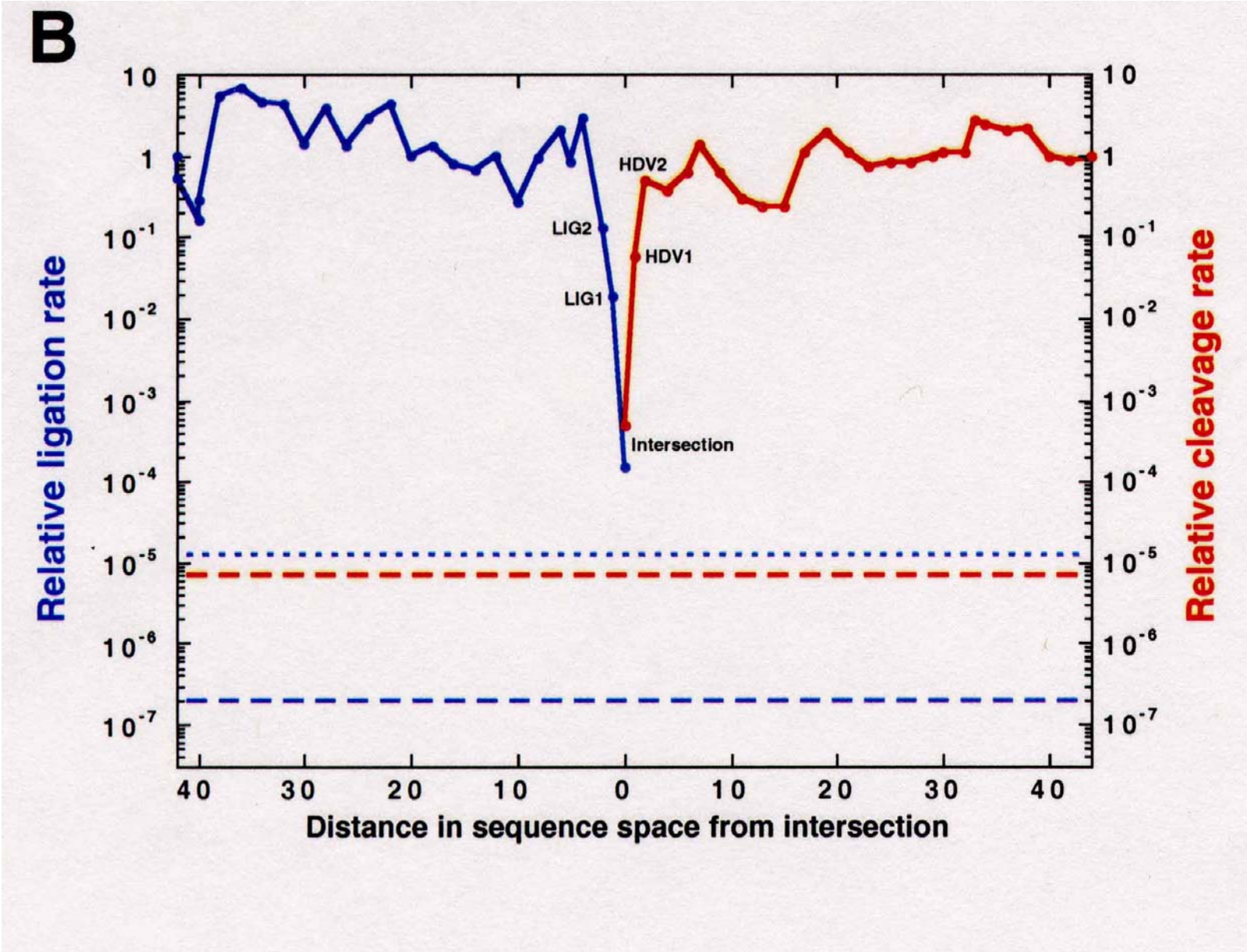


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

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

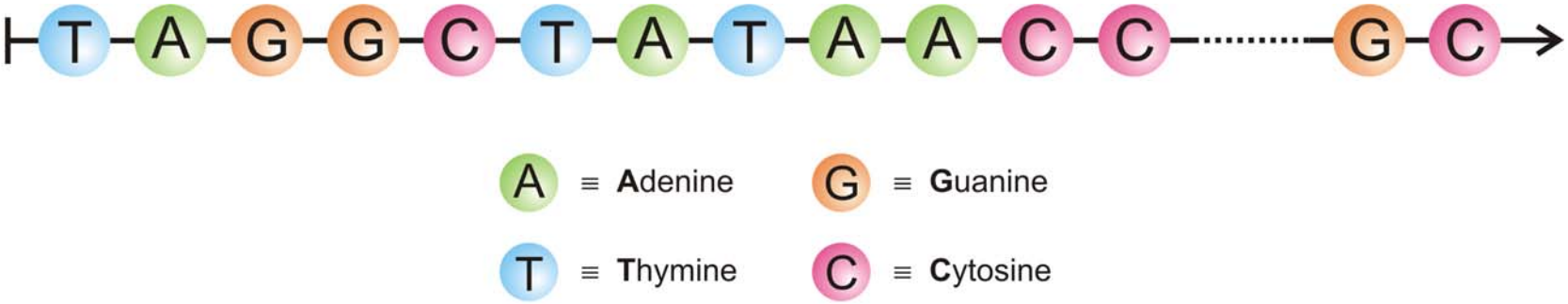
An RNA molecules which is 88 nucleotides long and can form both structures

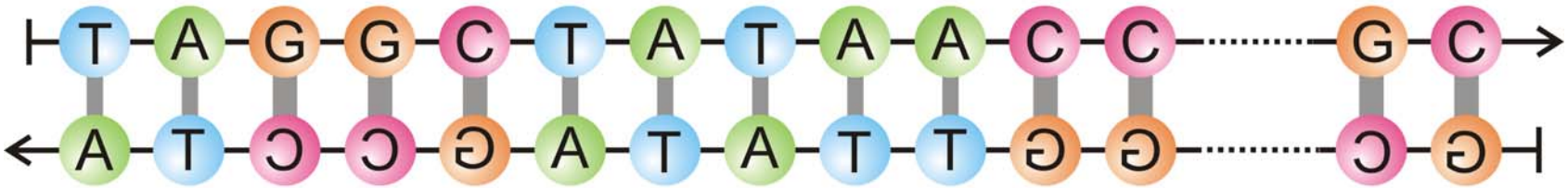




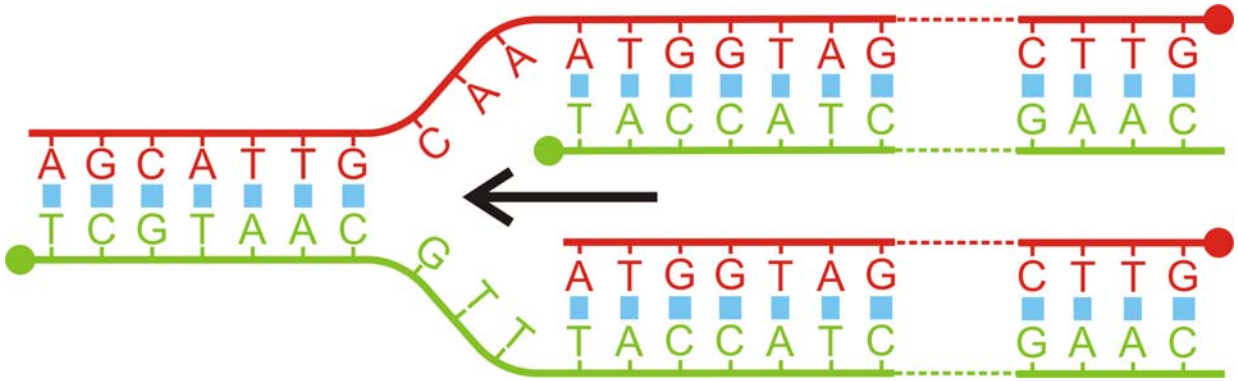
Two neutral walks through sequence space with conservation of structure and catalytic activity

1. Was ist Leben?
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. **Die DNA + Protein Welt**
7. Evolutionsmechanismen



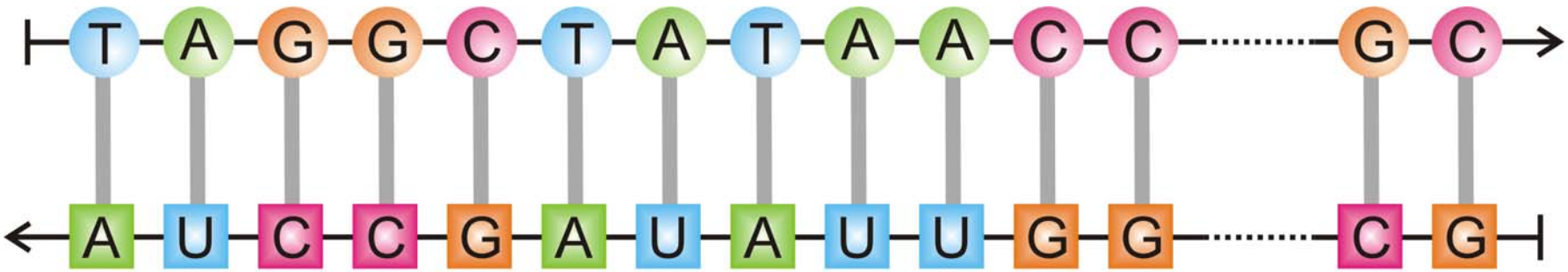


Deoxyribonucleic acid - DNA

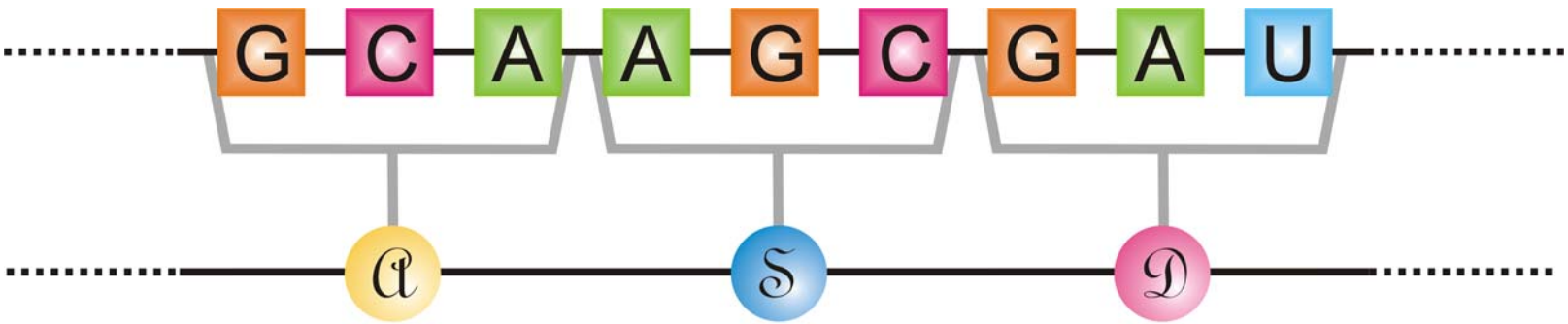


Die “Replikationsgabel”

Mechanismus der Replikation von doppelsträngigen DNA-Molekülen

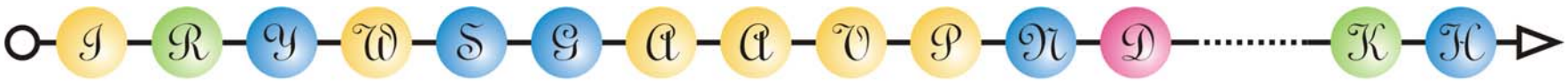


Transcription - DNA → RNA



Translation - RNA → Protein

Redundancy of the code: 4^3 = **64** codons versus **20** amino acids



𝔄 ≡ **ala**nine

𝔅 ≡ **cys**teine

𝔇 ≡ **asp**artic acid

𝔈 ≡ **glu**tamic acid

𝔉 ≡ **phe**nyl alanine

𝔊 ≡ **gly**cine

ℌ ≡ **his**tidine

ℑ ≡ **iso**l**eu**cine

𝔎 ≡ **lys**ine

𝔏 ≡ **leu**cine

𝔐 ≡ **met**hionine

𝔑 ≡ **as**paragi**n**e

𝔓 ≡ **pro**line

𝔔 ≡ **glu**tami**n**e

ℜ ≡ **arg**inine

𝔖 ≡ **ser**ine

𝔗 ≡ **thr**eonine

𝔙 ≡ **val**ine

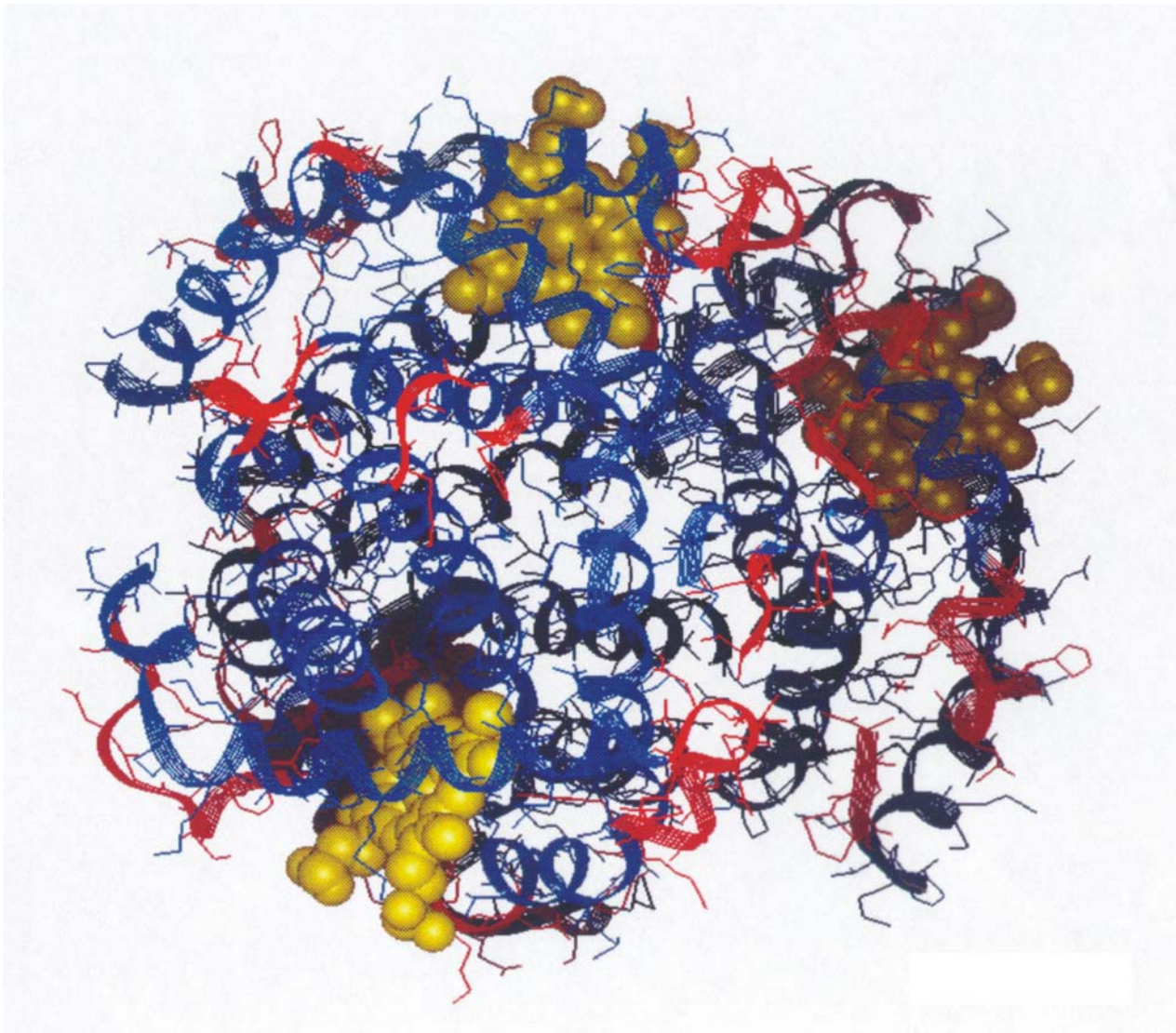
𝔚 ≡ **tryp**tophane

𝔜 ≡ **tyr**osine

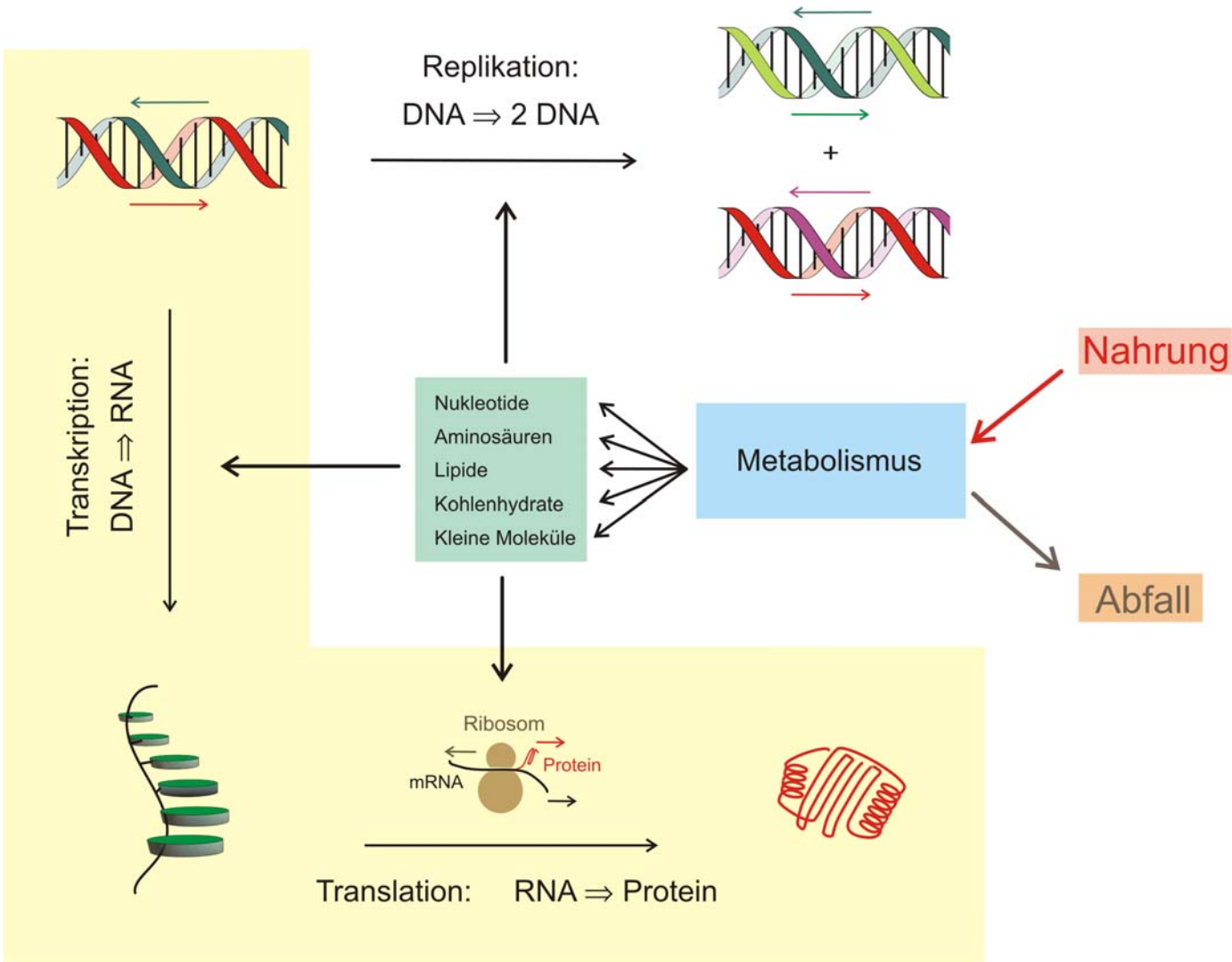
Protein

Max F. Perutz 1914-2002

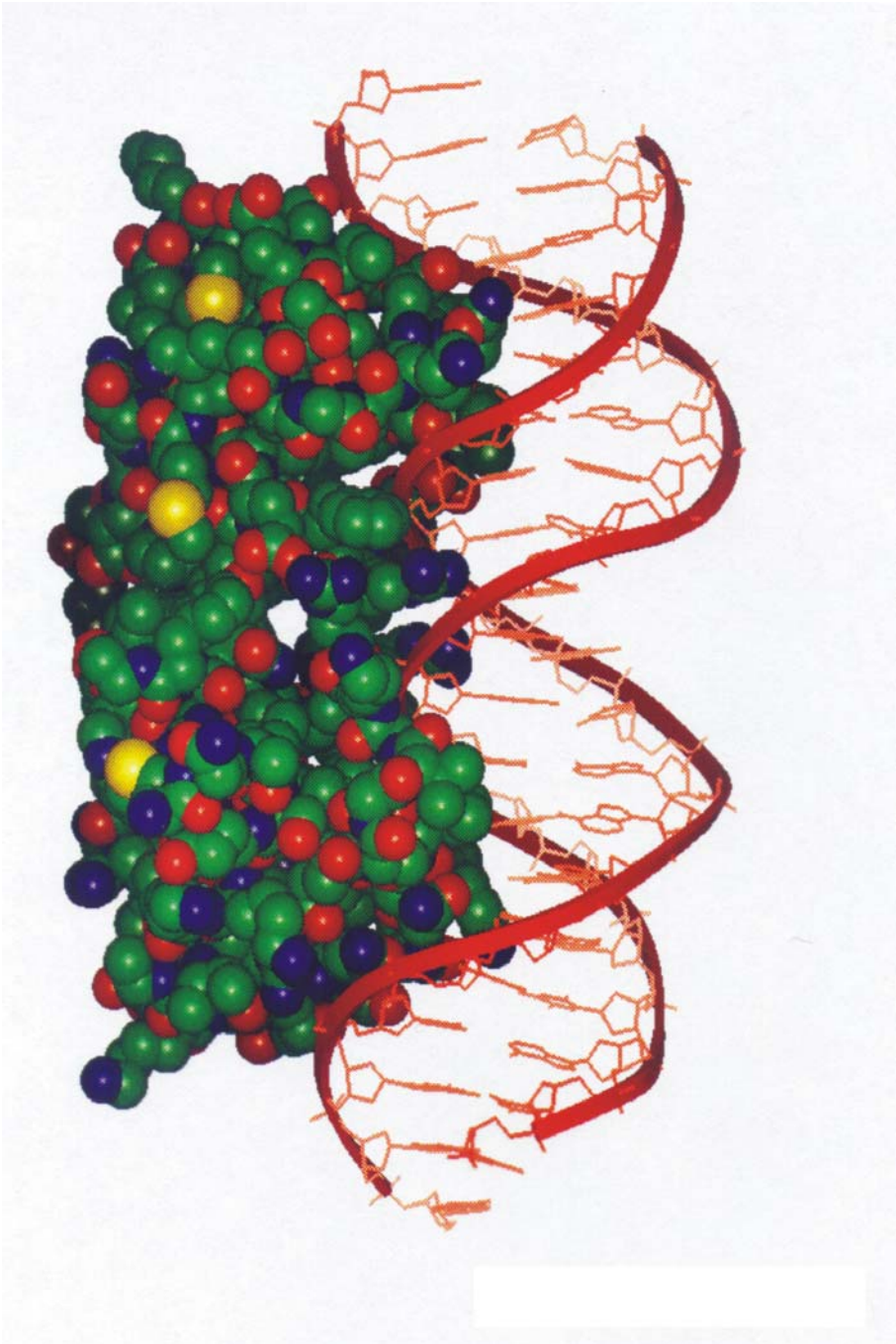
Nobel prize 1962



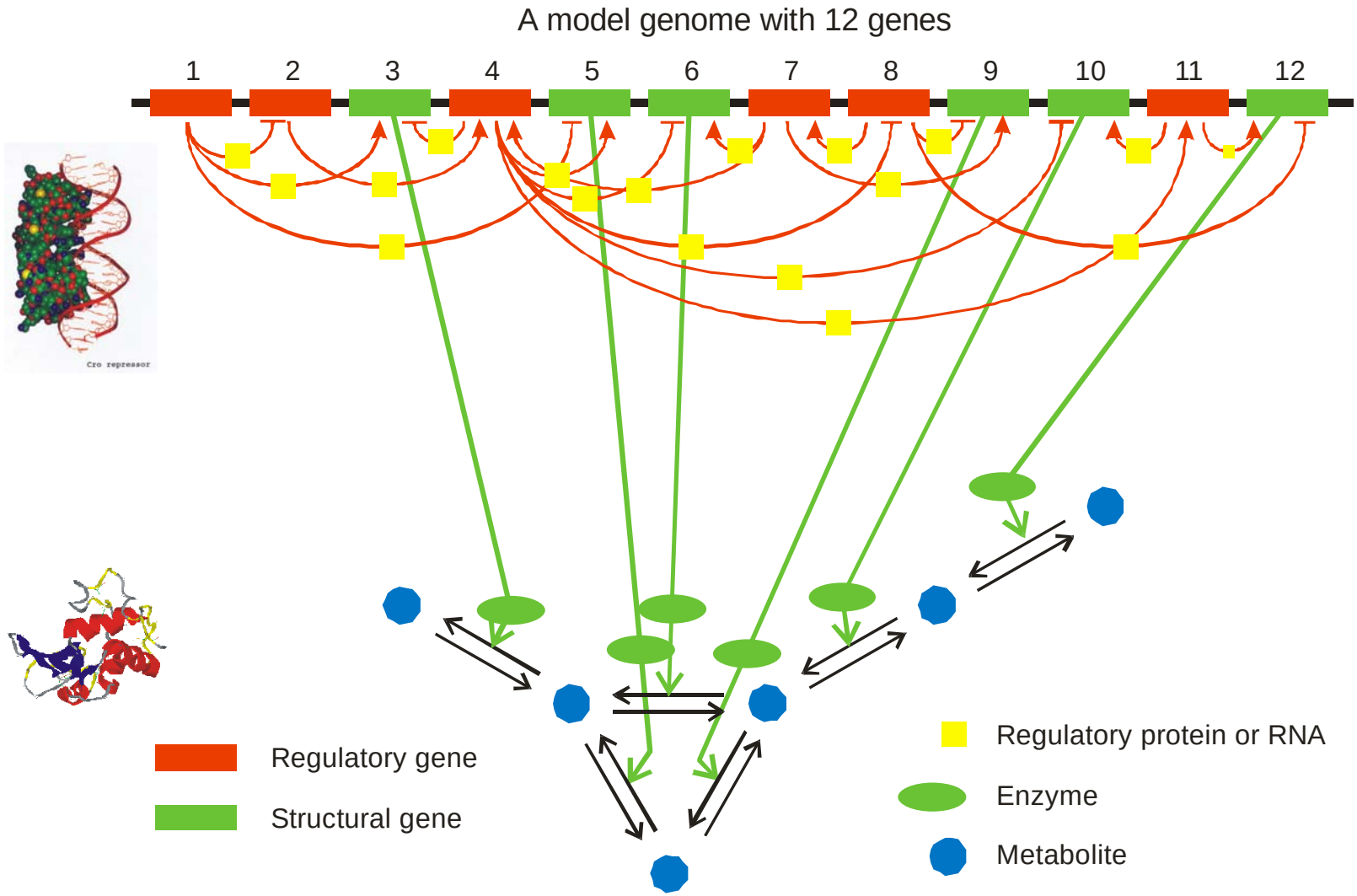
Three-dimensional
structure of hemoglobin



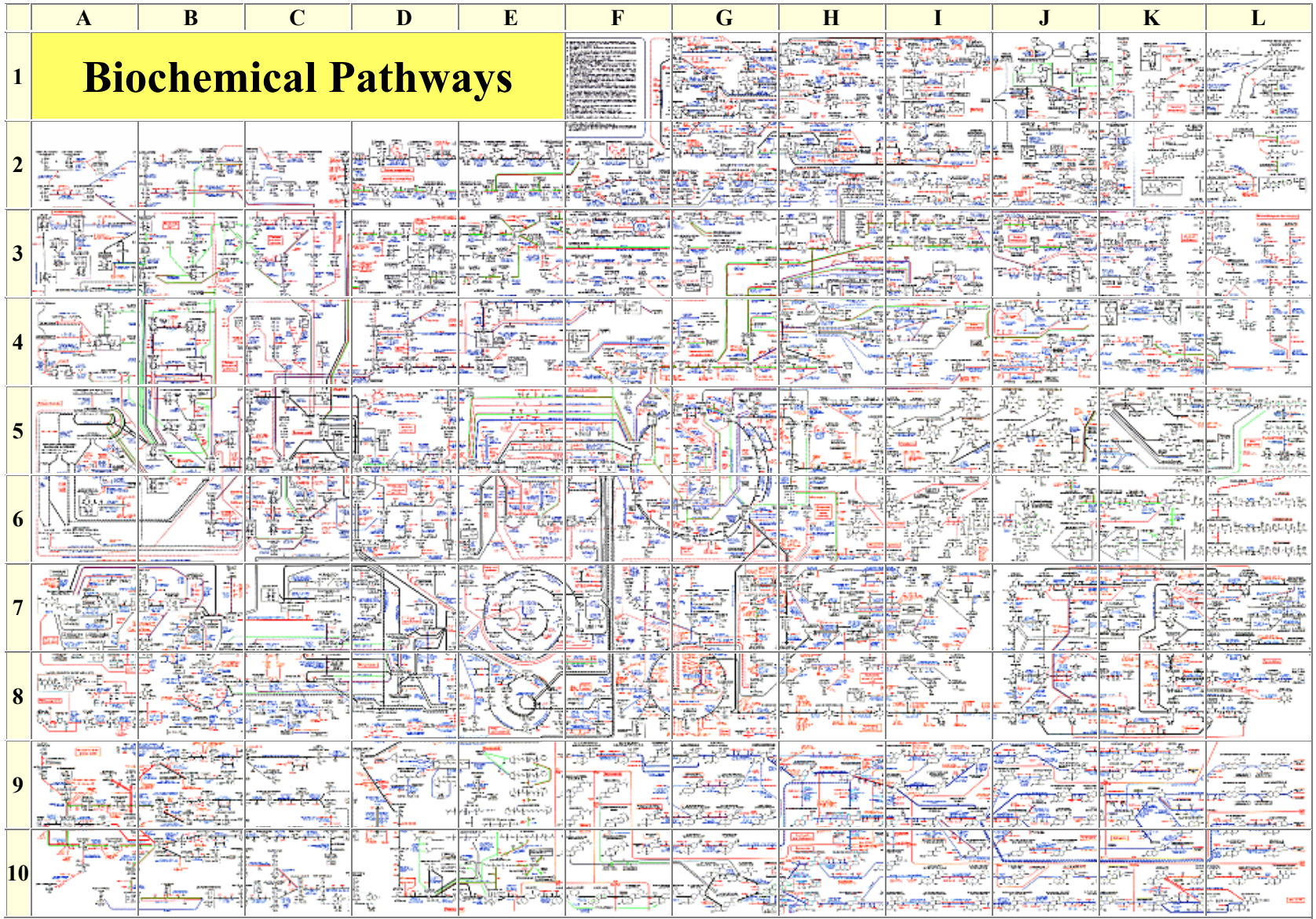
Skizze des zellulären Stoffwechsels



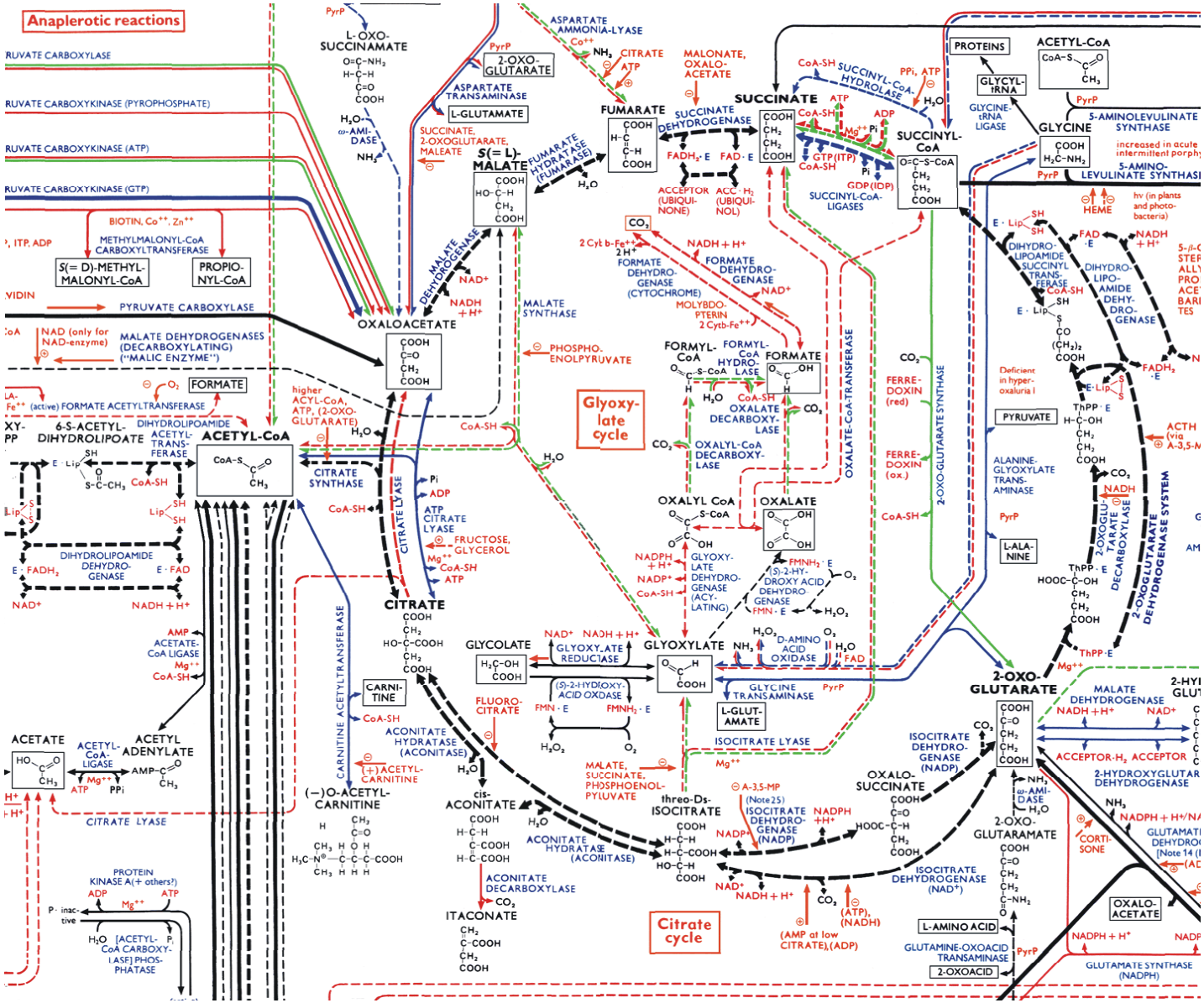
Three-dimensional structure of the complex between a specific DNA binding site and the regulatory protein cro-repressor



Skizze eines einfachen genetisch-metabolischen Regulationsnetzwerkes

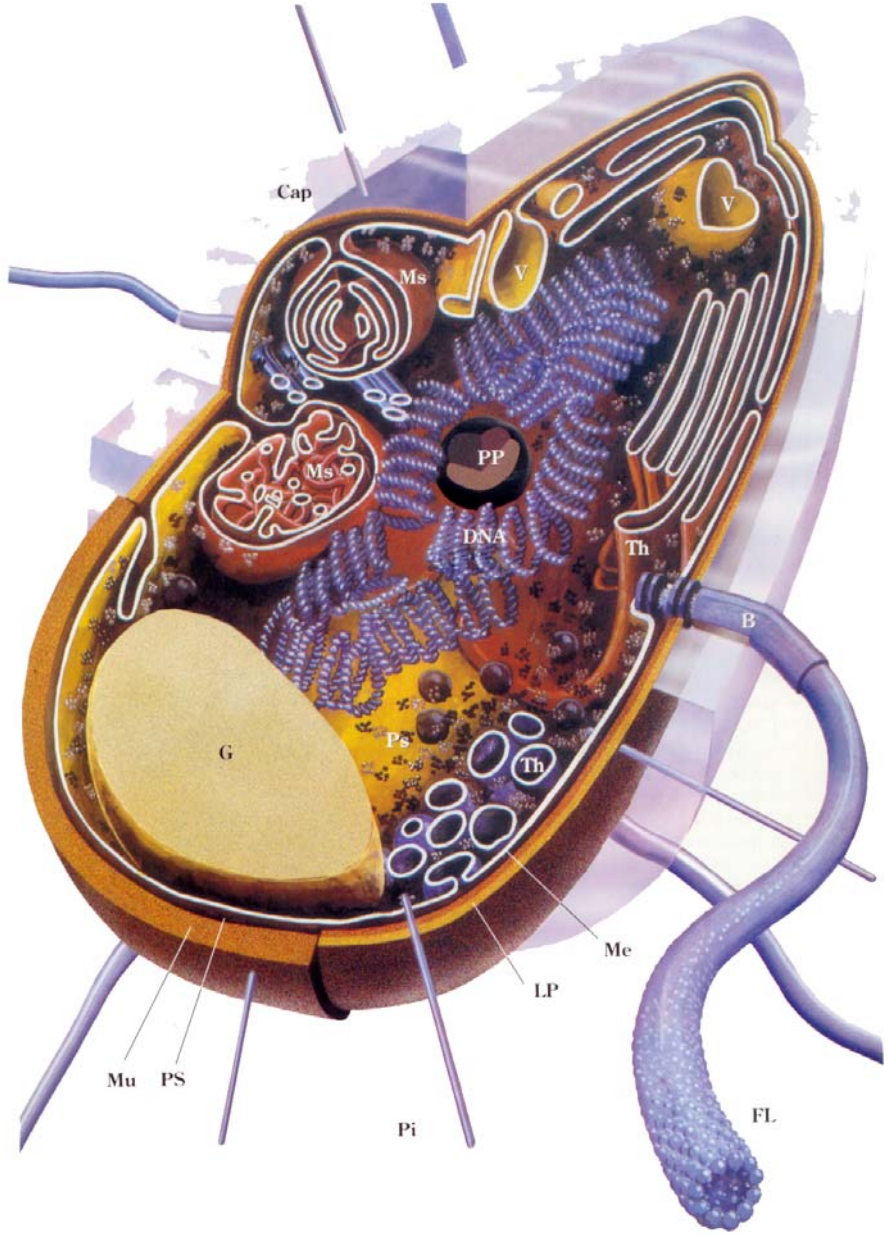


Das Reaktionsnetzwerk des zellulären Stoffwechsels publiziert von Boehringer-Ingelheim.



Der Zitronensäure- oder Krebszyklus (vergrößert aus der vorigen Abbildung).

Die Bakterienzelle als ein Beispiel für die einfachste Form autonomen Lebens.



Der menschliche Körper:

10^{14} Zellen =
10^{13} eukaryotische Zellen +
$\approx 9 \times 10^{13}$ prokaryotische Bakterienzellen

80 kg eukaryotische Zellen +
800 g Bakterienzellen

$\approx$ 200 eukaryotische Zelltypen

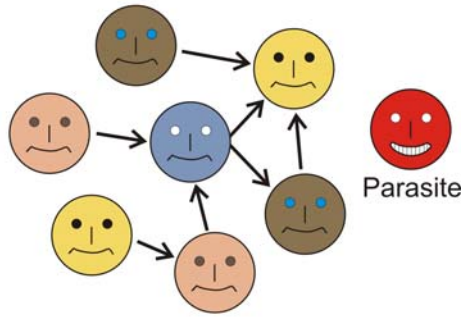
1. Was ist Leben?
2. Chemische Evolution
3. Der Ursprung biologischer Information
4. Darwinsche Evolution mit Molekülen
5. Evolutionsexperimente
6. Die DNA + Protein Welt
7. **Evolutionsmechanismen**

Die großen Evolutionsschritte (nach John Maynard Smith und Eörs Szathmáry)

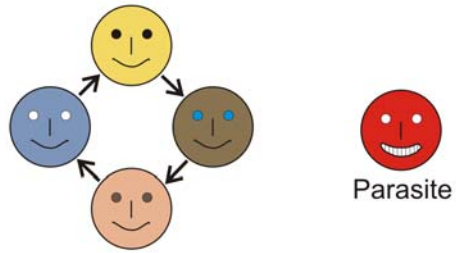
	Membranen, organisierte Teilung		
Replizierende Moleküle	⇒	Moleküle in Kompartments	
	Molekülverkettung, gemeinsame Replikation		
Unabhängige Replikatoren	⇒	Chromosomen	
	genetischer Code, Ribosom		
RNA als Gen und Enzyme	⇒	DNA und Protein	
	Zusammenschluß durch Endosymbiose		
Prokaryoten	⇒	Eukaryoten	
	Ursprung der sexuellen Vermehrung		
Asexuell vermehrende Klone	⇒	Sexuell vermehrende Populationen	
	Zelldifferenzierung und Entwicklung		
Protisten	⇒	Pflanzen, Pilze und Tiere	
	Entstehung nicht-reproduktiver Kasten		
Einzeln lebende Individuen	⇒	Tierkolonien	
	Sprache, Schrift, Kultur, …		
Primatengesellschaften	⇒	menschliche Gesellschaften	

Stufe I:
Unabhängige Replikatoren
in Konkurrenz

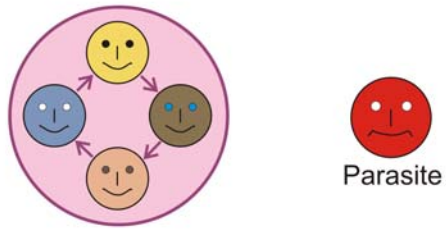
Stufe II:
Katalyse und Konkurrenz
bei der Replikation



Stufe III:
Funktionell verknüpfte
Replikatoren



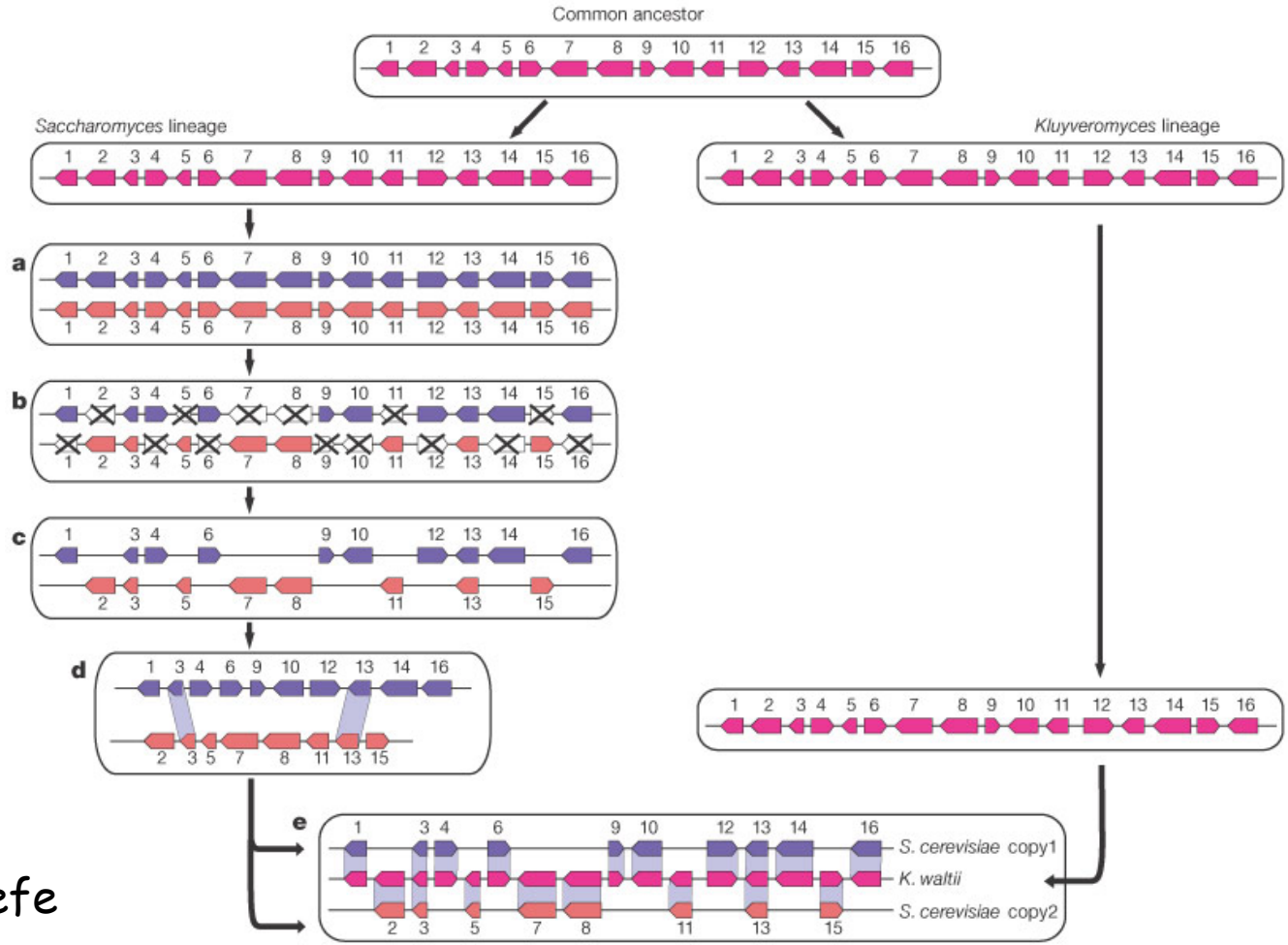
Stufe IV:
Neue Einheit der
Selektion



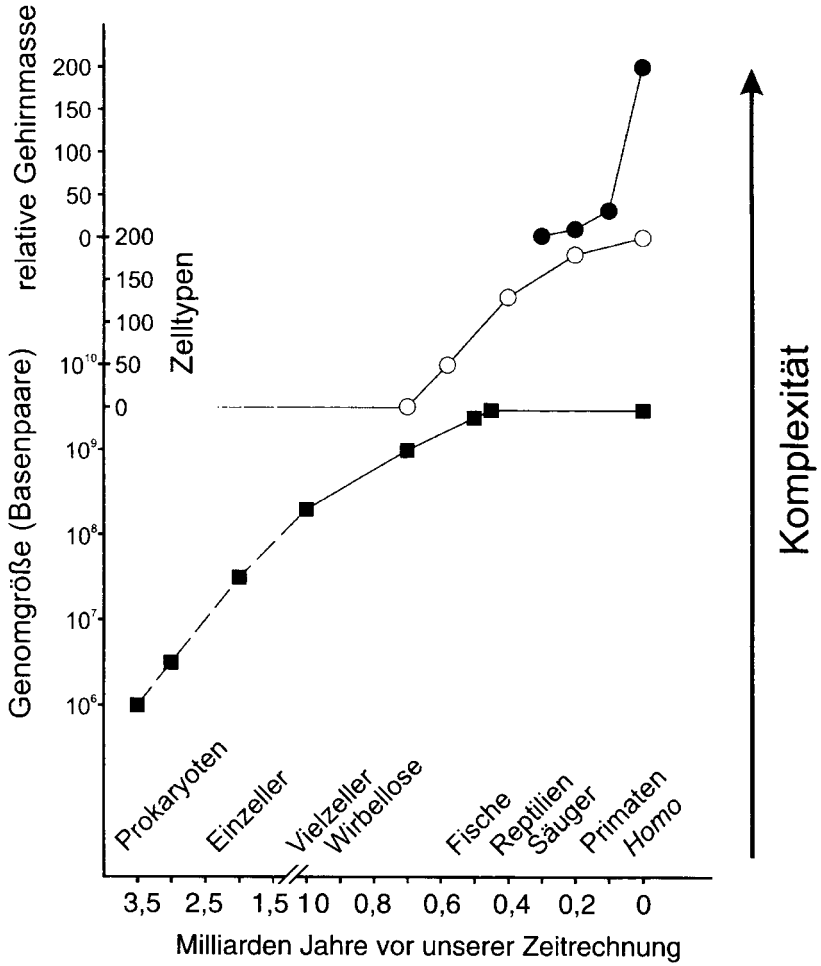
Stufe V:
Unabhängige Einheiten
in Konkurrenz

Ein Mechanismus zur Überwindung
hierarchischer Stufen in der Evolution
(nach Manfred Eigen und Peter Schuster)

Ein Modell für die Genverdopplung in Hefe vor ≈ 1×10^8 Jahren



Manolis Kellis, Bruce W. Birren, and Eric S. Lander. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**: 617-624, 2004



4.10 Die Zunahme der Komplexität ist ein wesentlicher Aspekt der biologischen Evolution, wobei höhere Komplexität sowohl durch Vergrößerung der Zahl von miteinander in Wechselwirkung stehenden Elementen als auch durch Differenzierung der Funktionen dieser Elemente entstehen kann. In dieser Abbildung wird zwischen drei Phasen oder Strategien der Evolution von Komplexität unterschieden. *Untere Kurve*: Zunahme der Genomgröße; logarithmische Auftragung der Zahl der Basenpaare im Genom von Zellen seit Beginn der biologischen Evolution (Daten aus Abbildung 2.3). *Mittlere Kurve*: Zunahme der Zahl der Zelltypen in der Evolution der Metazoa (Daten aus Abbildung 4.8). *Obere Kurve*: Zunahme des relativen Gehirngewichts (bezogen auf die Körperoberfläche) bei Säugetieren (Daten aus Wilson 1985). Für die Abszisse wurden zwei Skaleneinteilungen verwendet, eine für den Zeitraum $>10^9$ Jahre, eine andere für den Zeitraum $<10^9$ Jahre vor der Gegenwart. Oberhalb der Abszisse sind die Namen einiger wichtiger taxonomischer Einheiten angeführt, deren Evolution in etwa beim jeweiligen Wortbeginn einsetzt.

Wolfgang Wieser. *Die Erfindung der Individualität oder die zwei Gesichter der Evolution*. Spektrum Akademischer Verlag, Heidelberg 1998.

A.C.Wilson. *The Molecular Basis of Evolution*. Scientific American, Oct.1985, 164-173.

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