Origin of life and early evolution in the light of present day molecular biology

Peter Schuster

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

Theory of Evolution and the Belief in Creation

Wien, 23.– 26.02.2010

Web-Page for further information:

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

Peter Schuster. ***Evolution and design. The Darwinian theory of evolution is a scientific fact and not an ideology.*** Complexity **11**(1):12-15, 2006

Peter Schuster. ***Evolution und Design. Versuch einer Bestandsaufnahme der Evolutionstheorie***.

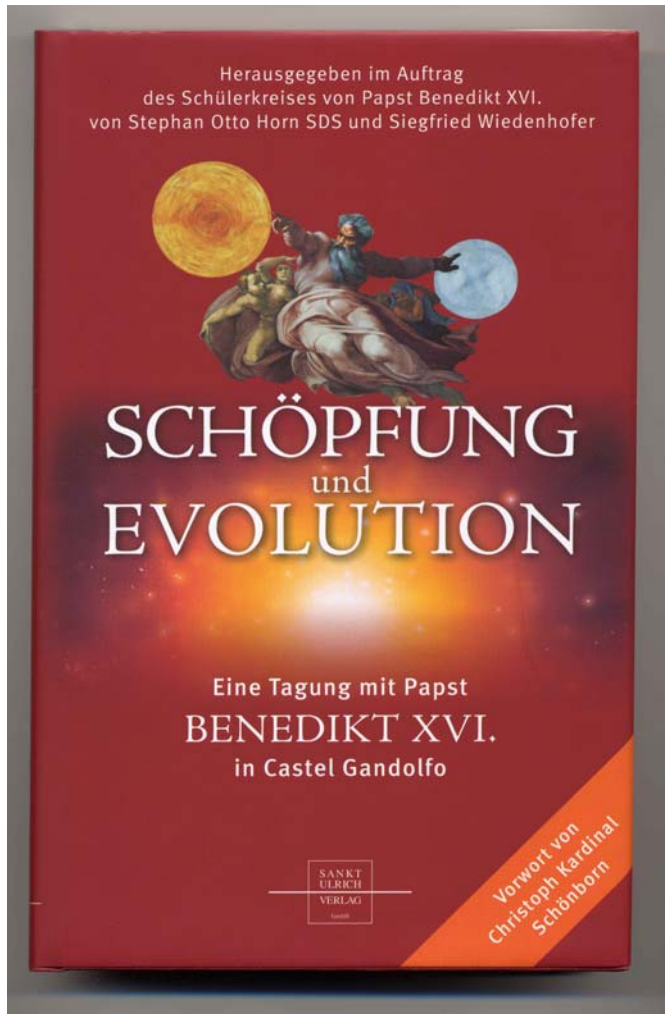
In: Stephan Otto Horn und Siegfried Wiedenhofer, Eds.

Schöpfung und Evolution. Eine Tagung mit Papst Benedikt XVI
in Castel Gandolfo. Sankt Ulrich Verlag, Augsburg 2007, pp.25-56.

English translation:

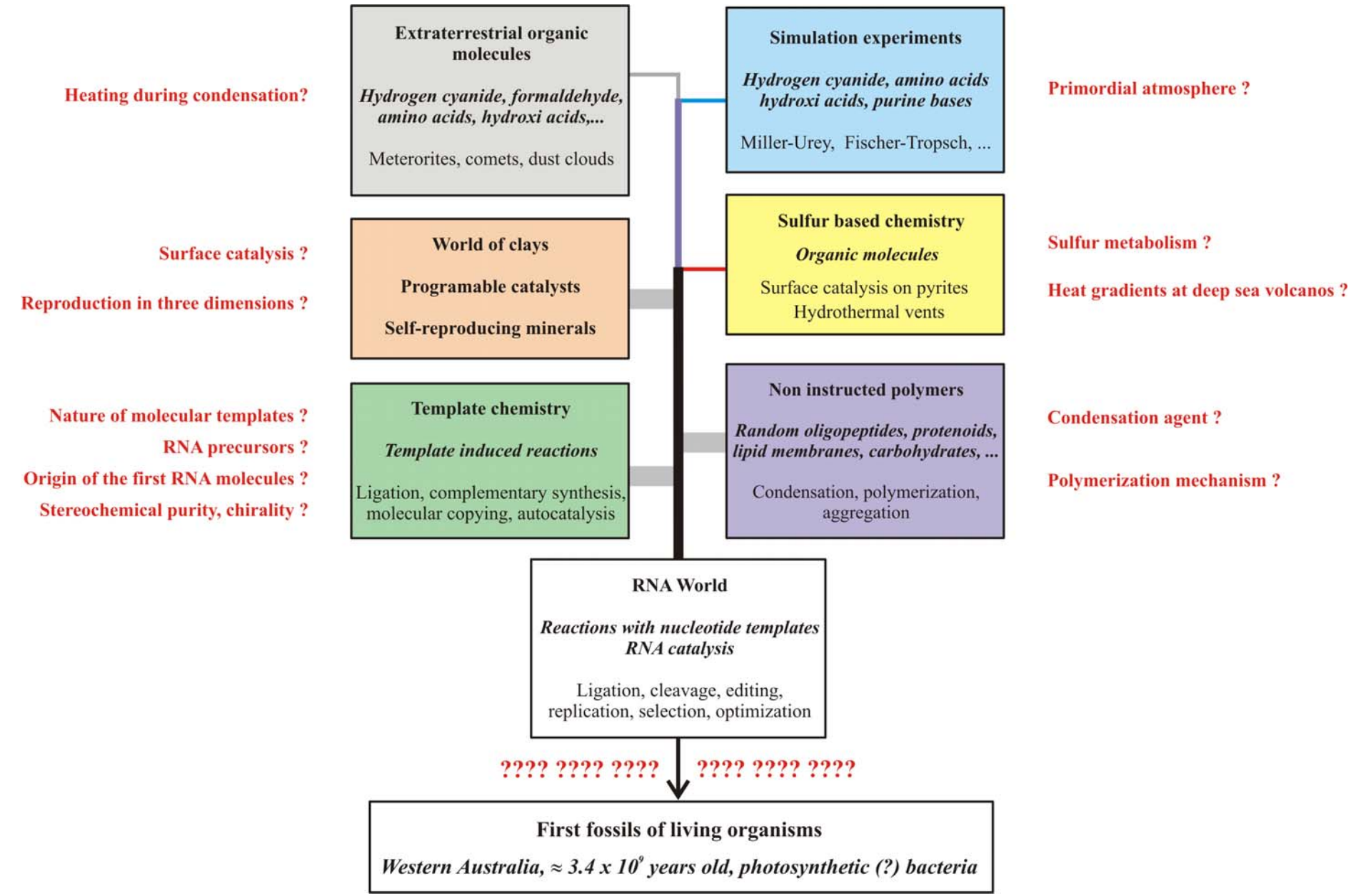
Creation and Evolution.

Ignatius Press, San Francisco, CA, 2008

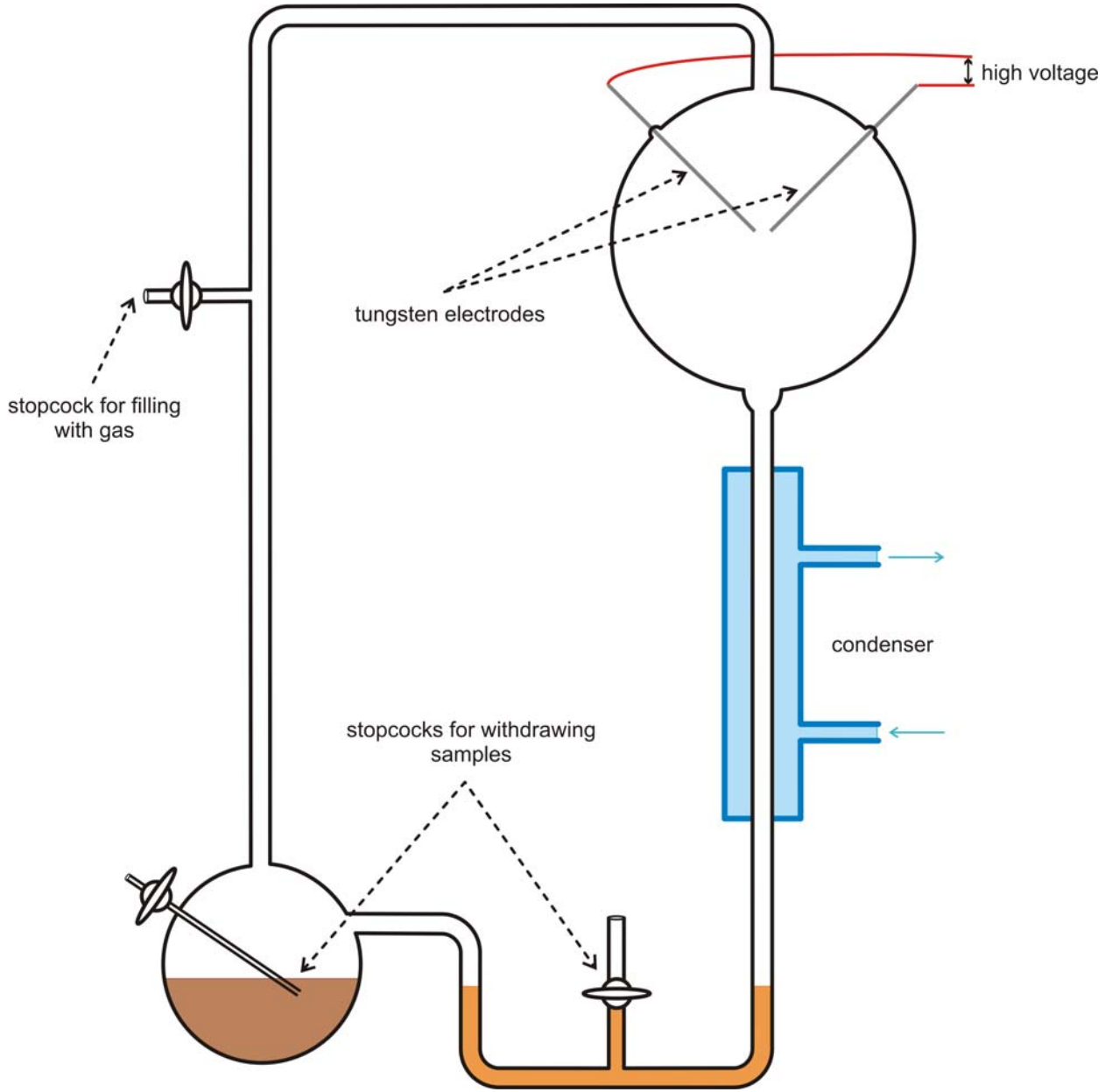


1. Origins of evolution
2. The „tree of life“
3. Probability and chance
4. Natural and artificial selection
5. Neutrality and random drift
6. Contingency and history

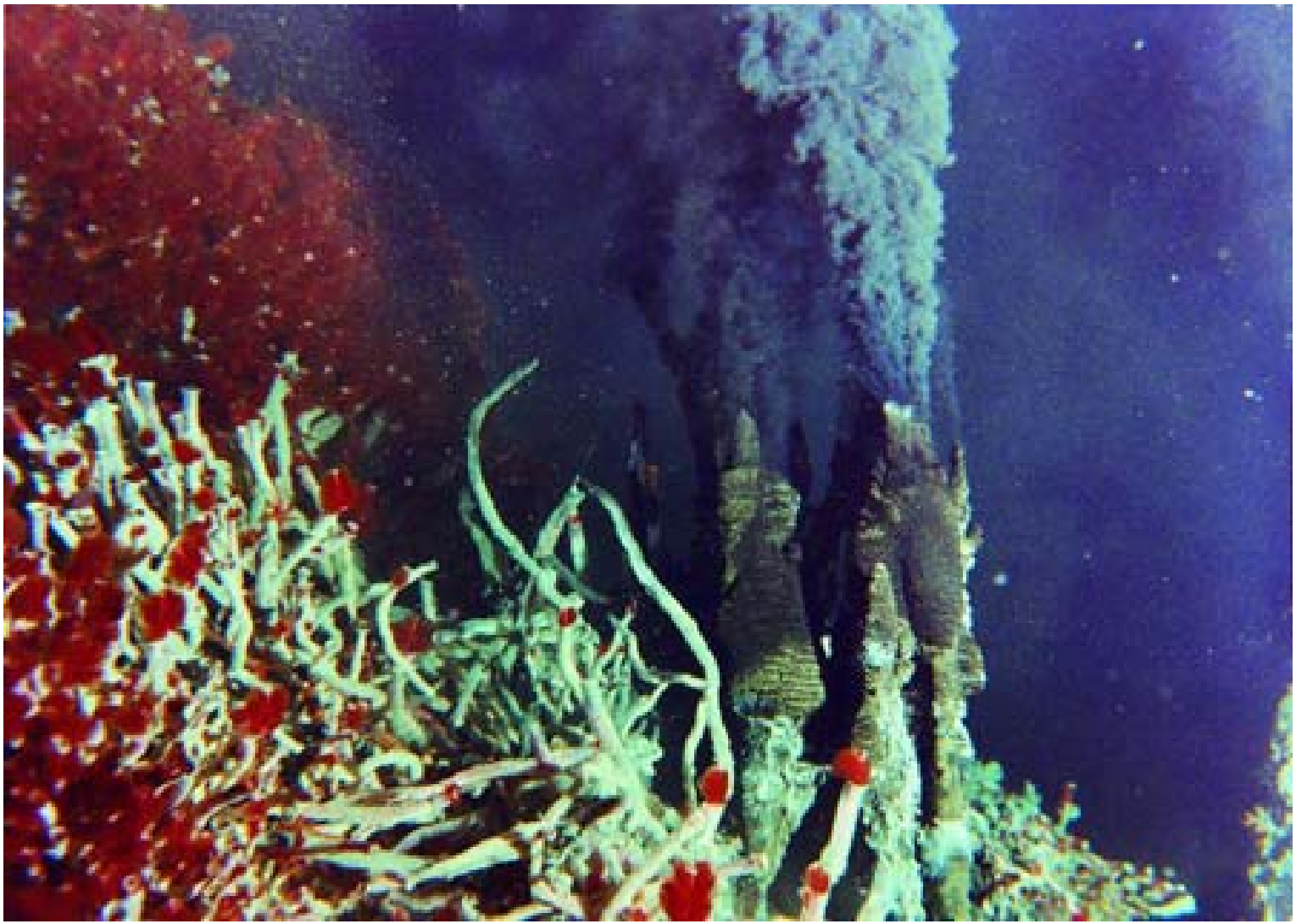
1. **Origins of evolution**
2. The „tree of life“
3. Probability and chance
4. Natural and artificial selection
5. Neutrality and random drift
6. Contingency and history


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 10^9 years old, photosynthetic (?) bacteria

‚Horsehead‘ nebula in orion contains a huge dark cloud

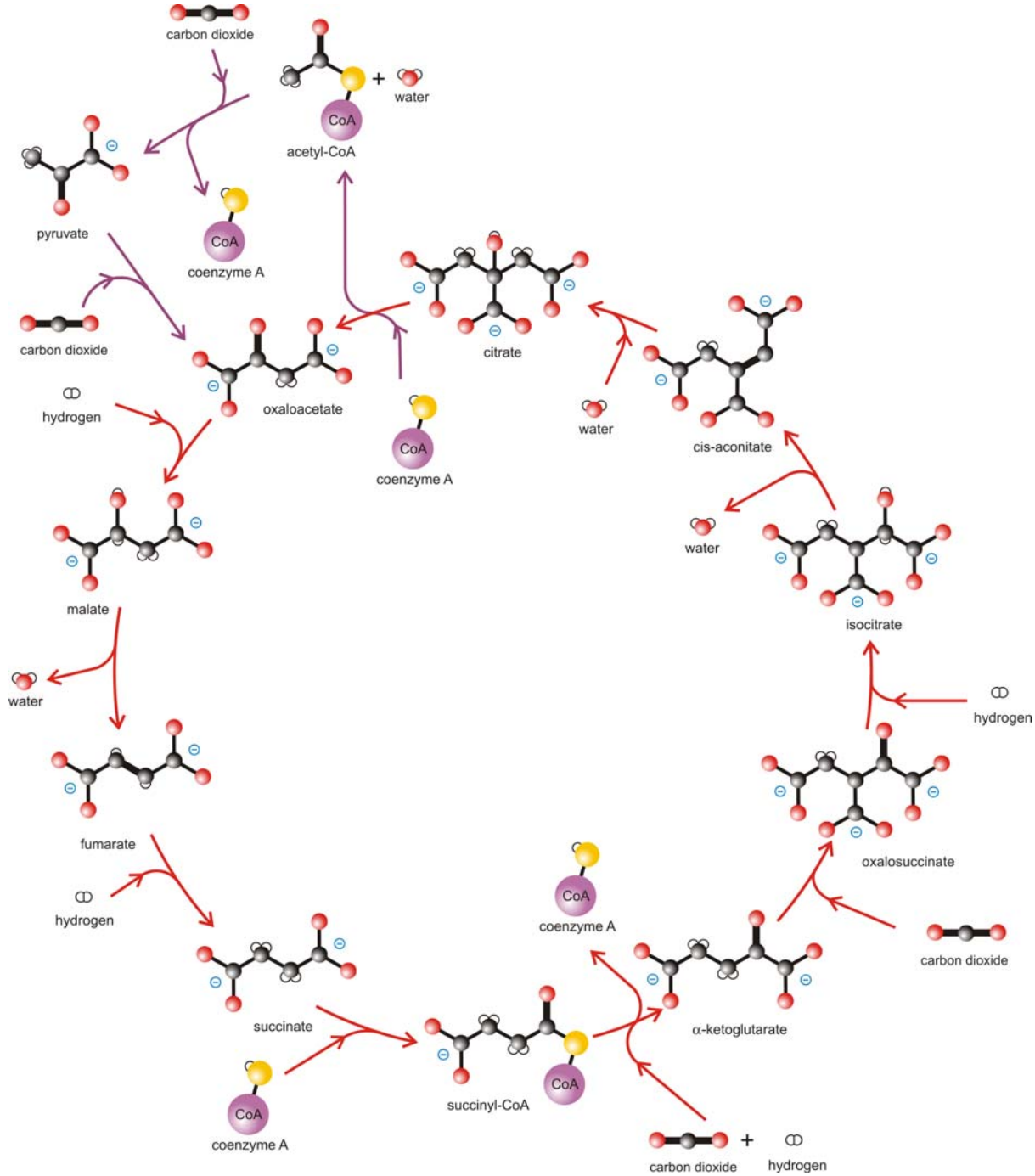


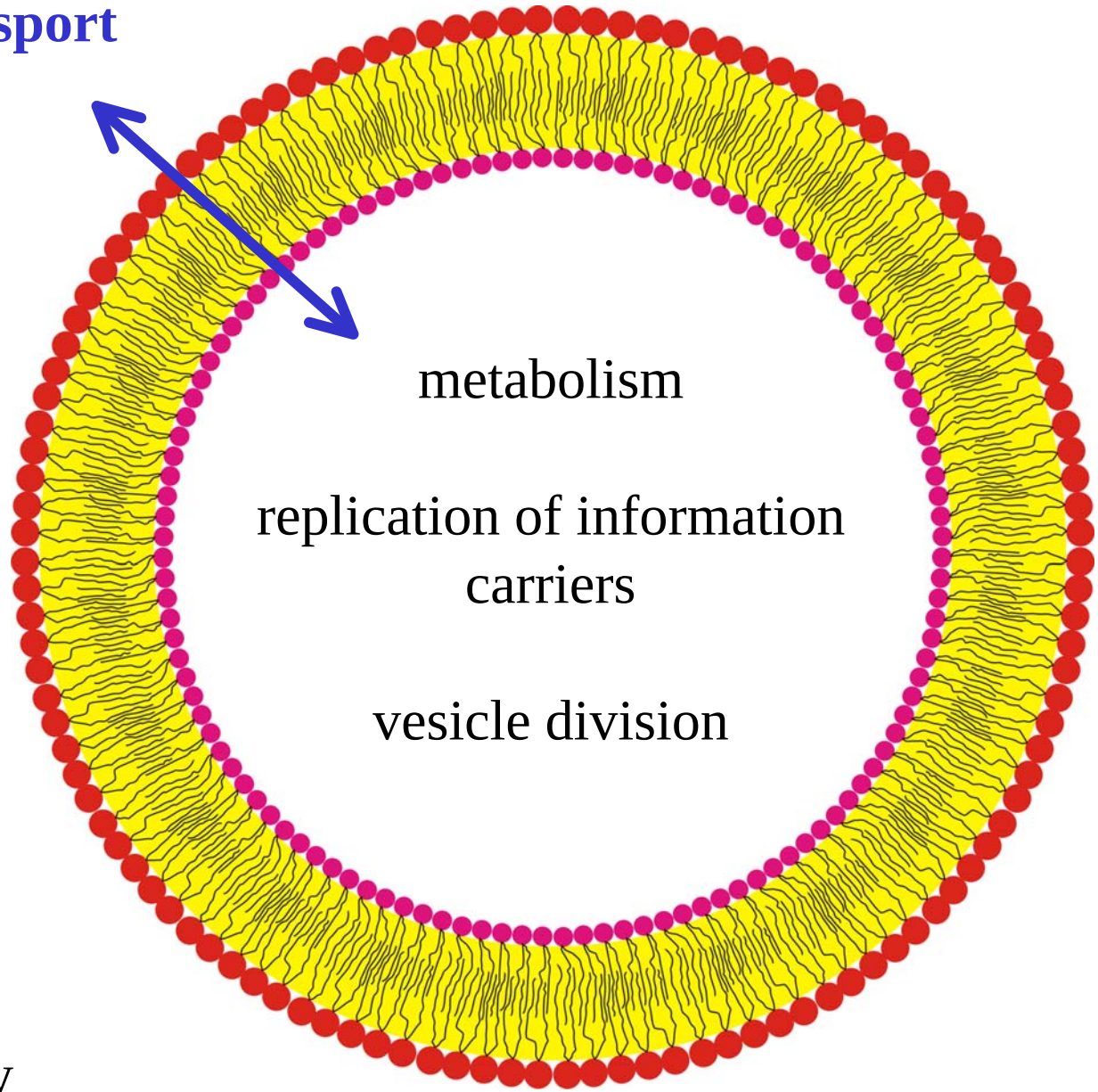
S.L.Miller. A production of amino acids under possible primitve earth conditions. Science 117, 528-529 (1953)



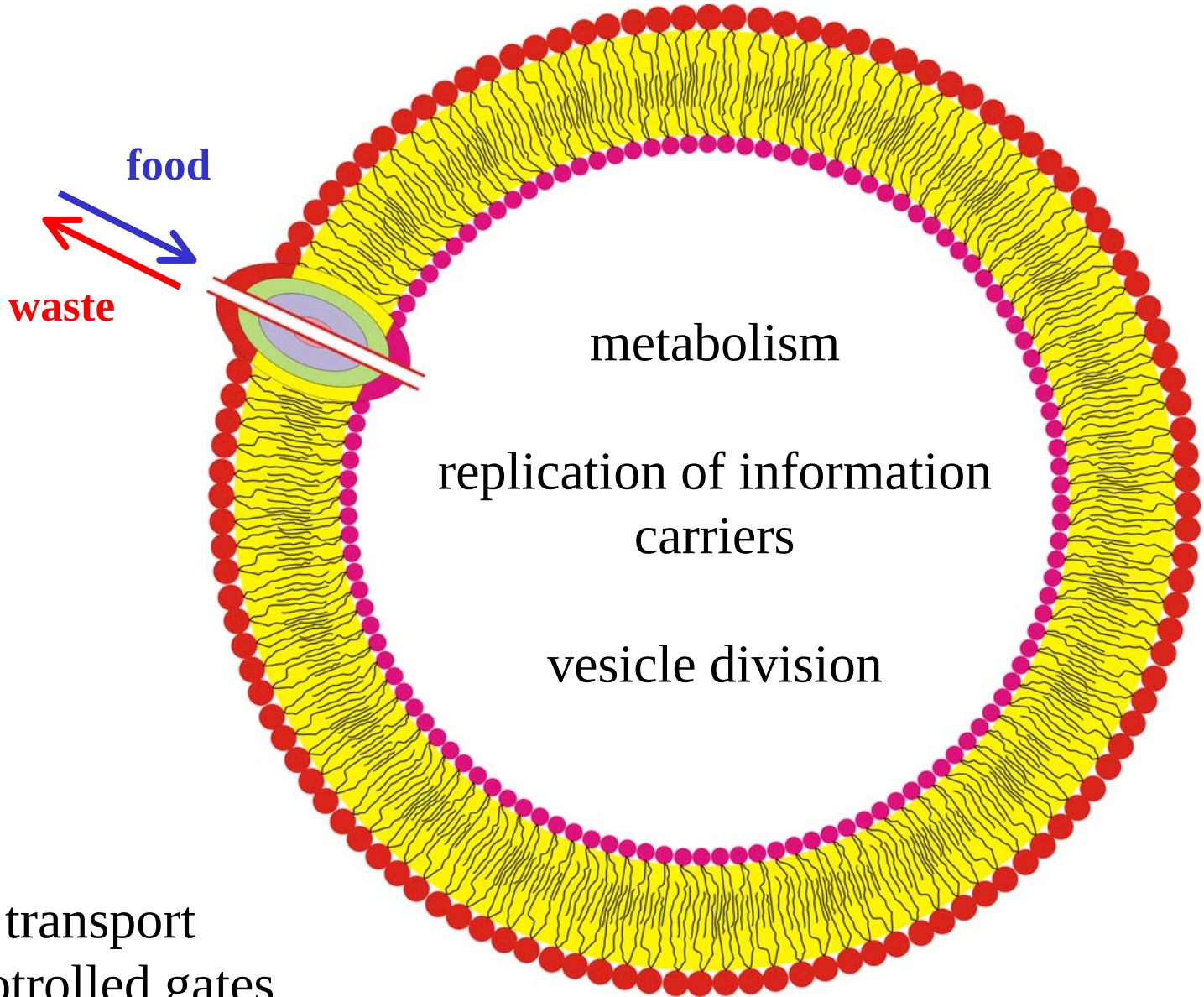
A hydrothermal vent

The reverse citrate cycle as a model for prebiotic synthesis of carbon compounds





compartments in prebiotic chemistry



organized transport
through cotrolled gates

1. Origins of evolution
2. **The „tree of life“**
3. Probability and chance
4. Natural and artificial selection
5. Neutrality and random drift
6. Contingency and history

ON

THE ORIGIN OF SPECIES

BY MEANS OF NATURAL SELECTION,

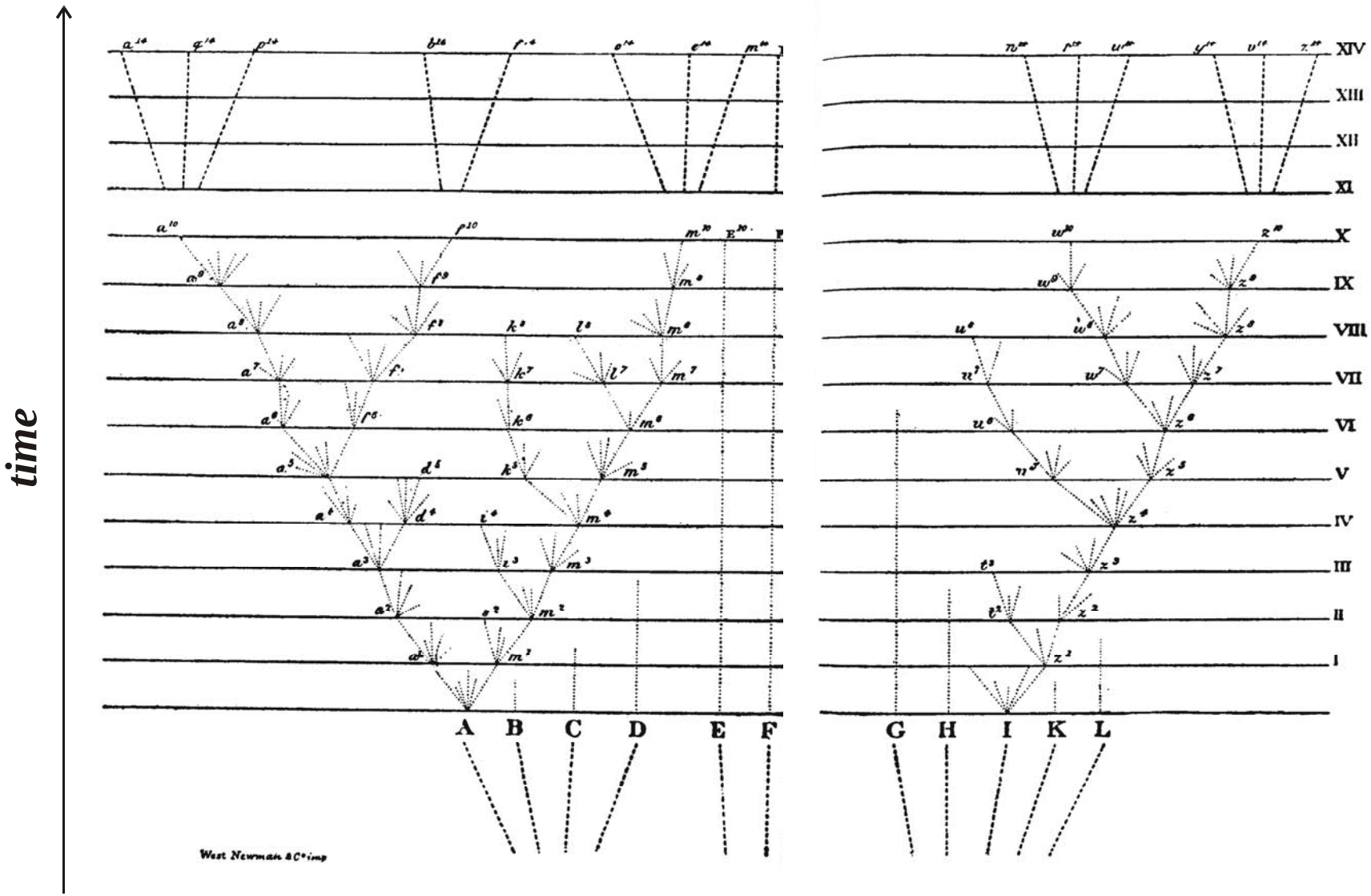
OR THE

PRESERVATION OF FAVOURED RACES IN THE STRUGGLE FOR LIFE.

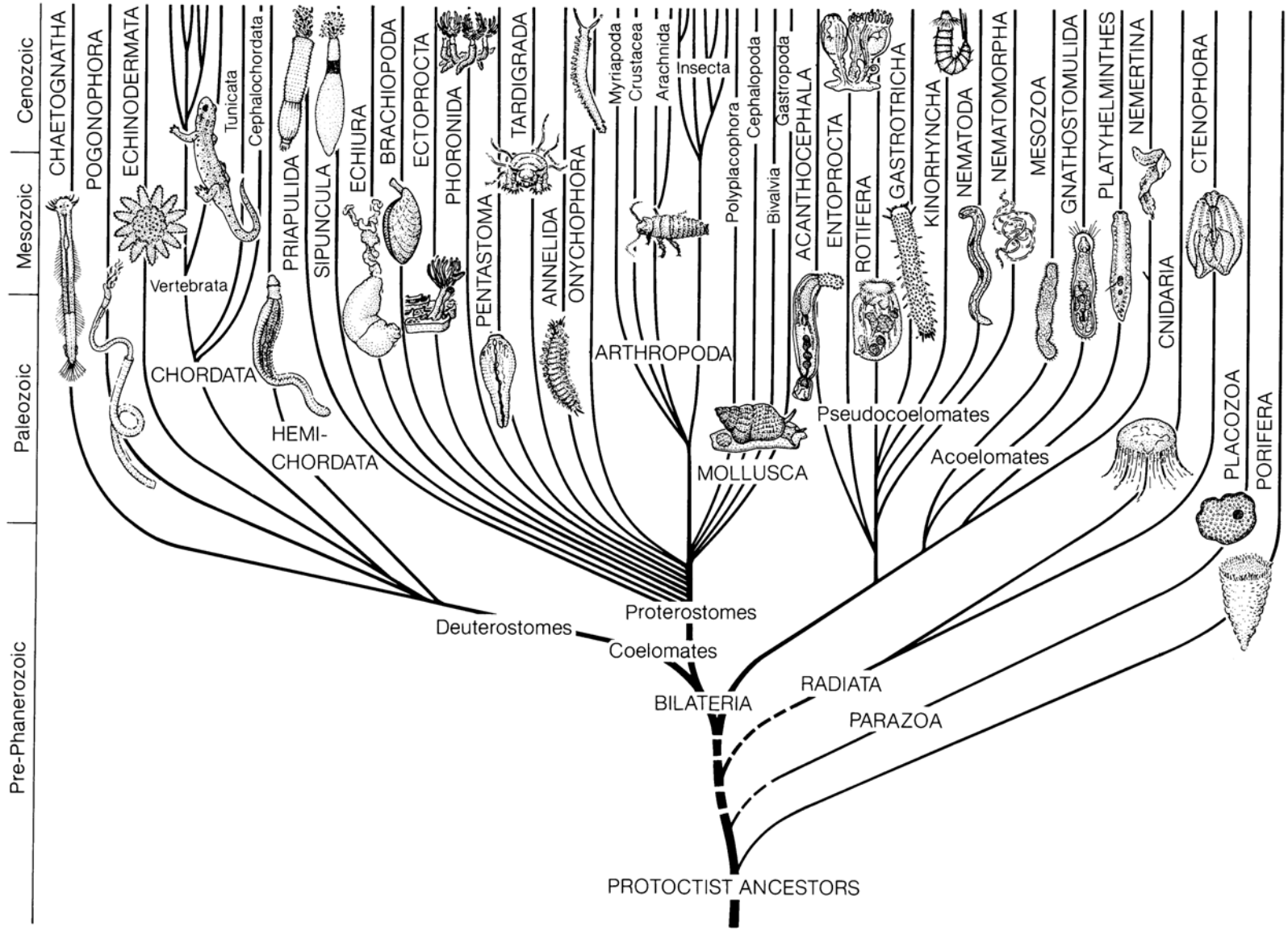
By CHARLES DARWIN, M.A.,

FELLOW OF THE ROYAL, GEOLOGICAL, LINNÆAN, ETC., SOCIETIES;
AUTHOR OF 'JOURNAL OF RESEARCHES DURING H. M. S. BEAGLE'S VOYAGE ROUND THE WORLD.'

LONDON:
JOHN MURRAY, ALBEMARLE STREET.
1859.



Charles Darwin, *The Origin of Species*, 6th edition.
Everyman‘s Library, Vol.811, Dent London, pp.121-122.



Modern phylogenetic tree: Lynn Margulis, Karlene V. Schwartz. *Five Kingdoms*. *An Illustrated Guide to the Phyla of Life on Earth*. W.H. Freeman, San Francisco, 1982.

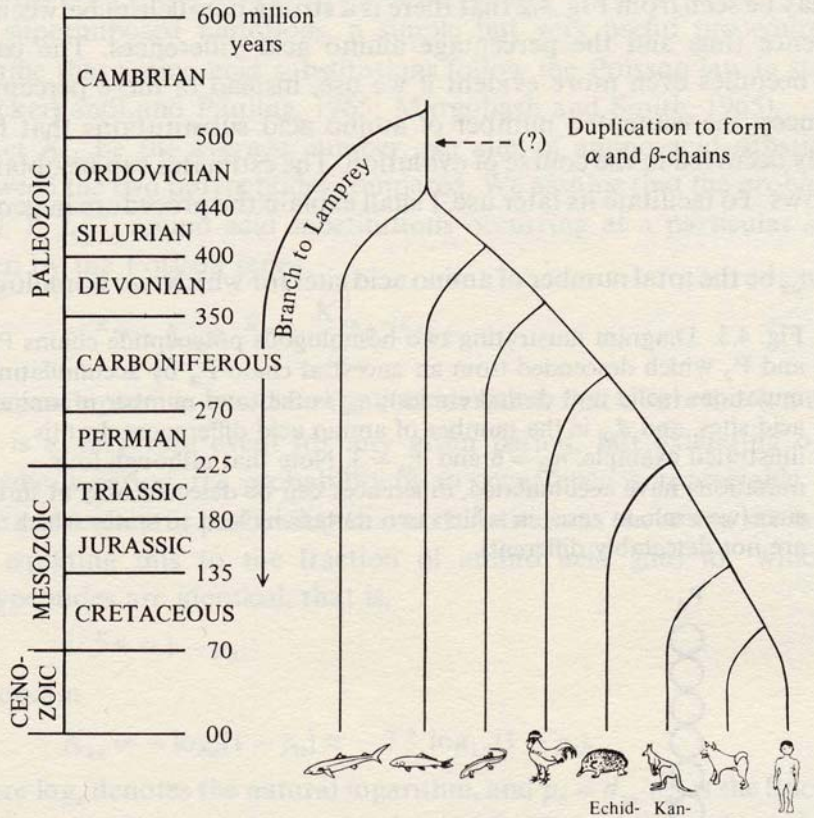
Motoo Kimura, 1924 - 1994

The molecular clock of evolution

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

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

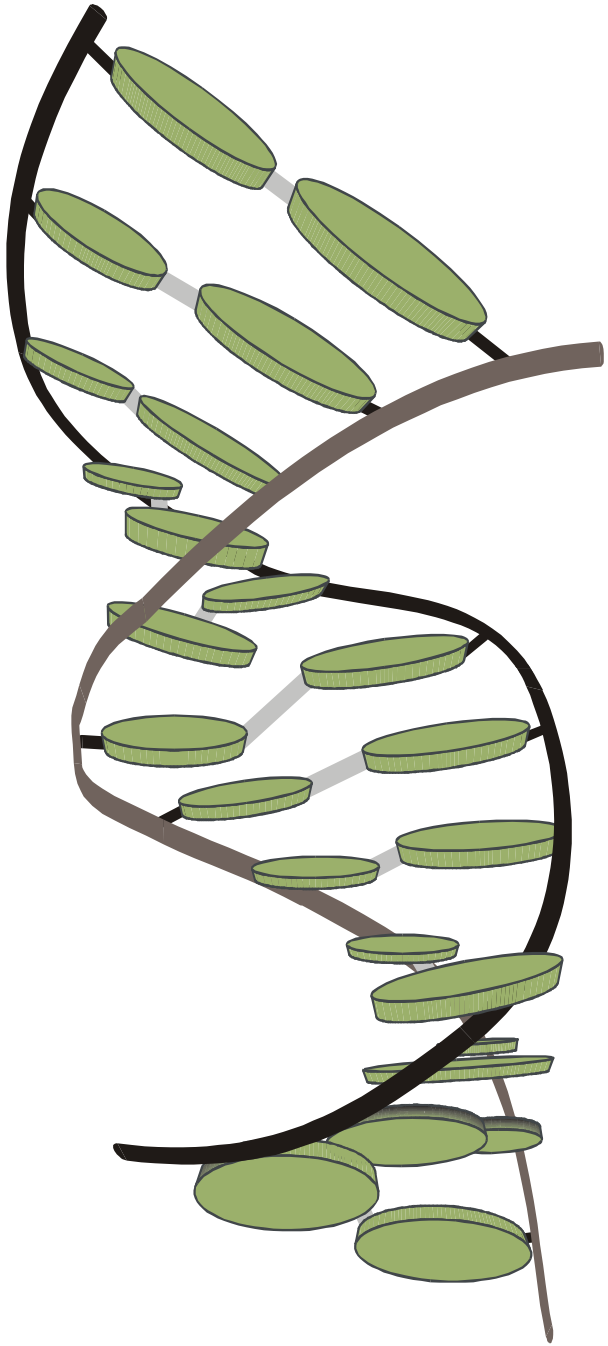
Fig. 4.2. Percentage amino acid differences when the α hemoglobin chains are compared among eight vertebrates together with their phylogenetic relationship and the times of divergence.

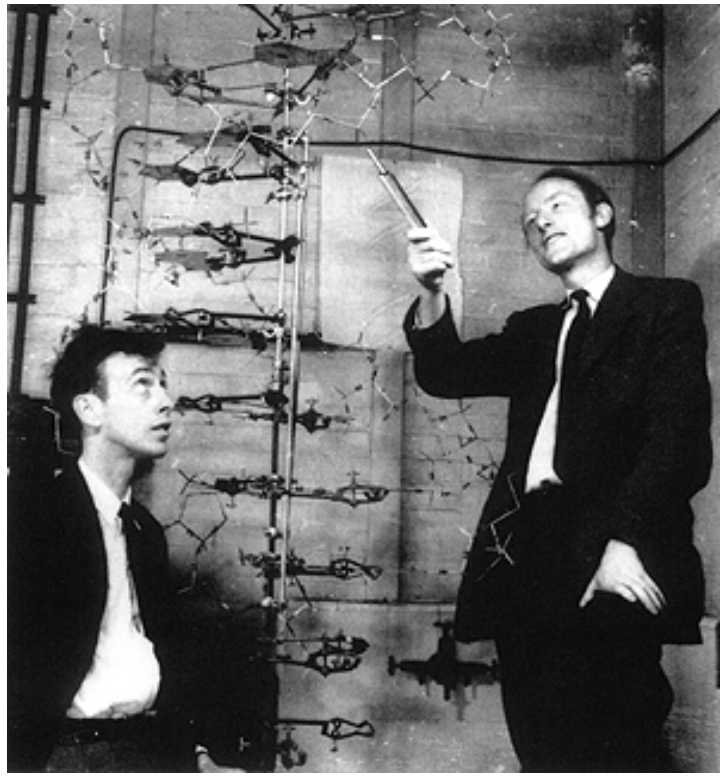


α-chain

Percent Amino Acid Differences

	Shark	Carp	Newt	Chicken	Echidna	Kangaroo	Dog	Human
Shark		59.4	61.4	59.7	60.4	55.4	56.8	53.2
Carp			53.2	51.4	53.6	50.7	47.9	48.6
Newt				44.7	50.4	47.5	46.1	44.0
Chicken					34.0	29.1	31.2	24.8
Echidna						34.8	29.8	26.2
Kangaroo							23.4	19.1
Dog								16.3
Human								



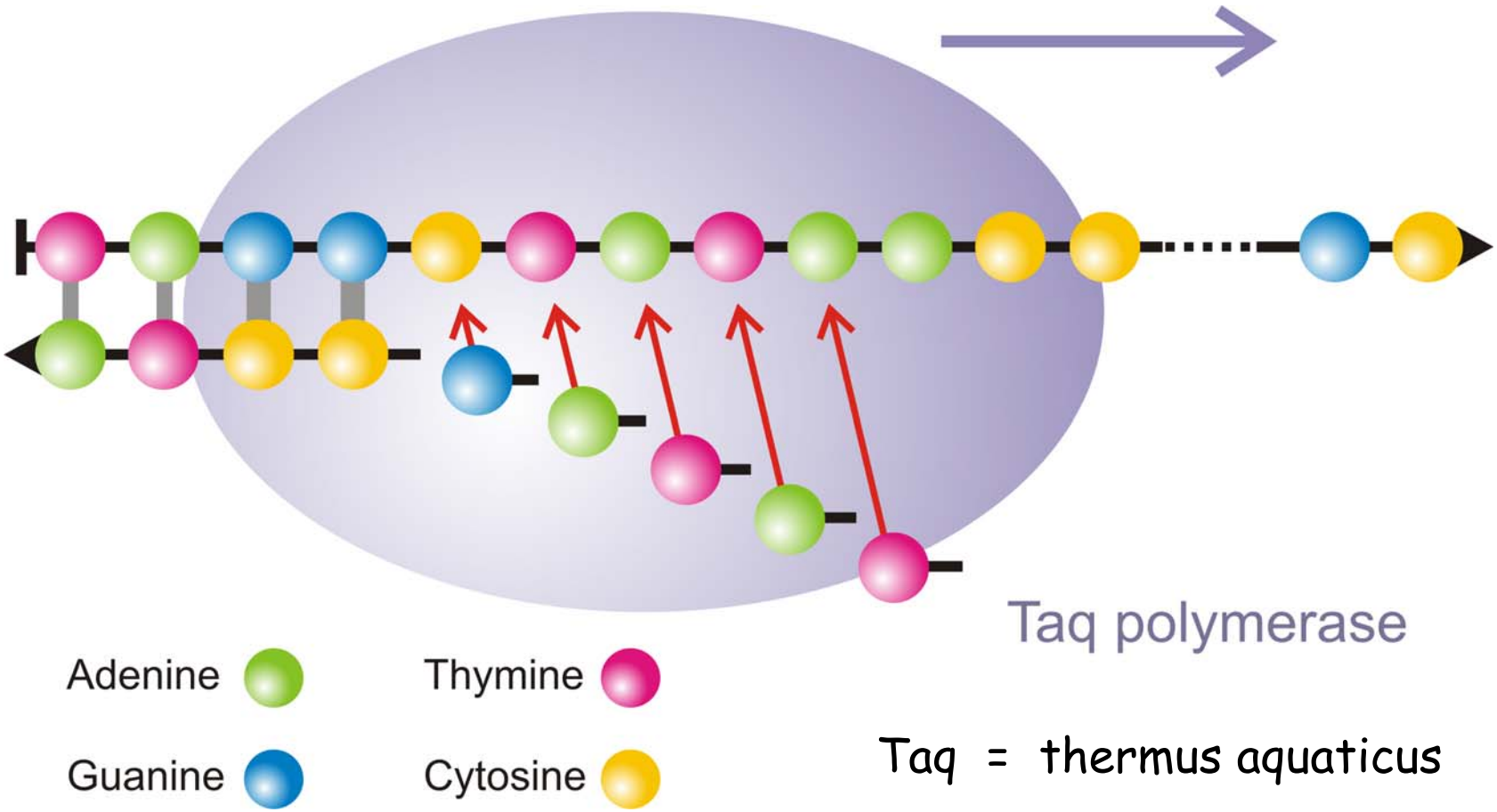


James D. Watson, 1928-, and Francis H.C. Crick, 1916-2004

Nobel prize 1962

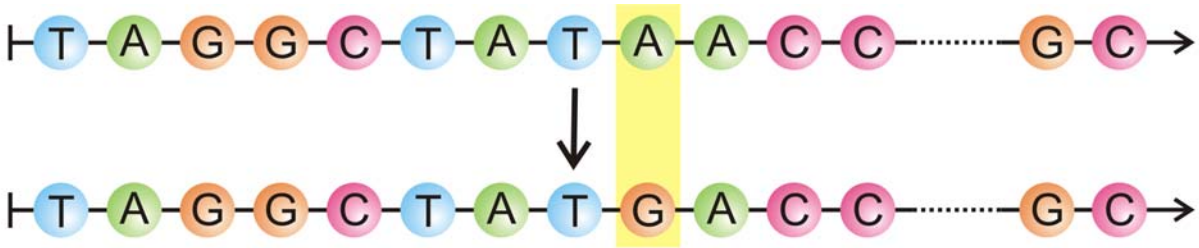
1953 – 2003 fifty years double helix

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

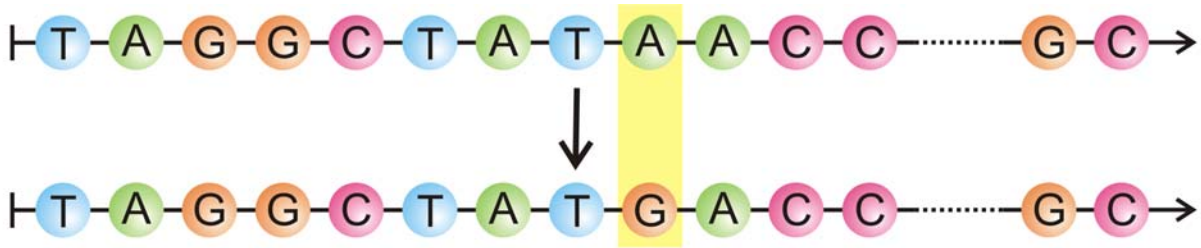


Taq = thermus aquaticus

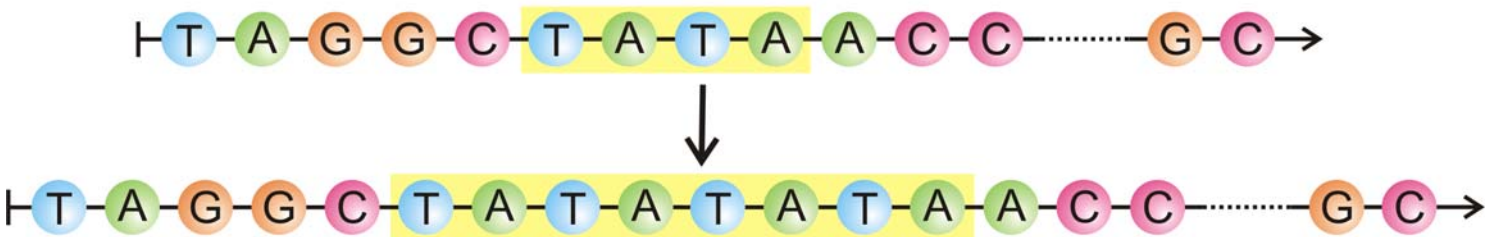
The logics of DNA replication



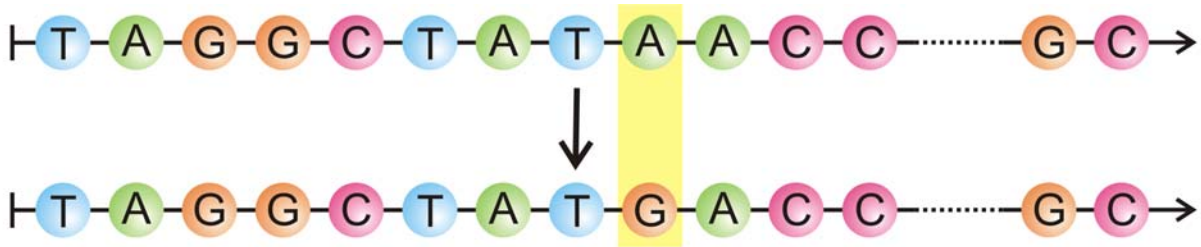
Punktmutation



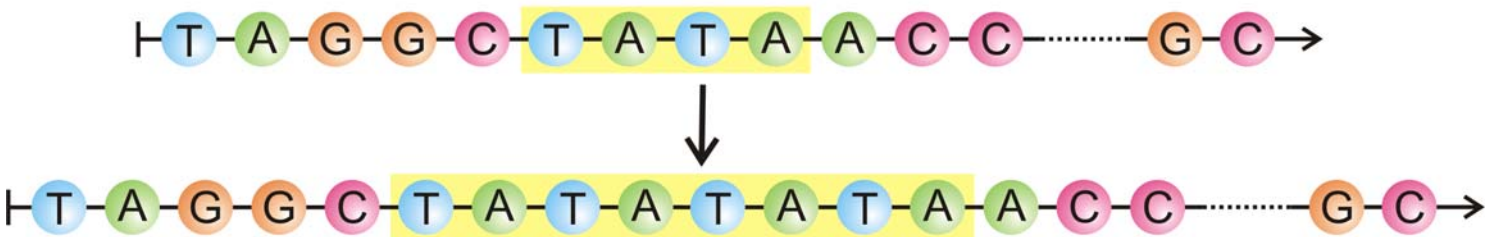
Punktmutation



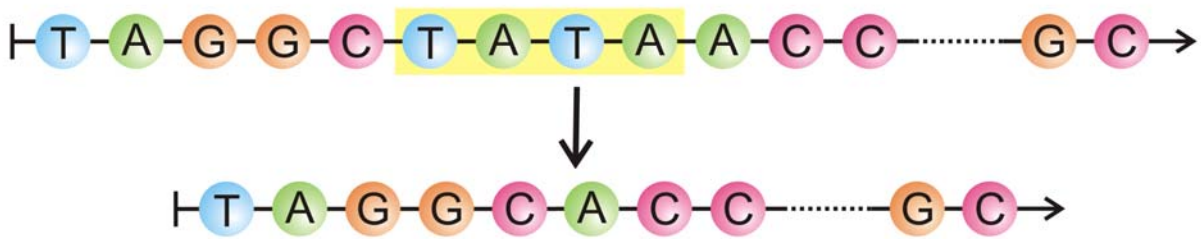
Insertion



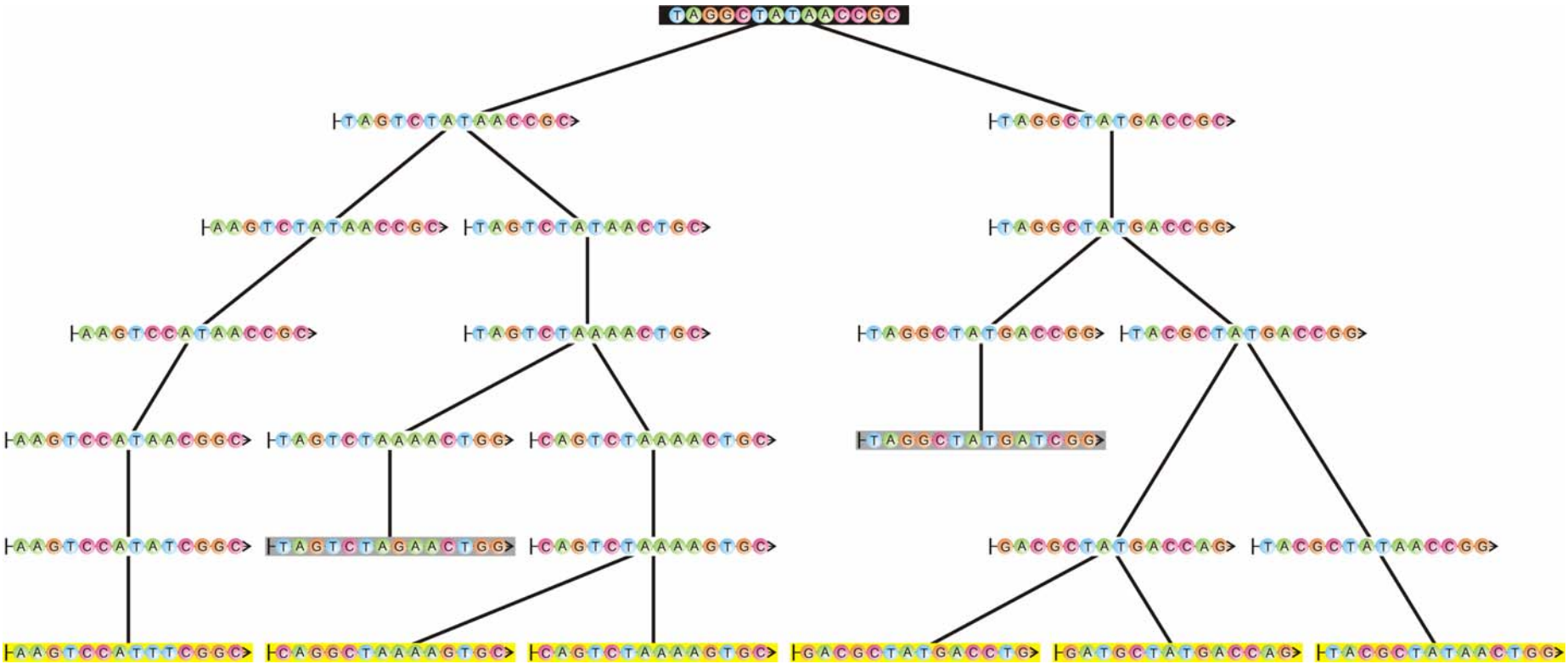
Punktmutation



Insertion



Deletion



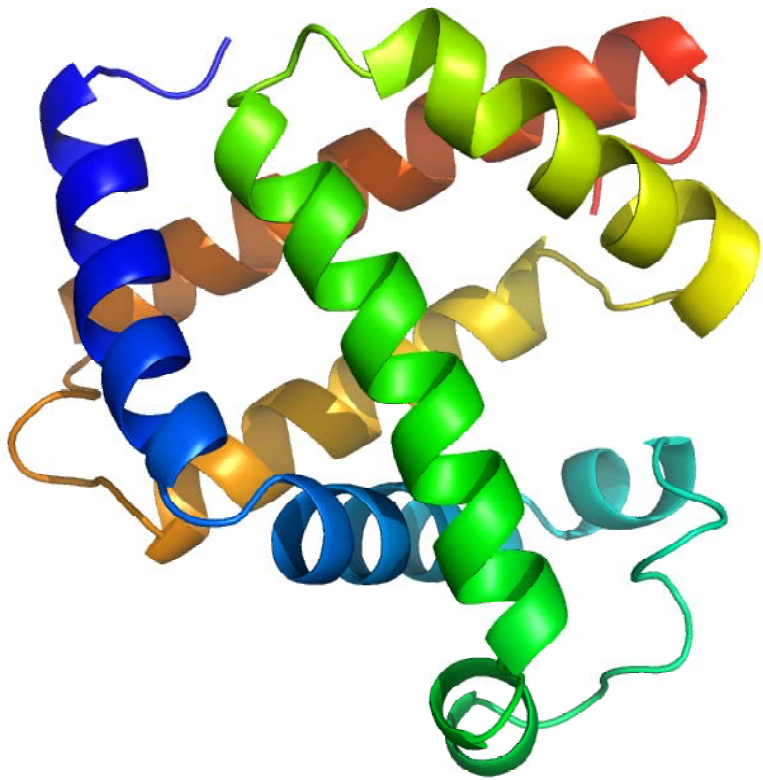
Reconstruction of phylogenies through comparison of molecular sequence data

1. Origins of evolution
2. The „tree of life“
3. **Probability and chance**
4. Natural and artificial selection
5. Neutrality and random drift
6. Contingency and history

Polymer chain of 153 amino acid residues with the sequence:

GLSDGEWQLVLNVWGKVEADIPGHGQEVLIRLFKGHPETLEKFDKFKHLK
SEDEMKASEDLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIP
VKYLEFISECIIQVLQSKHPGDFGADAQGAMNKALELFRKDMASNYKELG
FQG

The myglobin molecule



Eugene Wigner's or Fred Hoyle's argument applied to myoglobin:

All sequences have equal probability and all except the correct one have no survival value or are lethal

GLSDGEWQLVLNVWG.....FQG

Alphabet size: 20

Chain length: 153 amino acids

Number of possible sequences: $20^{153} = \mathbf{0.11 \times 10^{200}}$

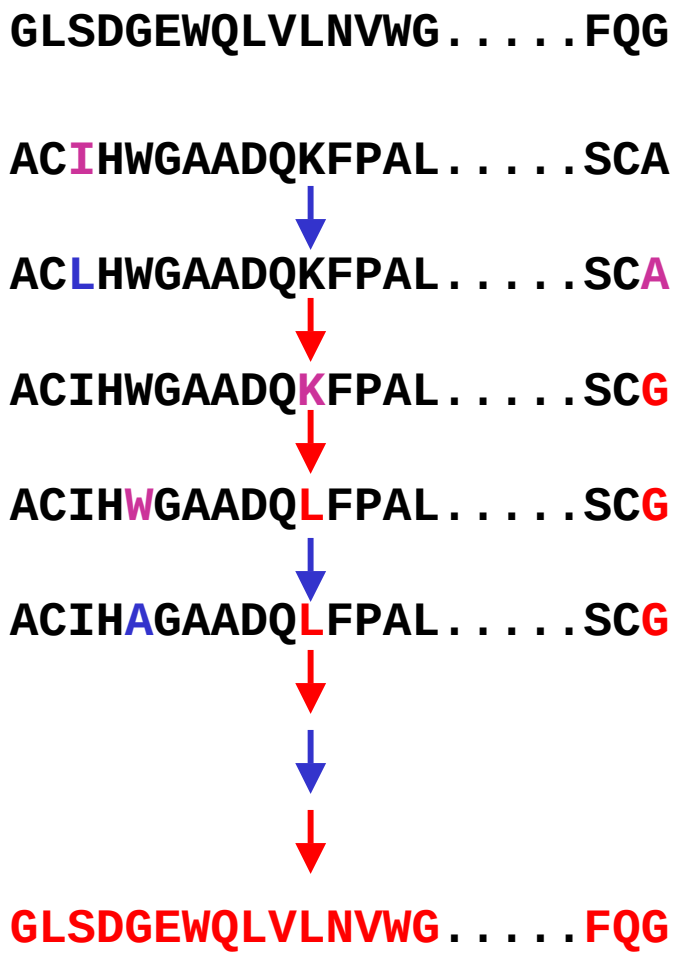
Probability to find the myoglobin sequence:

$20^{-153} = \mathbf{9 \times 10^{-200}} = 0.000\ldots\ldots009$

200

Eugene Wigner's and Fred Hoyle's arguments revisited:

Every single point mutation towards the target sequence leads to an improvement and is therefore selected



Alphabet size: 20

Chain length: 153 amino acids

Length of longest path to myoglobin sequence: $19 \times 153 = 2907$

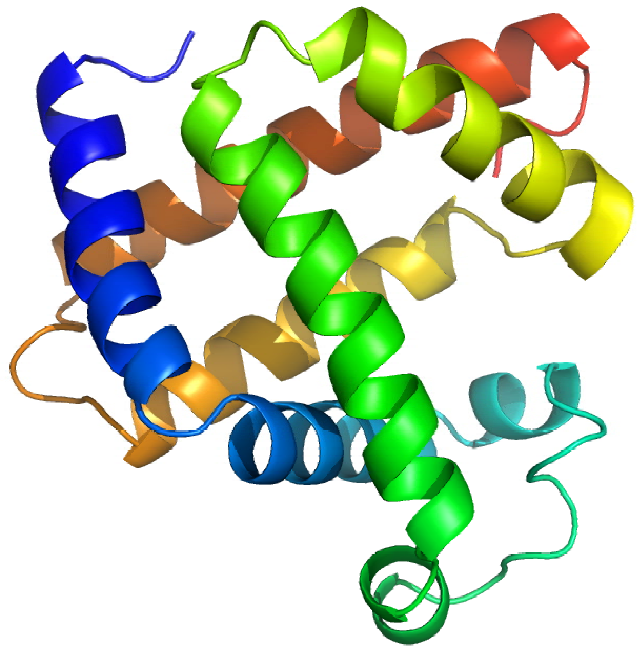
Probability to find the myoglobin sequence: **0.00034**

The folding problem of the myoglobin molecule:

A chain of 153 amino acid residues, each of which can adopt about 15 different geometries, can exist in

$15^{153} = 0.9 \times 10^{180}$ conformations.

One specific conformation – the most stable or minimum free energy conformation – has to be found in the folding process.



The Levinthal paradox of protein folding

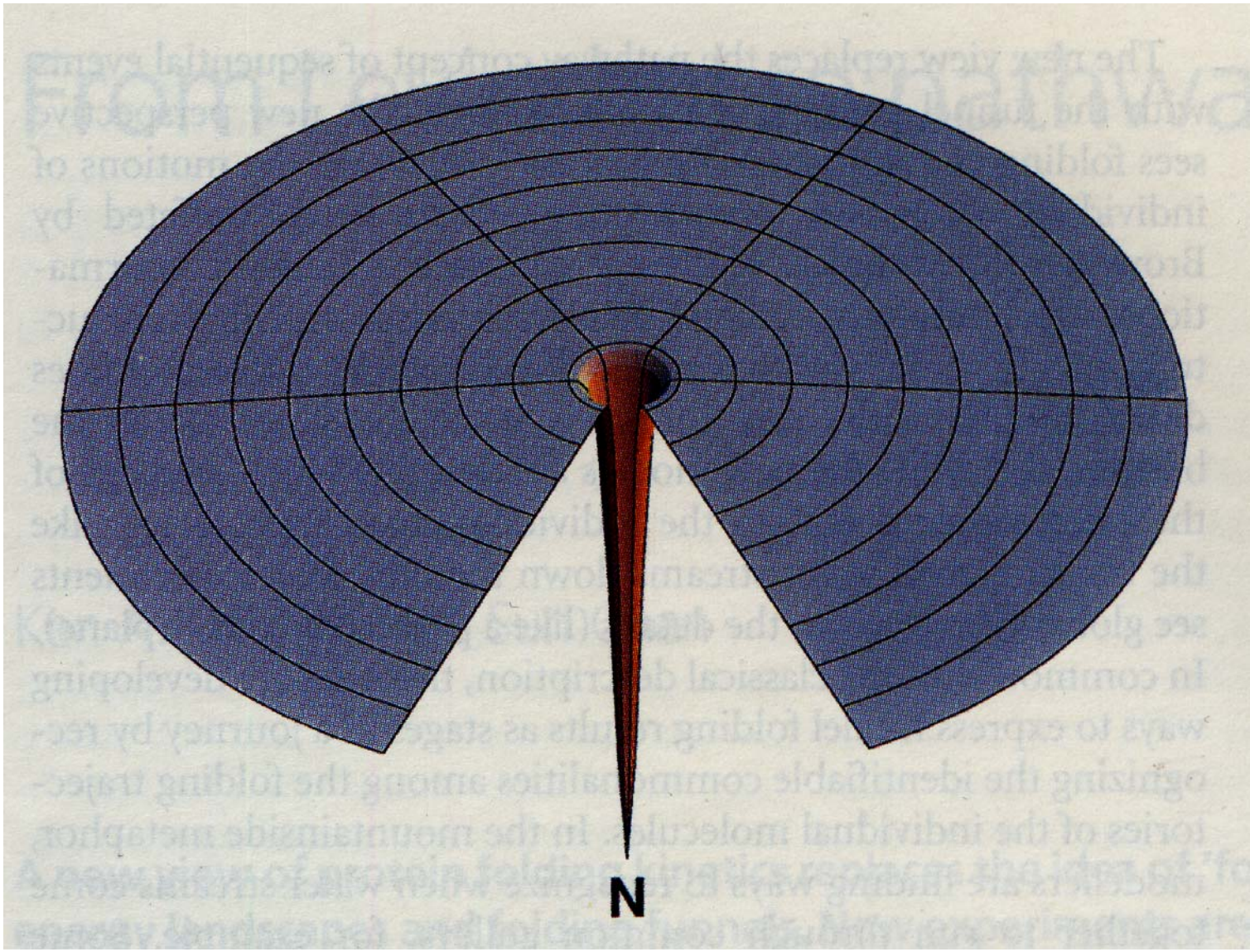
Three basic questions of the protein folding problem:

What is the folding code ?

What is the folding mechanism ?

Can we predict the native structure of a protein from its amino acid sequence?

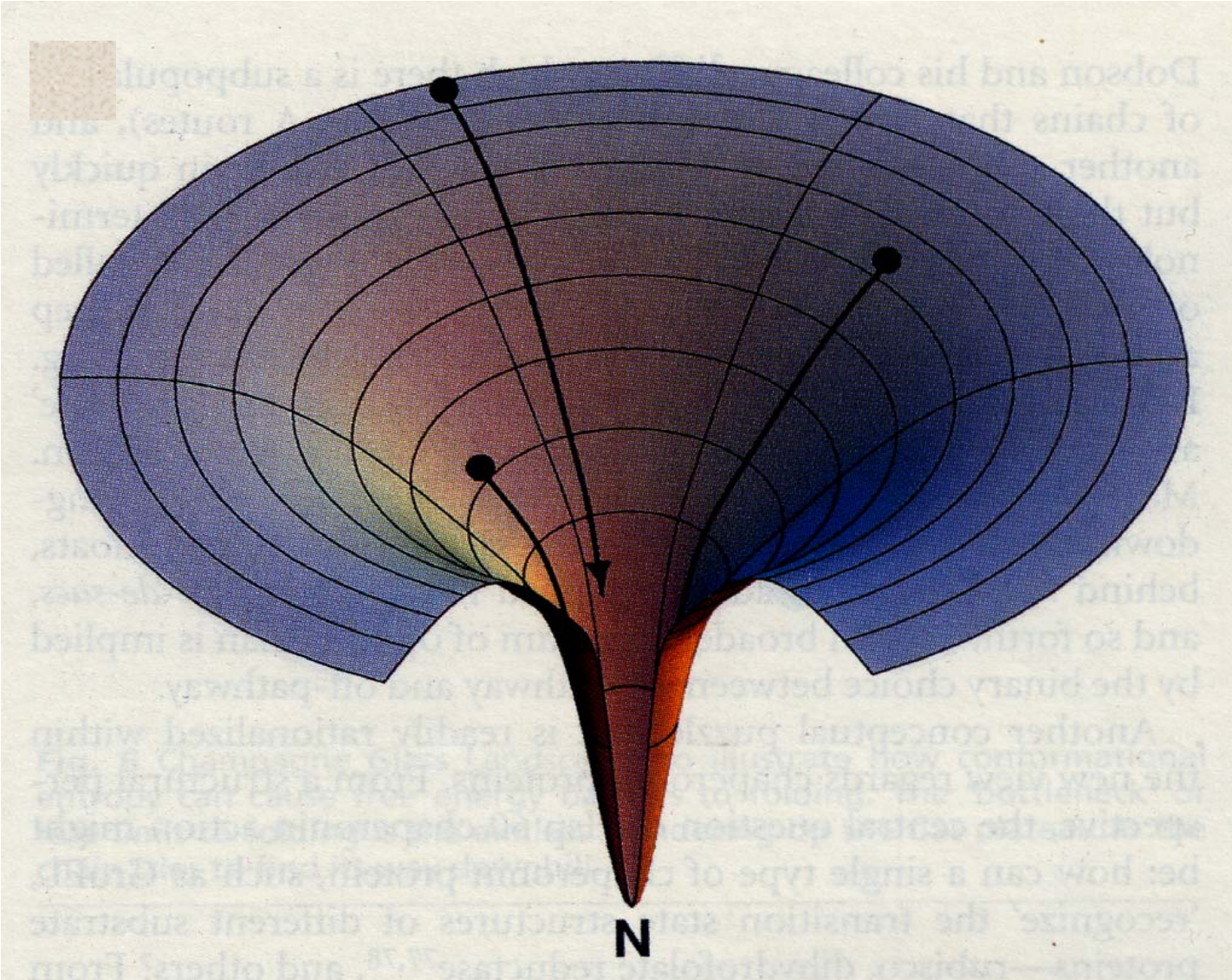
K.A. Dill, S.B. Ozkan, M.S. Shell, T.R. Weikl. 2008. The protein folding problem. Annu.Rev.Biophys. **37**:289-316.



The gulf course landscape

Solution to Levinthal's paradox

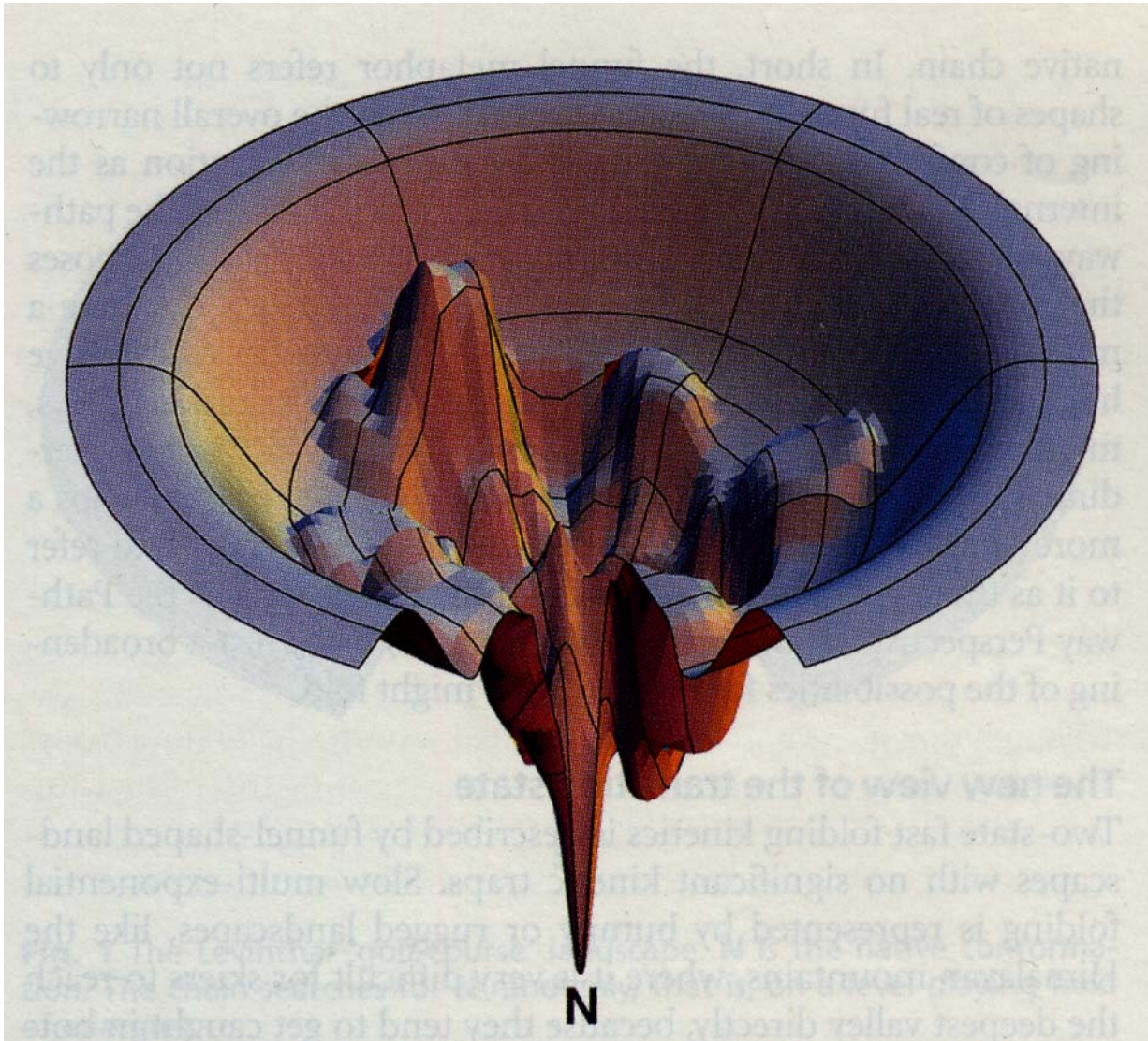
Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19



The funnel landscape

Solution to Levinthal’s paradox

Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19

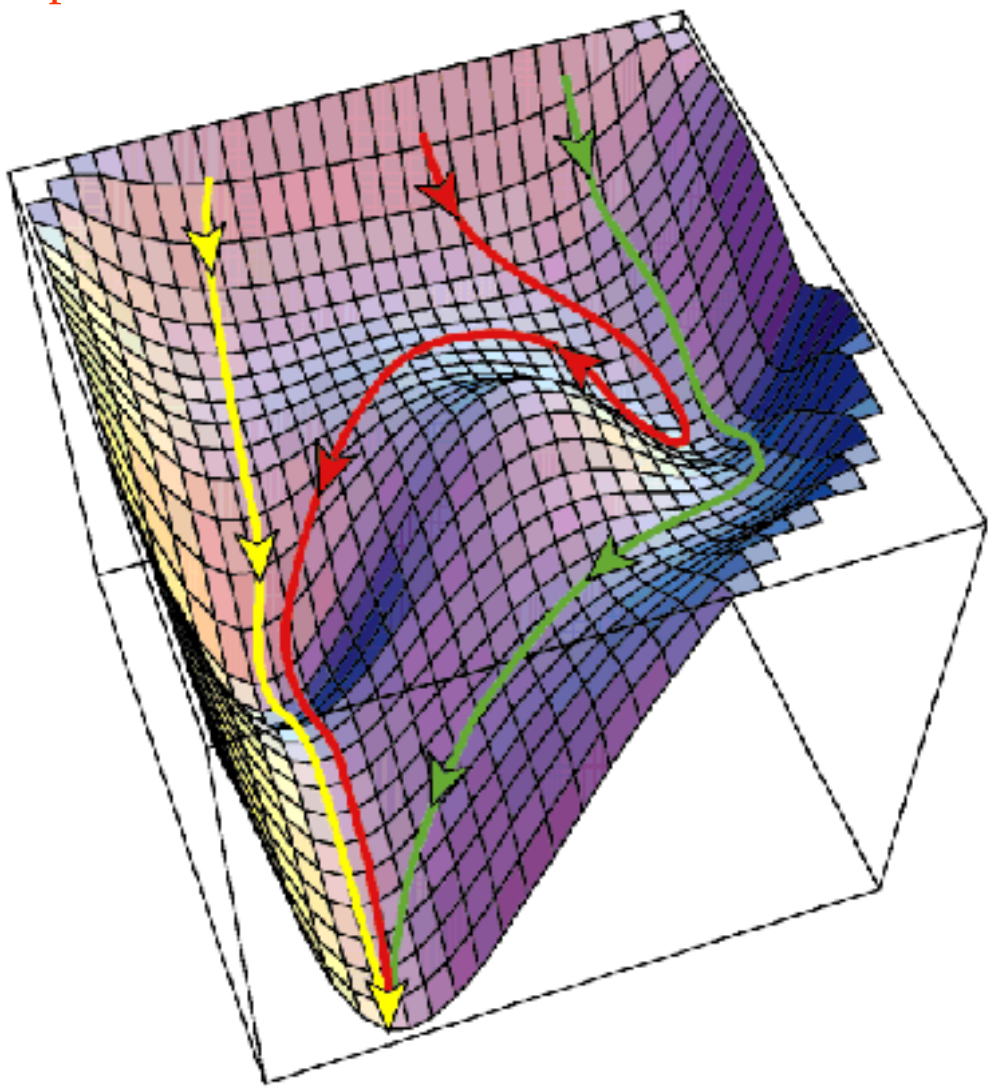


The structured funnel landscape

Solution to Levinthal’s paradox

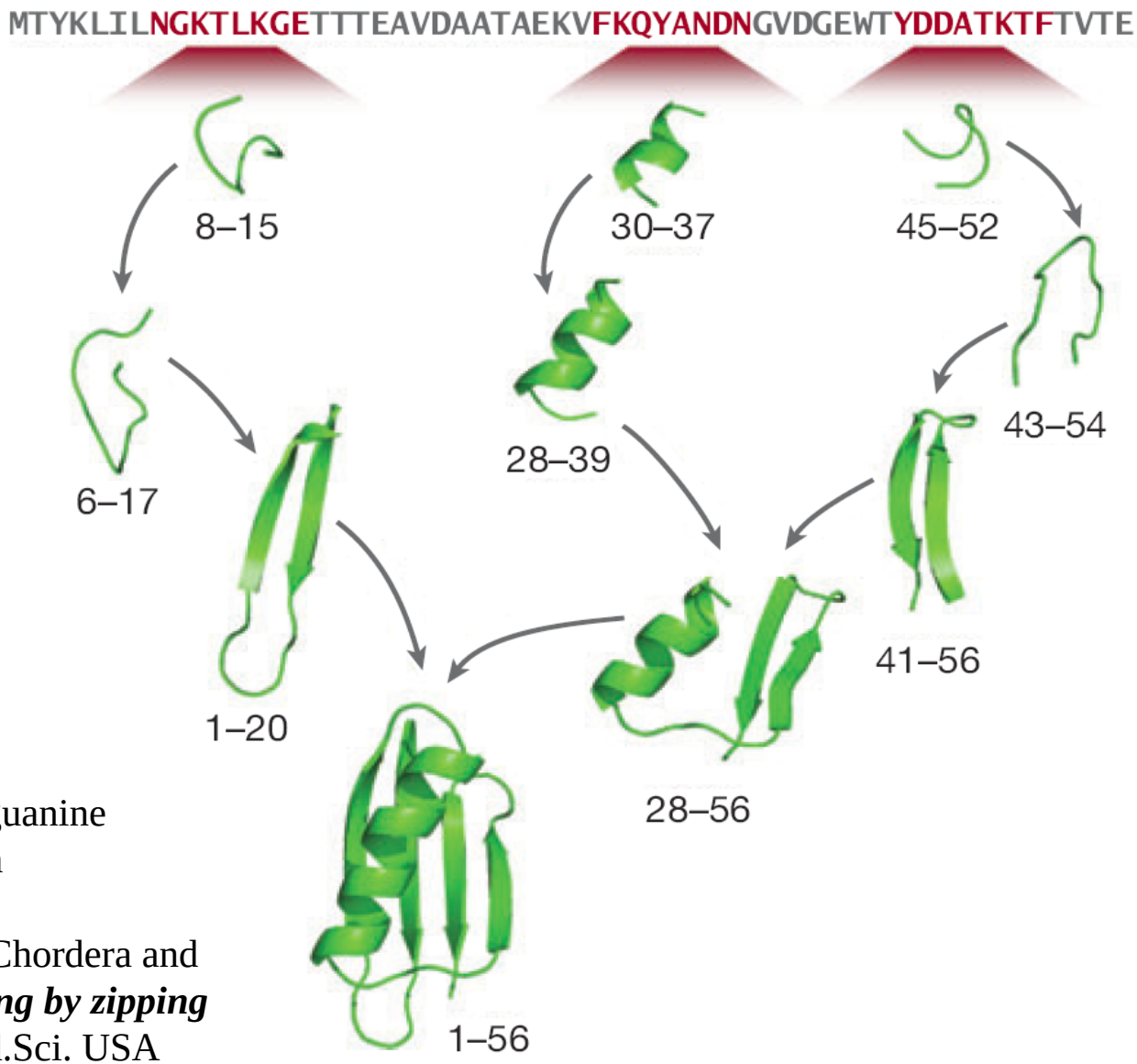
Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19

An “all-roads-lead-to-Rome” landscape



The reconstructed folding landscape
of a real biomolecule: “lysozyme”

Picture: C.M. Dobson, A. Šali, and M. Karplus, Angew.Chem.Internat.Ed. 37: 868-893, 1988



Computed folding routes for guanine nucleotide binding (G) protein

S.B. Ozkan, G.H.A. Wu, J.D.Chordera and K.A. Dill. 2007. ***Protein folding by zipping and assembly.*** Proc.Natl.Acad.Sci. USA **104**:11987-11992.

1. Origins of evolution
2. The „tree of life“
3. Probability and chance
4. **Natural and artificial selection**
5. Neutrality and random drift
6. Contingency and history

Charles Darwin, 1809-1882

Three necessary conditions for Darwinian evolution are:

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

All three conditions are fulfilled not only by cellular organisms but also by **nucleic acid molecules** – **DNA** or **RNA** – **in** suitable **cell-free experimental assays**:

Darwinian evolution in the test tube

Evolution in the test tube:

G.F. Joyce, *Angew.Chem.Int.Ed.* **46** (2007), 6420-6436

DOI: 10.1002/anie.200701369

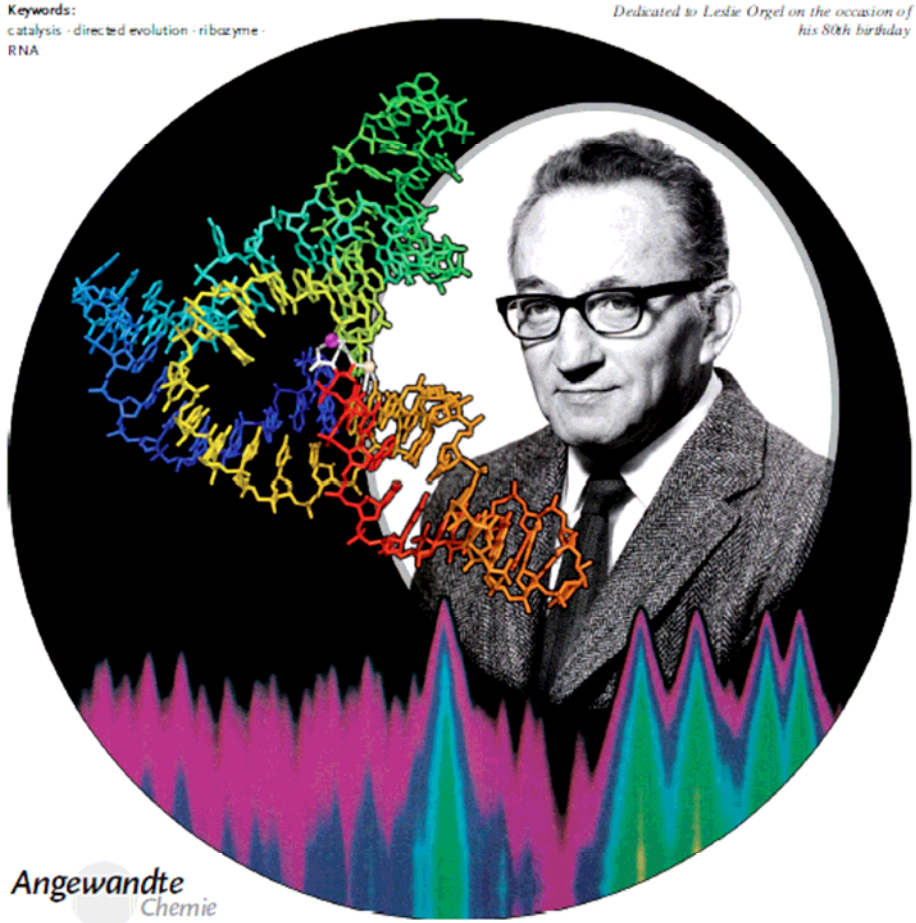
Molecular Evolution

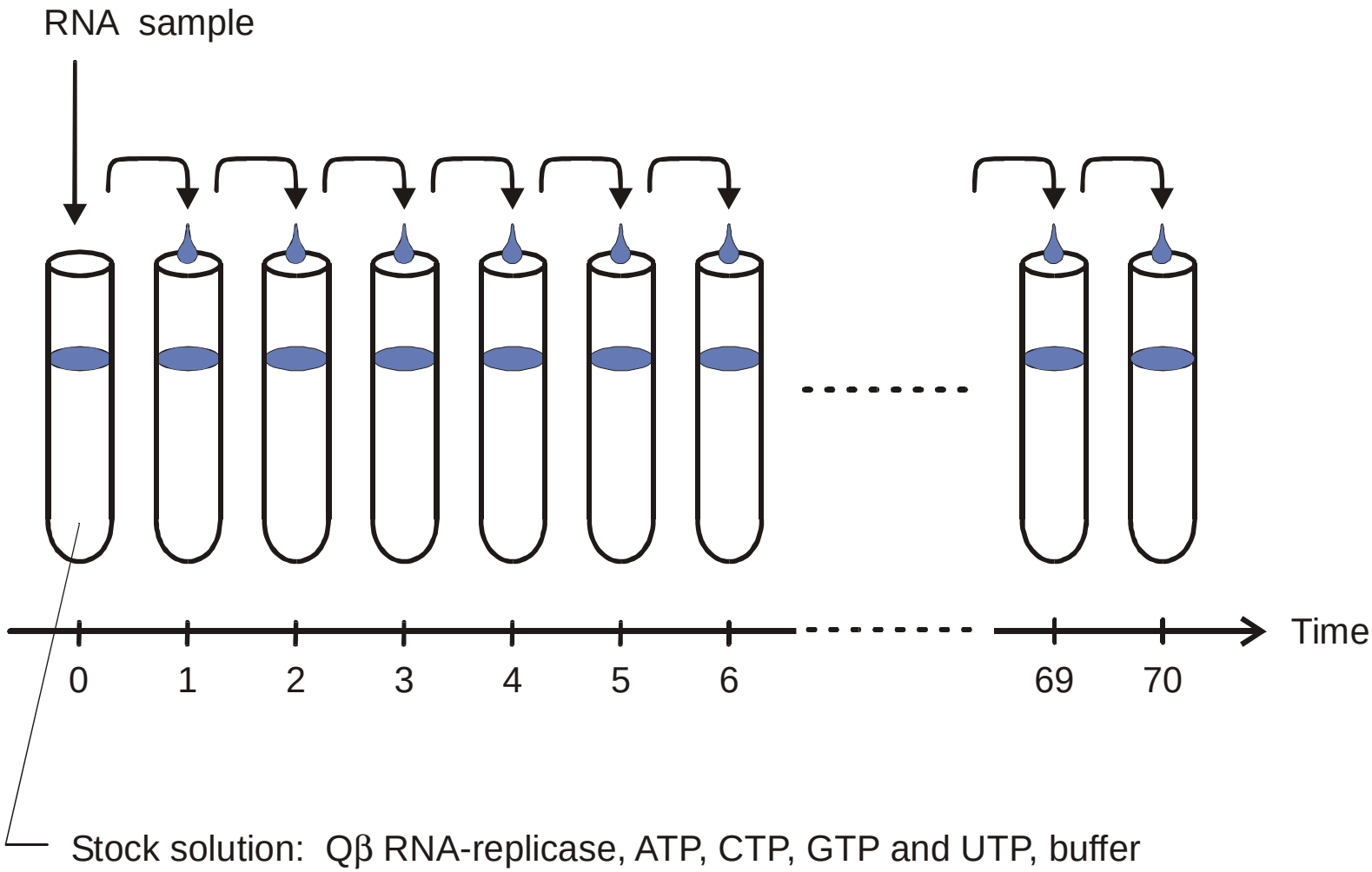
Forty Years of In Vitro Evolution**

*Gerald F. Joyce**

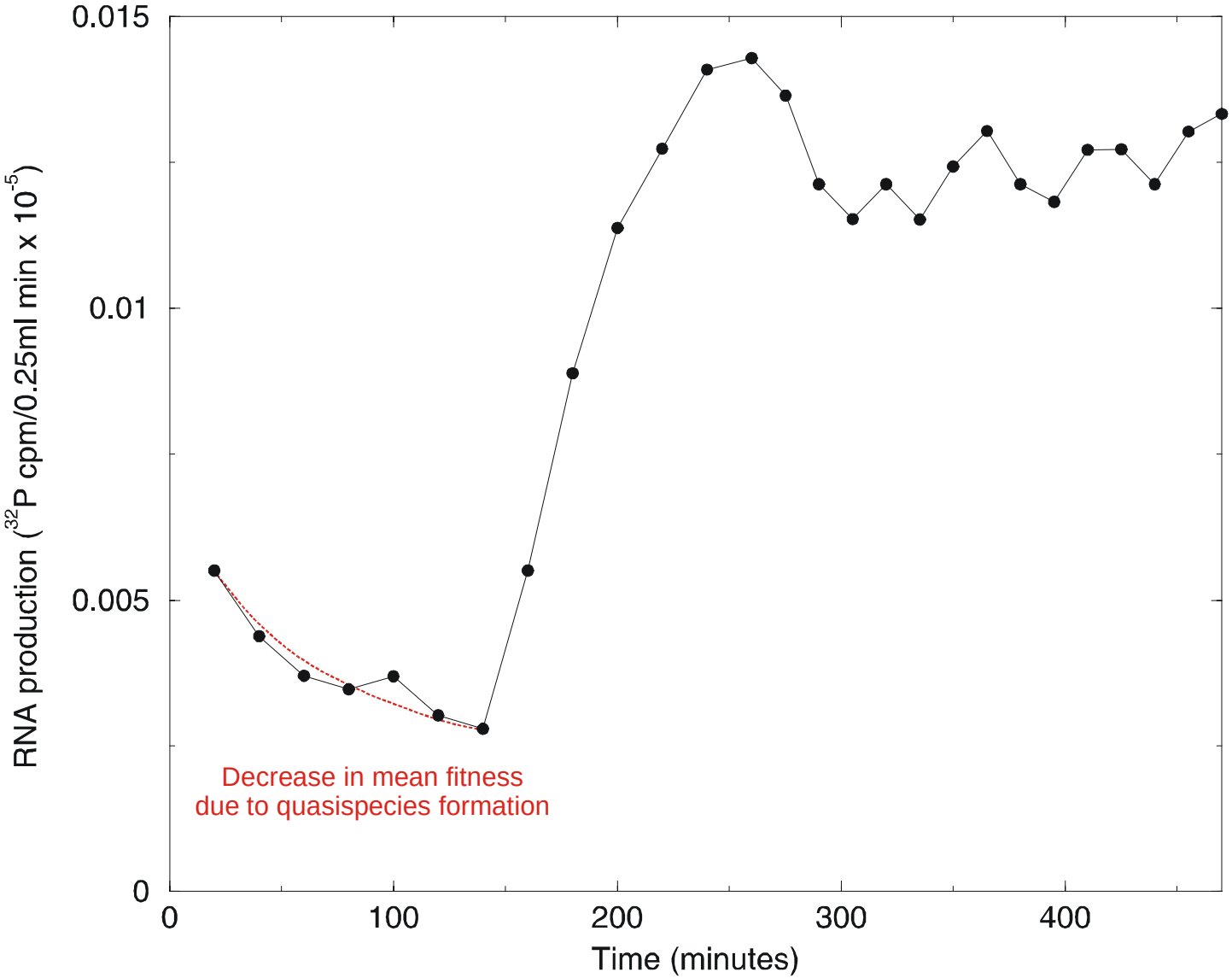
Keywords:
catalysis · directed evolution · ribozyme · RNA

Dedicated to Leslie Orgel on the occasion of his 80th birthday



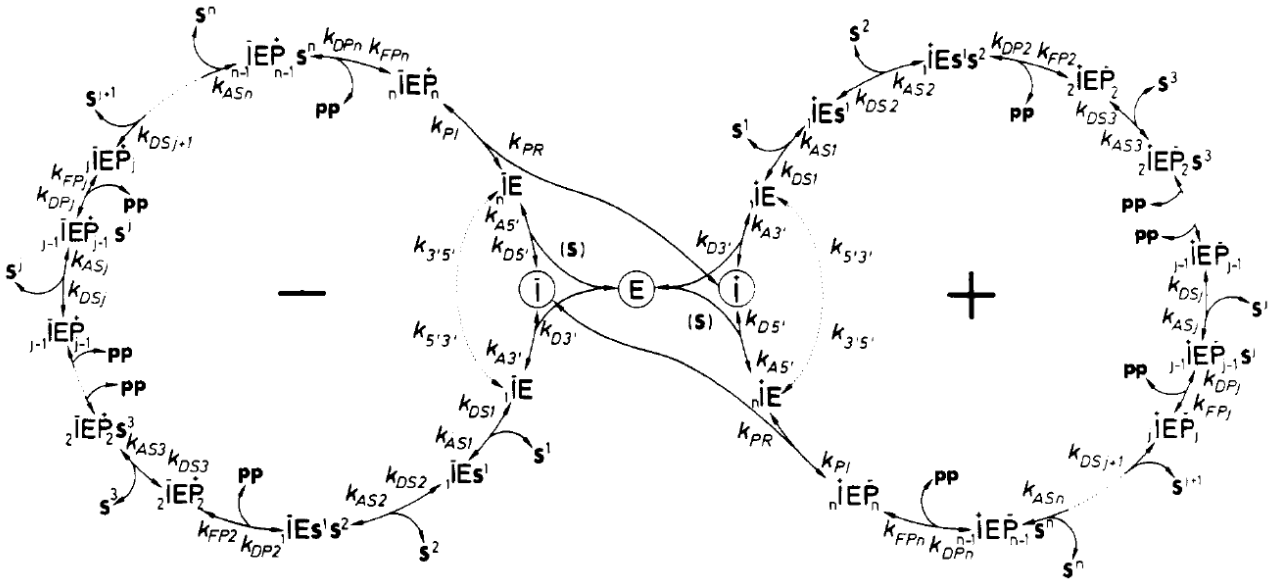


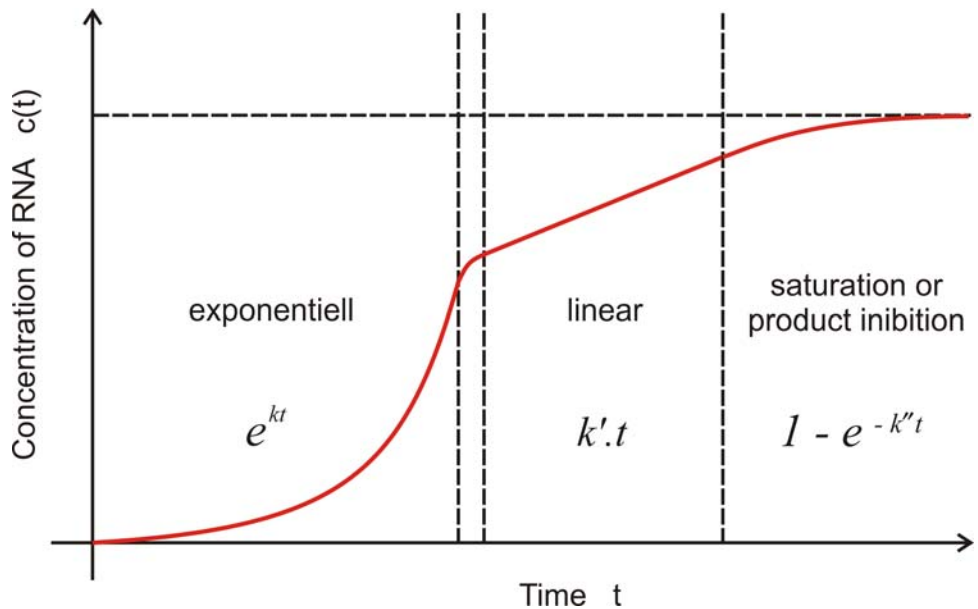
Application of serial transfer technique to evolution of RNA in the test tube



The increase in RNA production rate during a serial transfer experiment

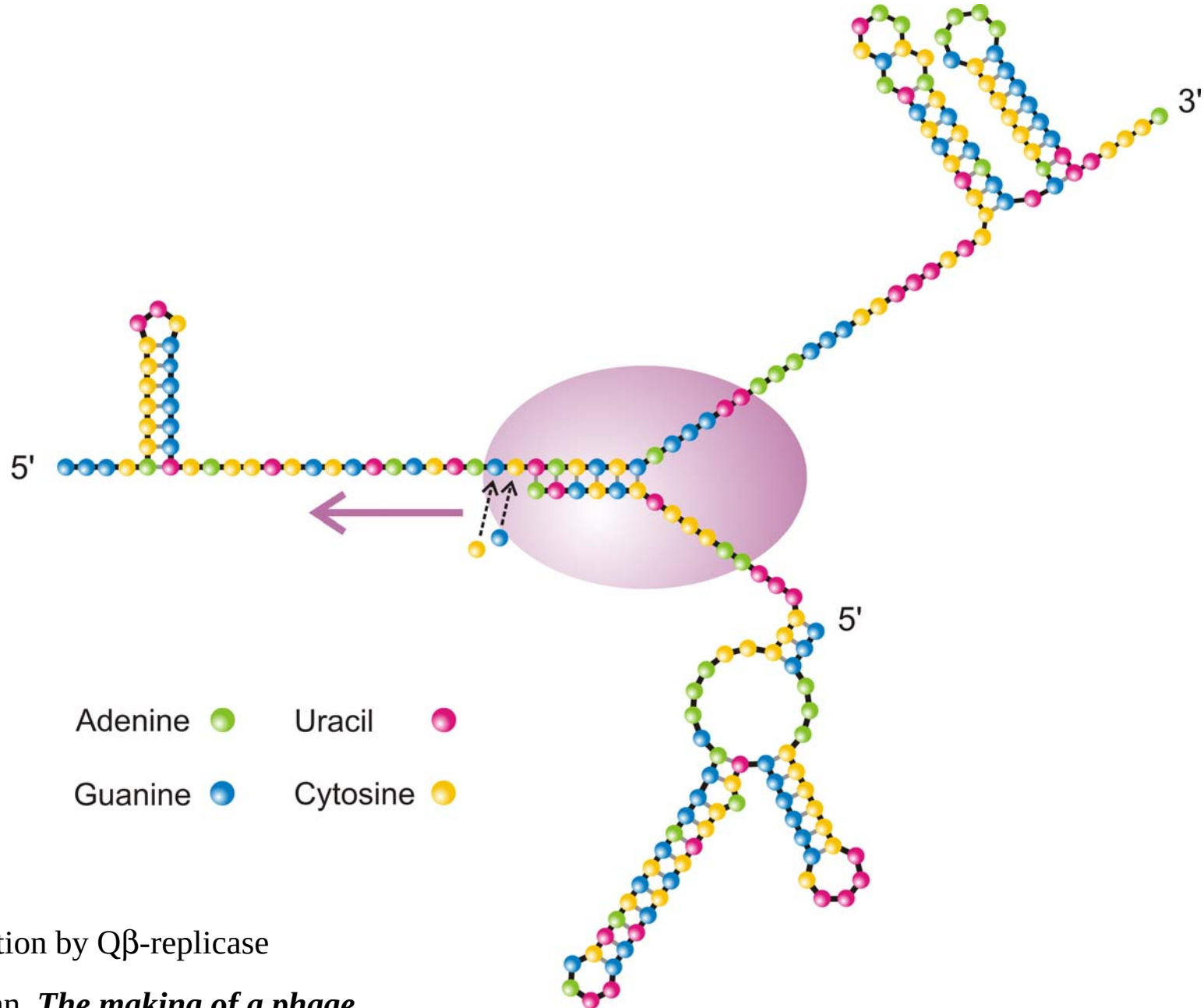
Christof K. Biebricher,
1941-2009





Kinetics of RNA replication

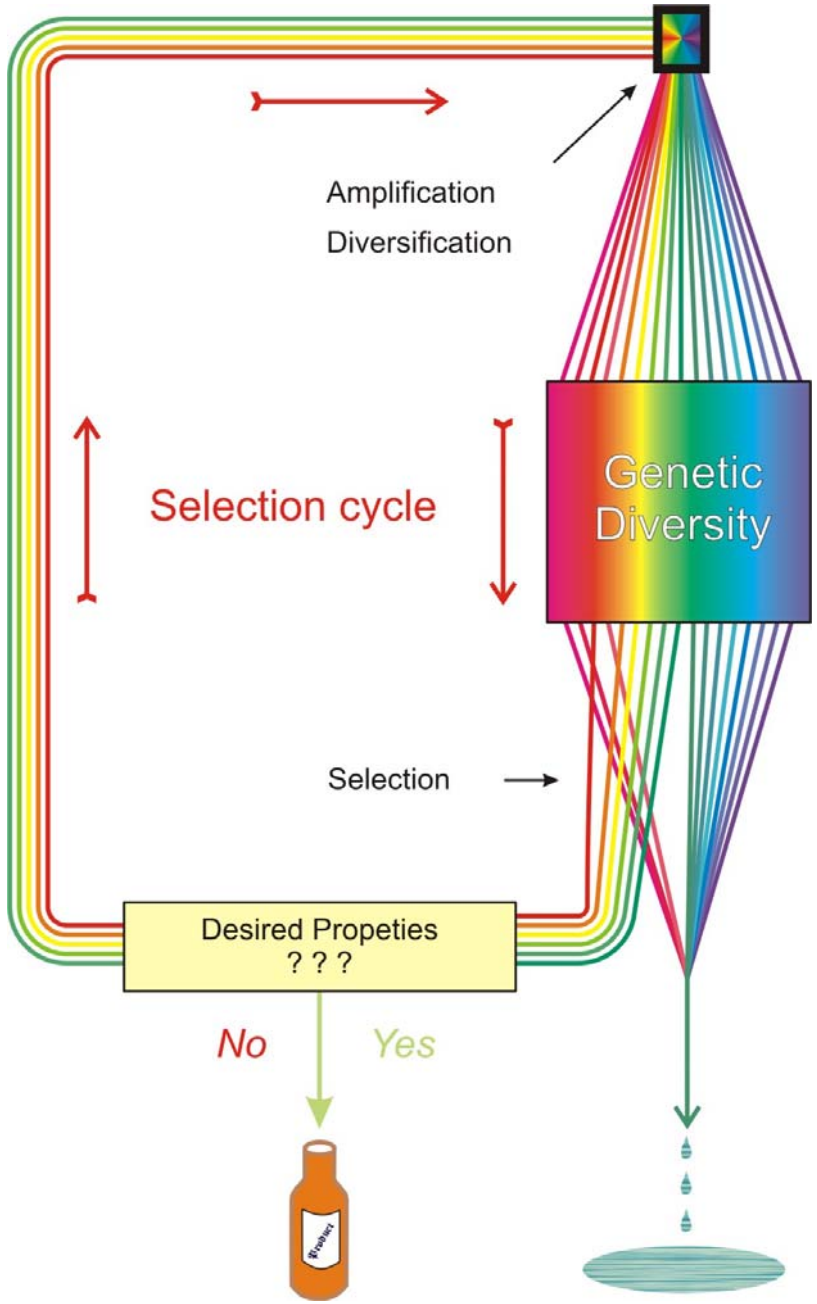
C.K. Biebricher, M. Eigen, W.C. Gardiner, Jr.
Biochemistry **22**:2544-2559, 1983

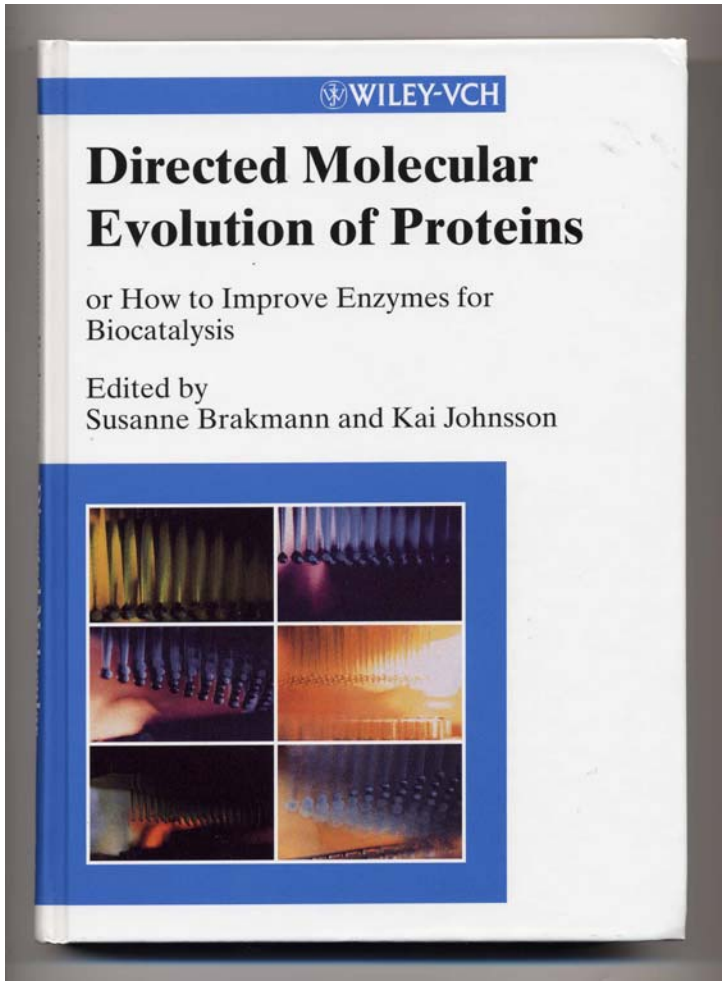


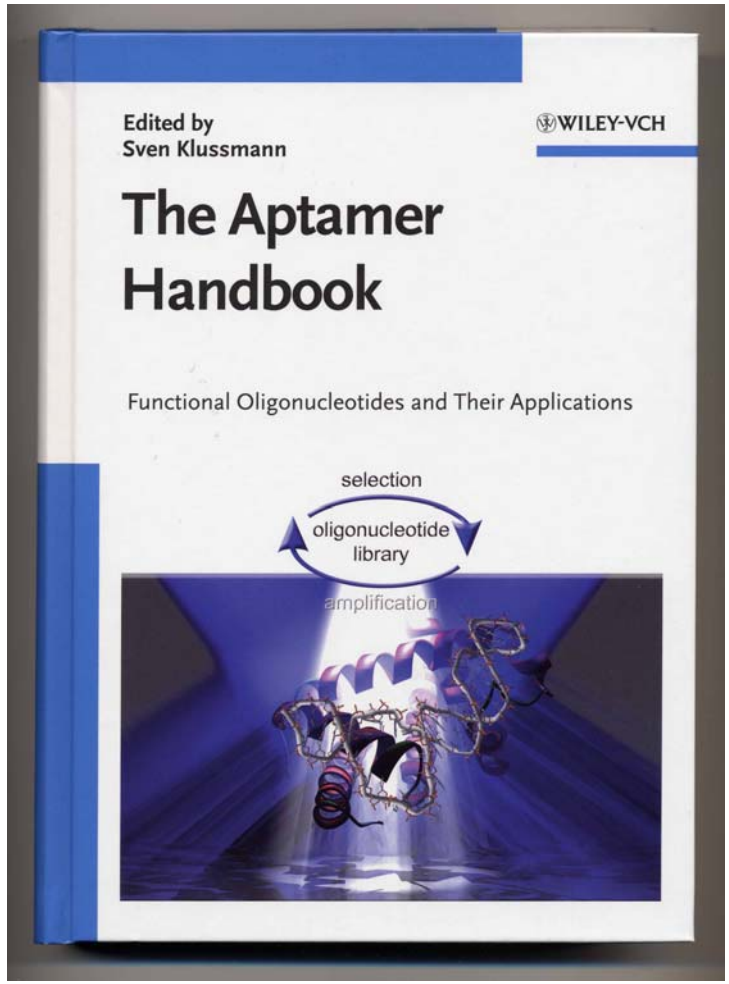
RNA replication by Qβ-replicase

C. Weissmann, ***The making of a phage***.
FEBS Letters **40** (1974), S10-S18

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





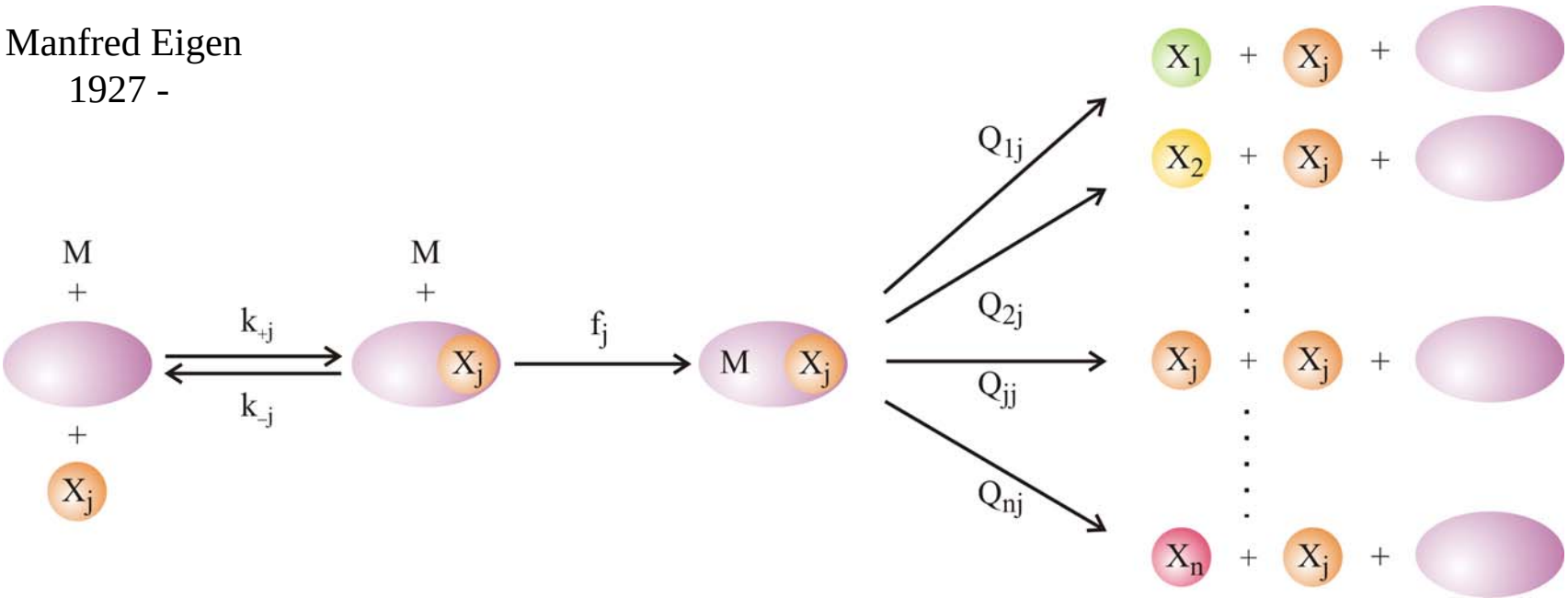


Application of molecular evolution to problems in biotechnology

Manfred Eigen
1927 -

$$\frac{dx_j}{dt} = \sum_{i=1}^{n} W_{ji}\, x_i - x_j\, \Phi\,; \quad j = 1, 2, \ldots, n$$

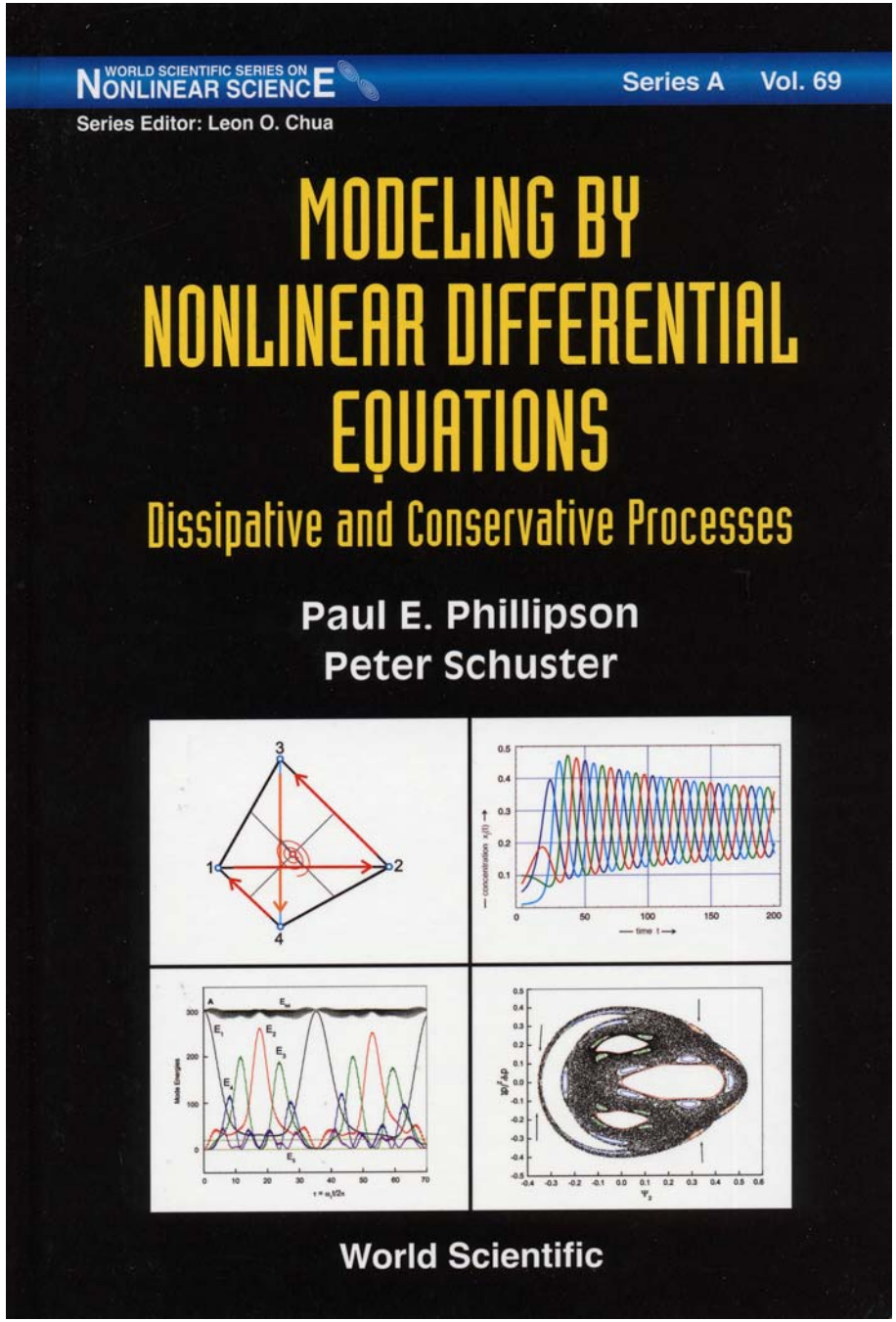
$$\Phi = \sum_{i=1}^{n} f_i\, x_i \Big/ \sum_{i=1}^{n} x_i$$



Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,
M. Eigen & P. Schuster.1977. *Naturwissenschaften* 64:541, 65:7 und 65:341

Modeling evolution at the molecular level



Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virus Research 107 (2005) 115–116
Virus Research
www.elsevier.com/locate/virusres

Preface

Antiviral strategy on the horizon

Error catastrophe had its conceptual origins in the middle of the XXth century, when the consequences of mutations on enzymes involved in protein synthesis, as a theory of aging. In those times biological processes were generally perceived differently from today. Infectious diseases were regarded as a fleeting nuisance which would be eliminated through the use of antibiotics and antiviral agents. Microbial variation, although known in some cases, was not thought to be a significant problem for disease control. Variation in differentiated organisms was seen as resulting essentially from exchanges of genetic material associated with sexual reproduction. The problem was to unveil the mechanisms of inheritance, expression of genetic information and metabolism. Few saw that genetic change is occurring at present in all organisms, and still fewer recognized Darwinian principles as essential to the biology of pathogenic viruses and cells. Population geneticists rarely used bacteria or viruses as experimental systems to define concepts in biological evolution. The extent of genetic polymorphism among individuals of the same biological species came as a surprise when the first results on comparison of electrophoretic mobility of enzymes were obtained. With the advent of in vitro DNA recombination, and rapid nucleic acid sequencing techniques, molecular analyses of genomes reinforced the conclusion of extreme inter-individual genetic variation within the same species. Now, due largely to spectacular progress in comparative genomics, we see cellular DNAs, both prokaryotic and eukaryotic, as highly dynamic. Most cellular processes, including such essential information-bearing and transferring events as genome replication, transcription and translation, are increasingly perceived as inherently inaccurate. Viruses, and in particular RNA viruses, are among the most extreme examples of exploitation of replication inaccuracy for survival.

Error catastrophe, or the loss of meaningful genetic information through excess genetic variation, was formulated in quantitative terms as a consequence of quasispecies theory, which was first developed to explain self-organization and adaptability of primitive replicons in early stages of life. Recently, a conceptual extension of error catastrophe that could be defined as "induced genetic deterioration" has emerged as a possible antiviral strategy. This is the topic of the current special issue of ***Virus Research***.

Few would nowadays doubt that one of the major obstacles for the control of viral disease is short-term adaptability of viral pathogens. Adaptability of viruses follows the same Darwinian principles that have shaped biological evolution over eons, that is, repeated rounds of reproduction with genetic variation, competition and selection, often perturbed by random events such as statistical fluctuations in population size. However, with viruses the consequences of the operation of these very same Darwinian principles are felt within very short times. Short-term evolution (within hours and days) can be also observed with some cellular pathogens, with subsets of normal cells, and cancer cells. The nature of RNA viral pathogens begs for alternative antiviral strategies, and forcing the virus to cross the critical error threshold for maintenance of genetic information is one of them.

The contributions to this volume have been chosen to reflect different lines of evidence (both theoretical and experimental) on which antiviral designs based on genetic deterioration inflicted upon viruses are being constructed. Theoretical studies have explored the copying fidelity conditions that must be fulfilled by any information-bearing replication system for the essential genetic information to be transmitted to progeny. Closely related to the theoretical developments have been numerous experimental studies on quasispecies dynamics and their multiple biological manifestations. The latter can be summarized by saying that RNA viruses, by virtue of existing as mutant spectra rather than defined genetic entities, remarkably expand their potential to overcome selective pressures intended to limit their replication. Indeed, the use of antiviral inhibitors in clinical practice and the design of vaccines for a number of major RNA virus-associated diseases, are currently presided by a sense of uncertainty. Another line of growing research is the enzymology of copying fidelity by viral replicases, aimed at understanding the molecular basis of mutagenic activities. Error catastrophe as a potential new antiviral strategy received an important impulse by the observation that ribavirin (a licensed antiviral nucleoside analogue) may be exerting, in some systems, its antiviral activity through enhanced mutagenesis. This has encouraged investigations on new mutagenic base analogues, some of them used in anticancer chemotherapy. Some chapters summarize these important biochemical studies on cell entry pathways and metabolism of mutagenic agents, that may find new applications as antiviral agents.

This volume intends to be basically a progress report, an introduction to a new avenue of research, and a realistic appraisal of the many issues that remain to be investigated. In this respect, I can envisage (not without many uncertainties) at least three lines of needed research: (i) One on further understanding of quasispecies dynamics in infected individuals to learn more on how to apply combinations of virus-specific mutagens and inhibitors in an effective way, finding synergistic combinations and avoiding antagonistic ones as well as severe clinical side effects. (ii) Another on a deeper understanding of the metabolism of mutagenic agents, in particular base and nucleoside analogues. This includes identification of the transporters that carry them into cells, an understanding of their metabolic processing, intracellular stability and alterations of nucleotide pools, among other issues. (iii) Still another line of needed research is the development of new mutagenic agents specific for viruses, showing no (or limited) toxicity for cells. Some advances may come from links with anticancer research, but others should result from the designs of new molecules, based on the structures of viral polymerases. I really hope that the reader finds this issue not only to be an interesting and useful review of the current situation in the field, but also a stimulating exposure to the major problems to be faced.

The idea to prepare this special issue came as a kind invitation of Ulrich Desselberger, former Editor of ***Virus Research***, and then taken enthusiastically by Luis Enjuanes, recently appointed as Editor of ***Virus Research***. I take this opportunity to thank Ulrich, Luis and the Editor-in-Chief of ***Virus Research***, Brian Mahy, for their continued interest and support to the research on virus evolution over the years.

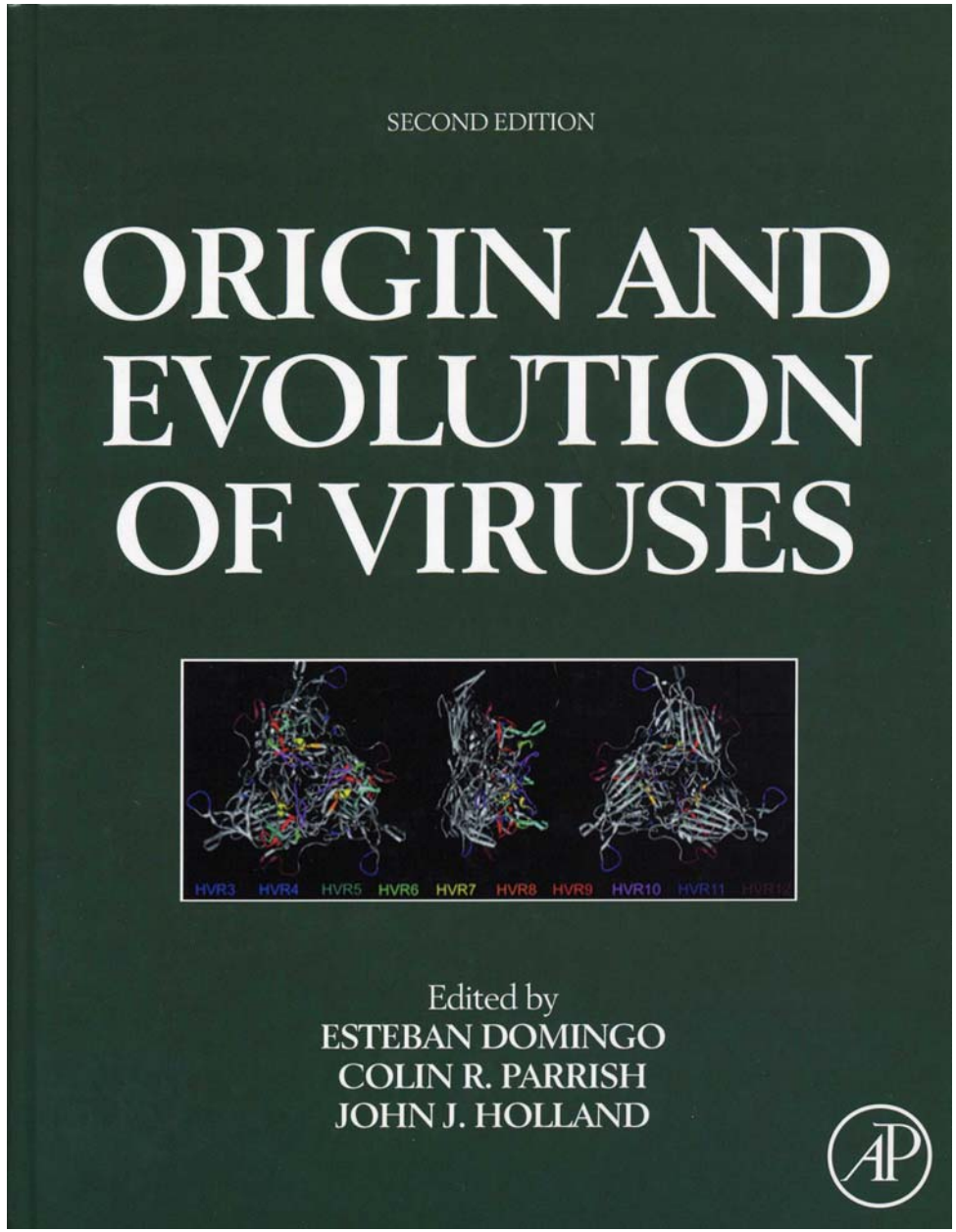
My thanks go also to the 19 authors who despite their busy schedules have taken time to prepare excellent manuscripts, to Elsevier staff for their prompt responses to my requests, and, last but not least, to Ms. Lucía Horrillo from Centro de Biología Molecular "Severo Ochoa" for her patient dealing with the correspondence with authors and the final organization of the issue.

Esteban Domingo
Universidad Autónoma de Madrid
Centro de Biología Molecular "Severo Ochoa"
Consejo Superior de Investigaciones Científicas
Cantoblanco and Valdeolmos
Madrid, Spain
Tel.: +34 91 497 84858/9; fax: +34 91 497 4799
E-mail address: edomingo@cbm.uam.es

Available online 8 December 2004

doi:10.1016/j.virusres.2004.11.001

Molecular evolution of viruses



1. Origins of evolution
2. The „tree of life“
3. Probability and chance
4. Natural and artificial selection
5. **Neutrality and random drift**
6. Contingency and history

Evolution *in silico*

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

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

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

9 March 1998; accepted 17 April 1998

Continuity in Evolution: On the Nature of Transitions

Walter Fontana and Peter Schuster

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

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

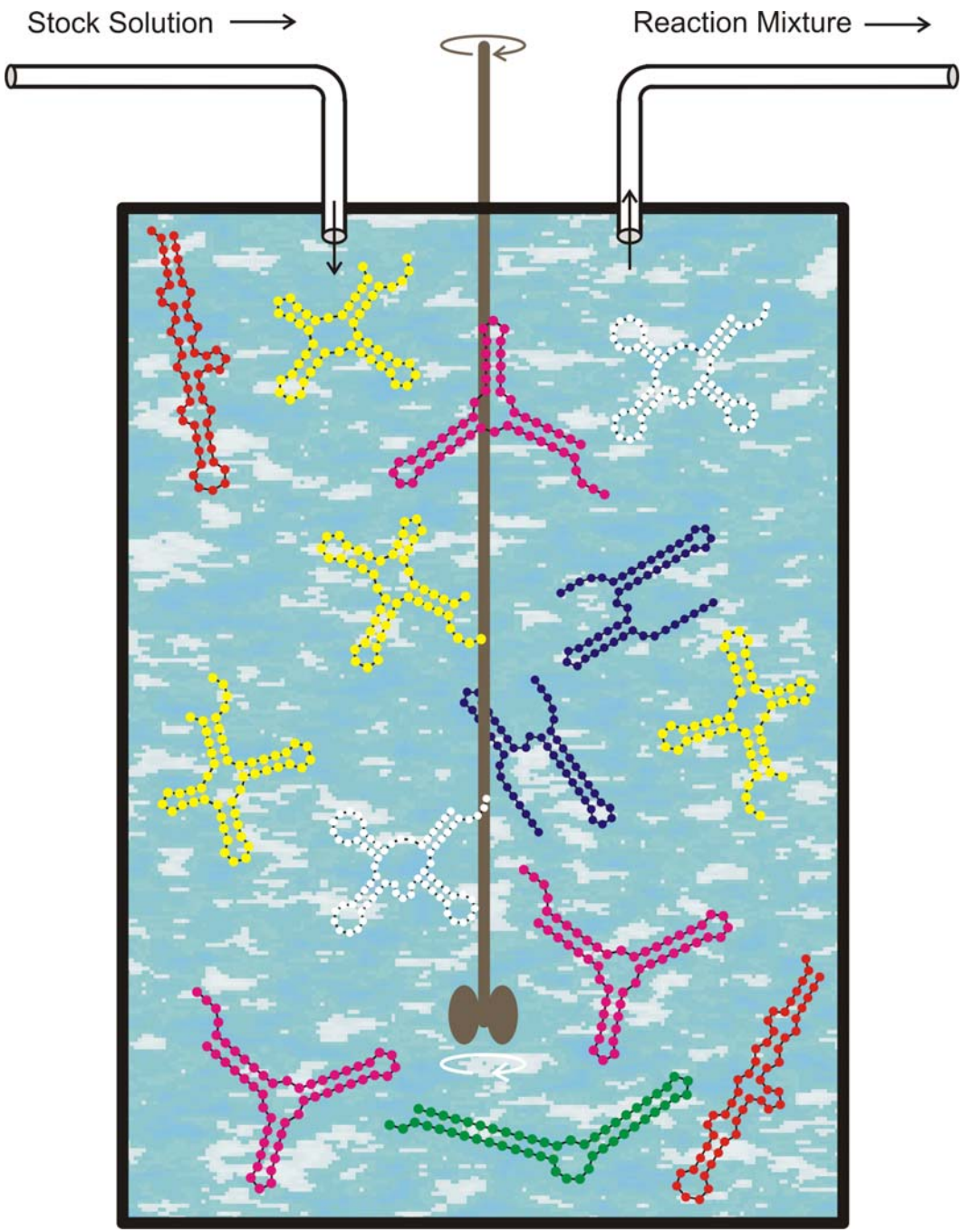
We focus on the narrow domain of RNA secondary structure, which is currently the simplest computationally tractable, yet realistic phenotype (*2*). This choice enables the definition and exploration of concepts that may prove useful in a wider context. RNA secondary structures represent a coarse level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

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

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

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

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



Replication rate constant:

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

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

Selection constraint:

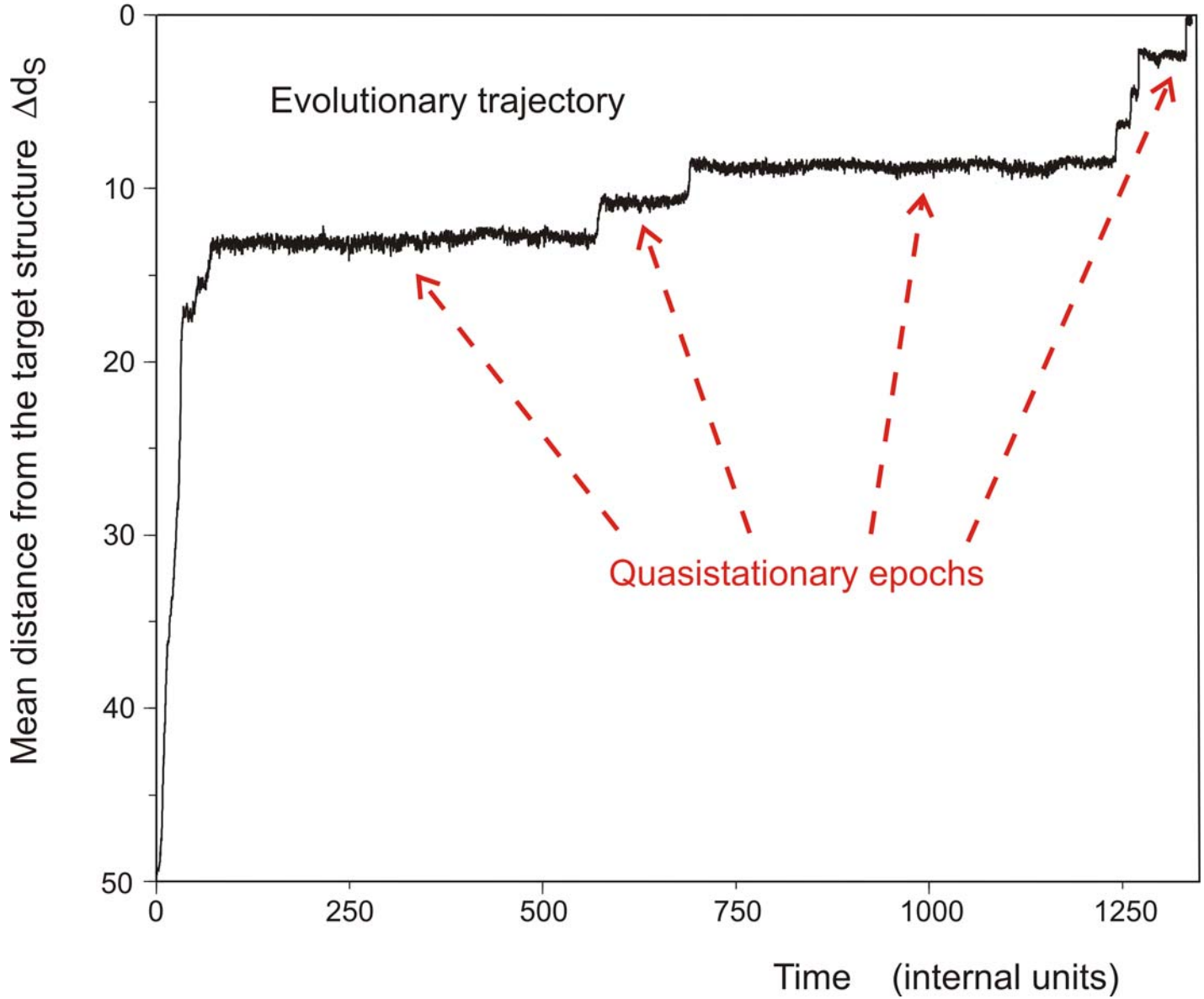
Population size, N = # RNA molecules, is controlled by the flow

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

Mutation rate:

p = 0.001 / site × replication

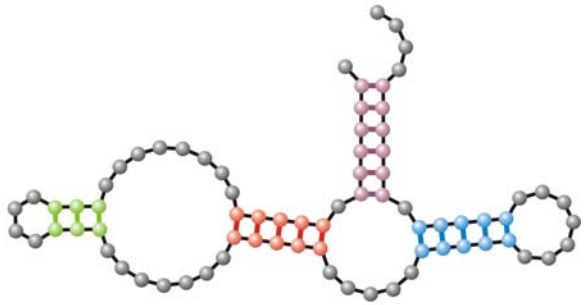
The flowreactor as a device for **studies** of evolution ***in vitro*** and ***in silico***



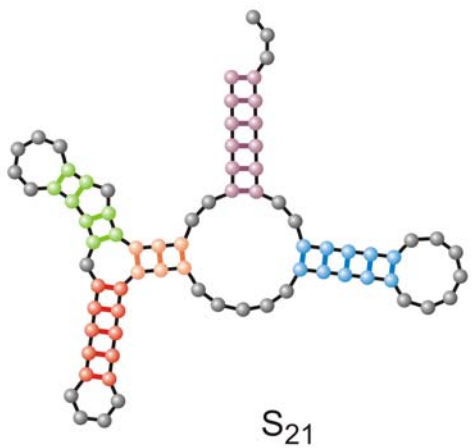
In silico optimization in the flow reactor: Evolutionary Trajectory

Randomly chosen
initial structure

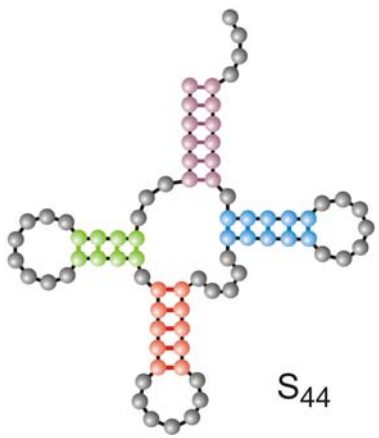
S_0



S_9



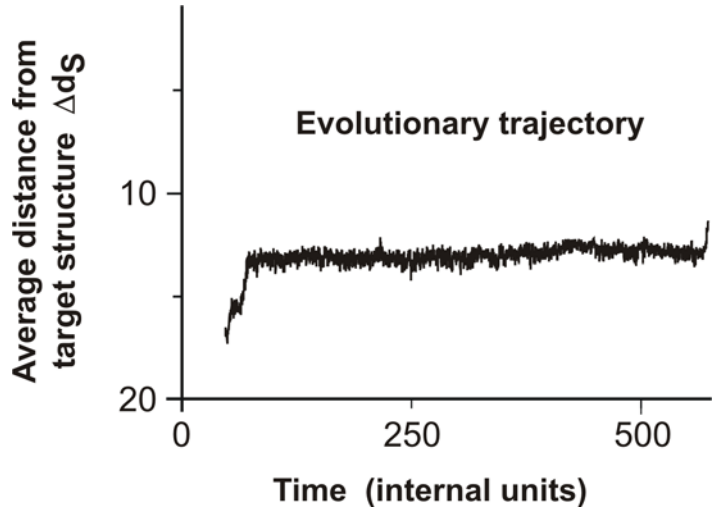
S_{21}



S_{44}

Phenylalanyl-tRNA
as target structure

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

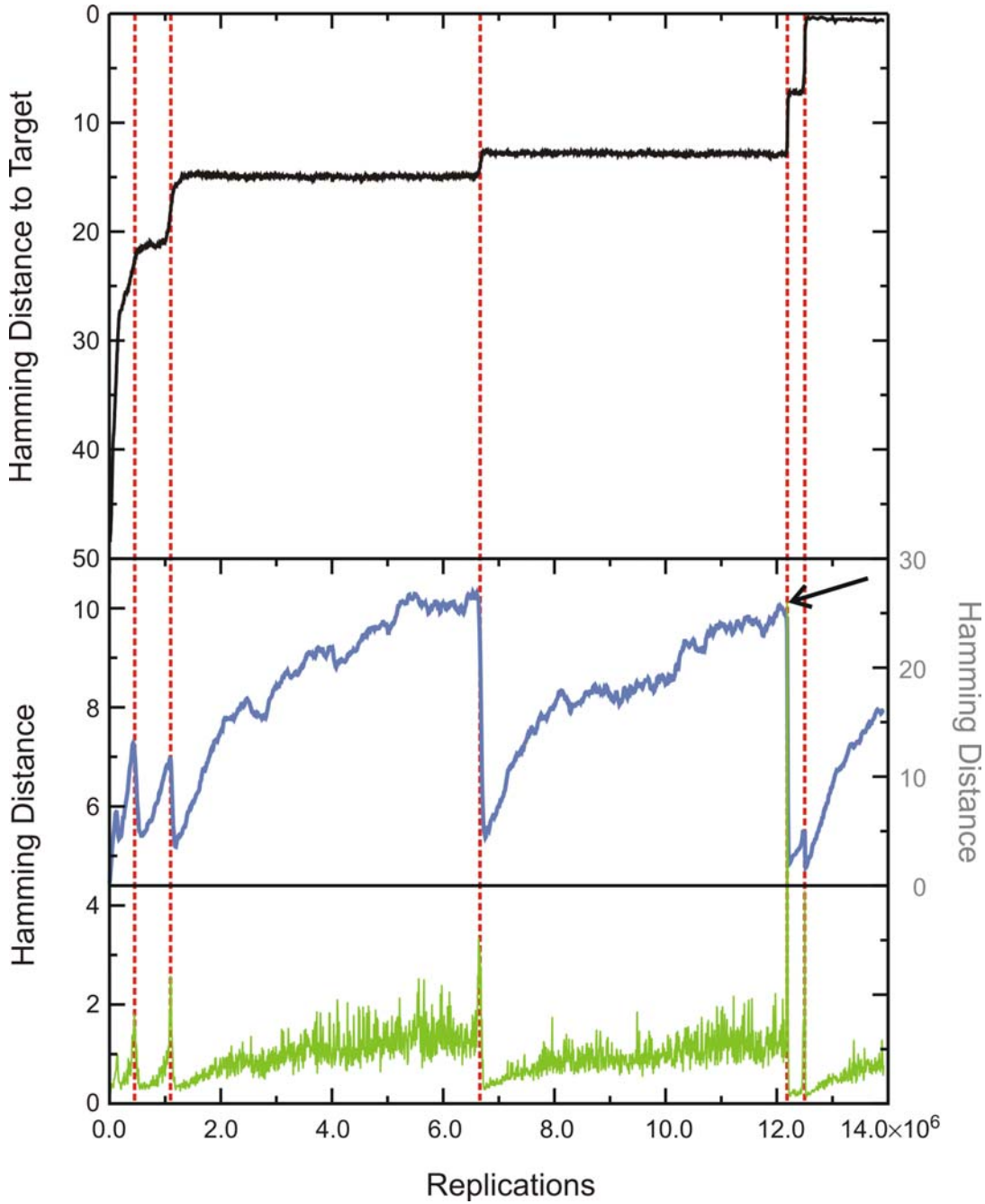
Neutral point mutations leave the molecular structure unchanged

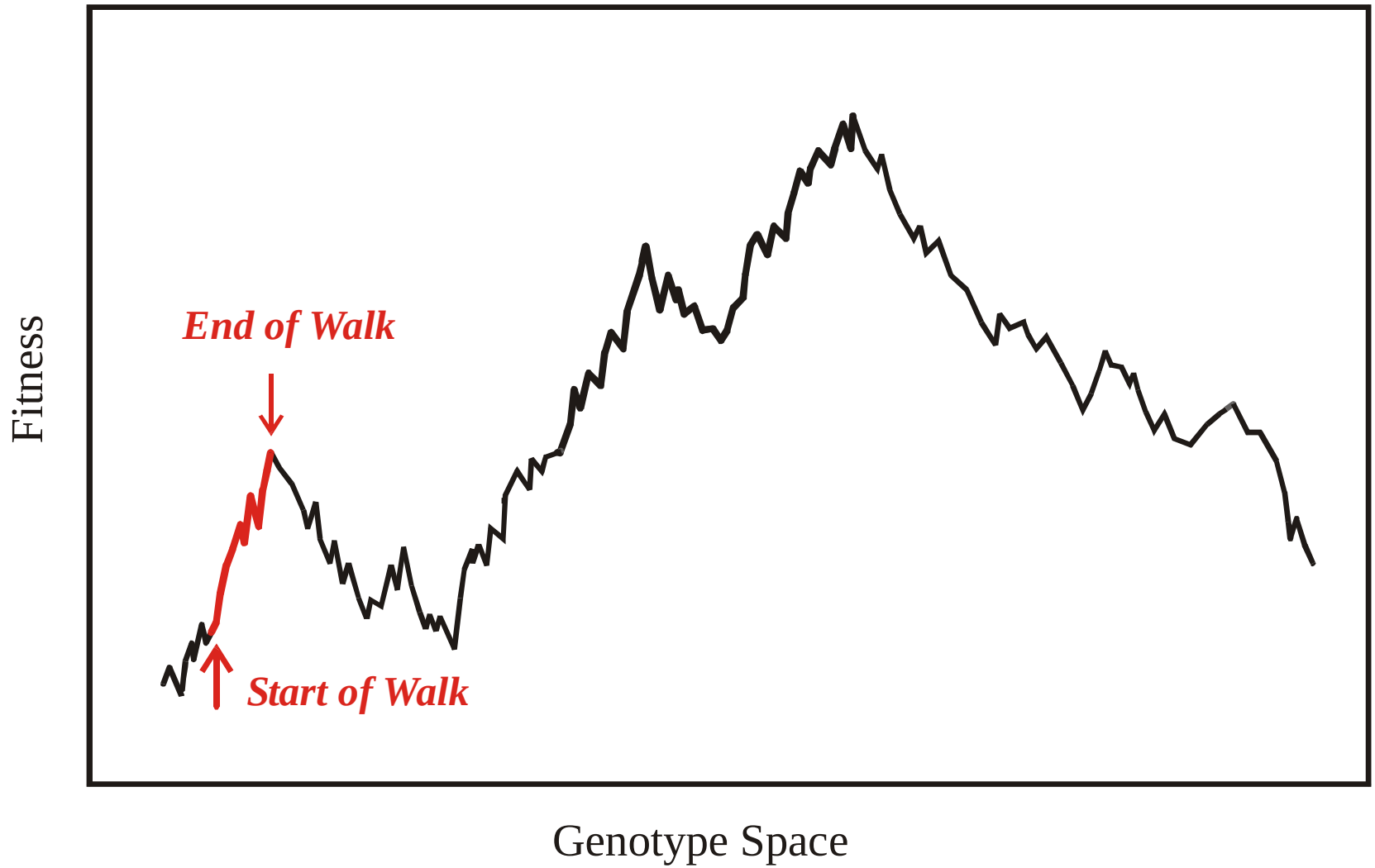
Neutral genotype evolution during phenotypic stasis

Evolutionary trajectory

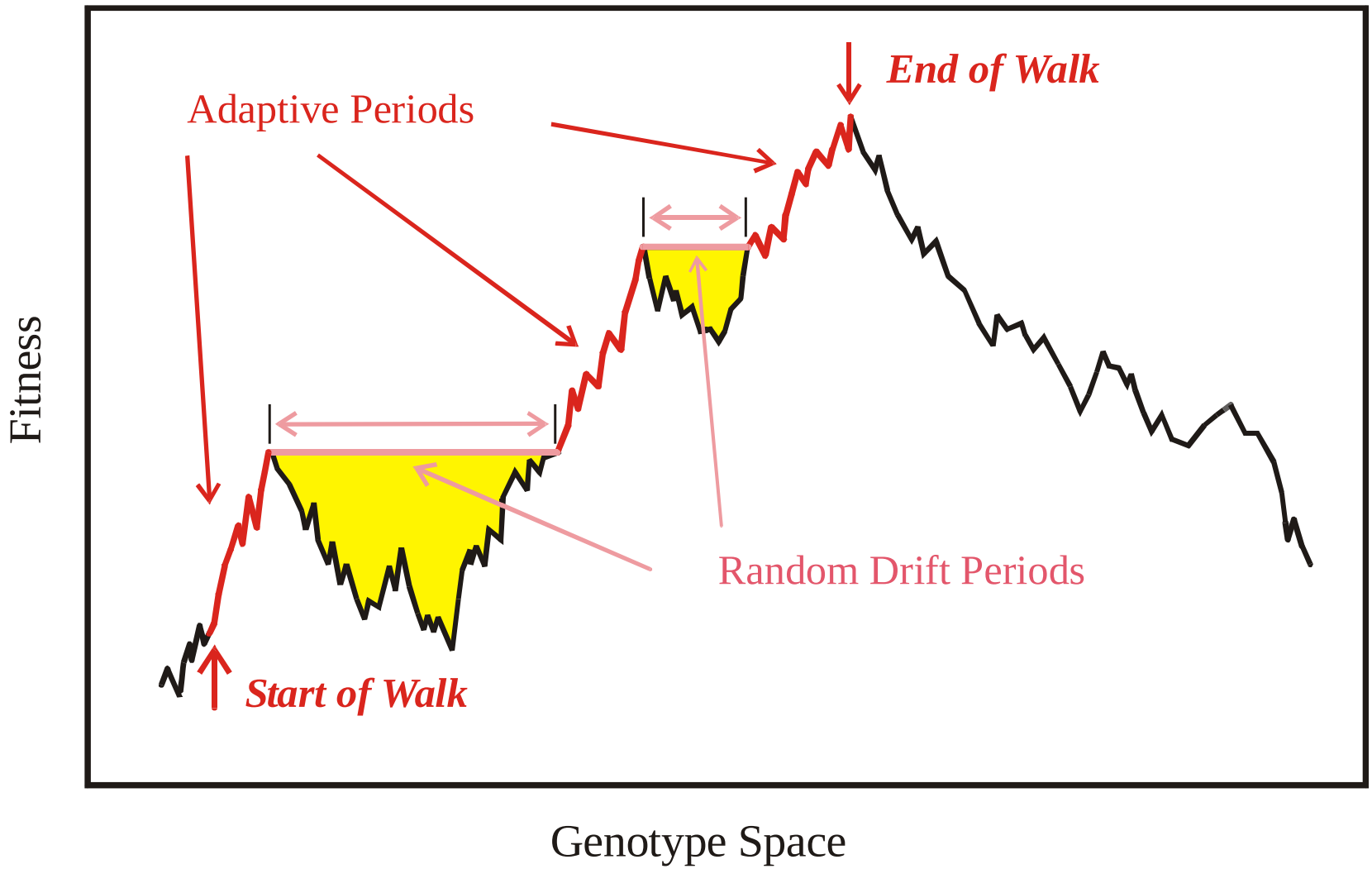
Spreading of the population on neutral networks

Drift of the population center in sequence space





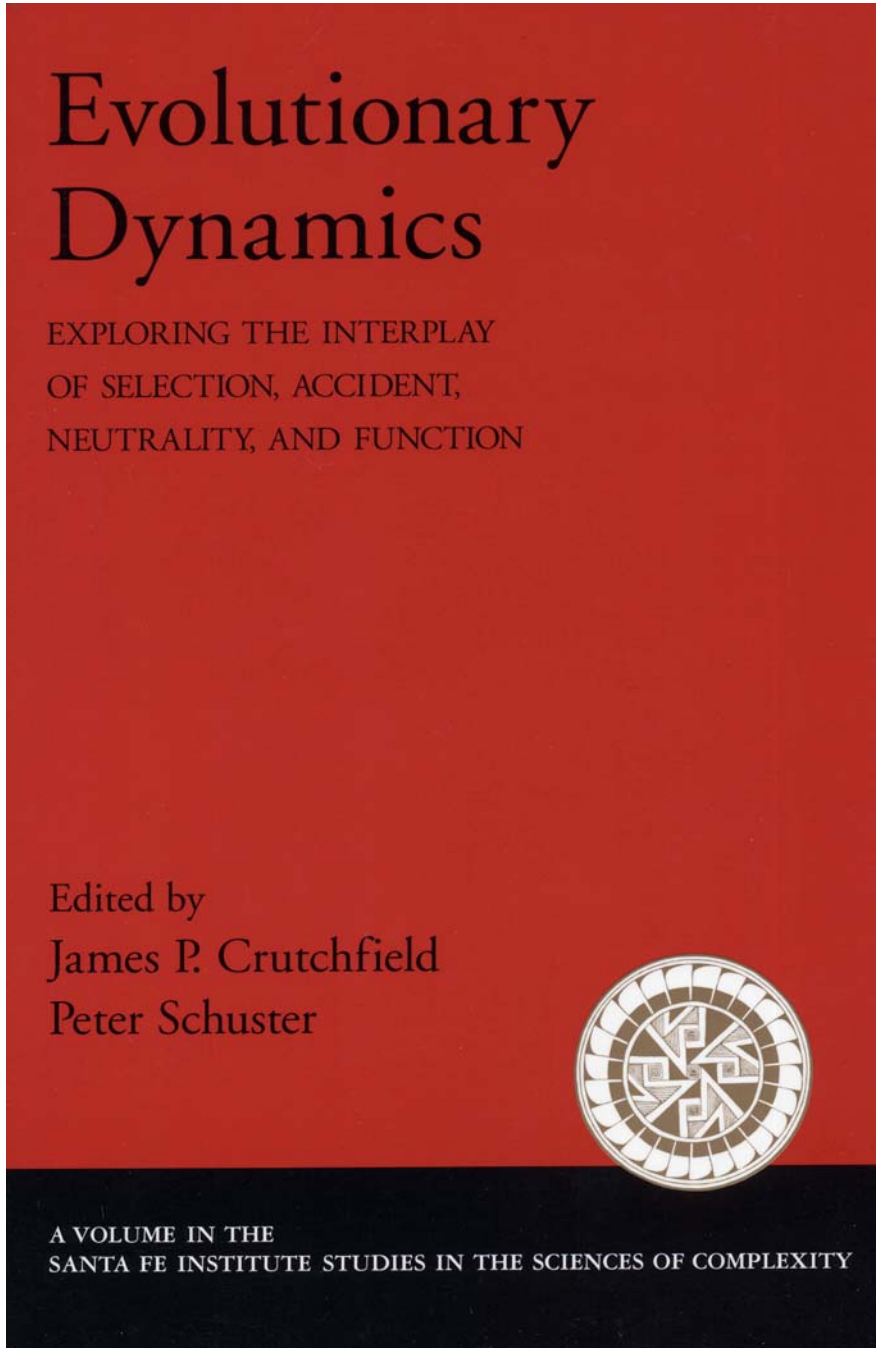
Evolutionary optimization in absence of neutral paths in sequence space



Evolutionary optimization including neutral paths in sequence space

All interesting games have deterministic and random components

Evolution is a result of deterministic and accidental components

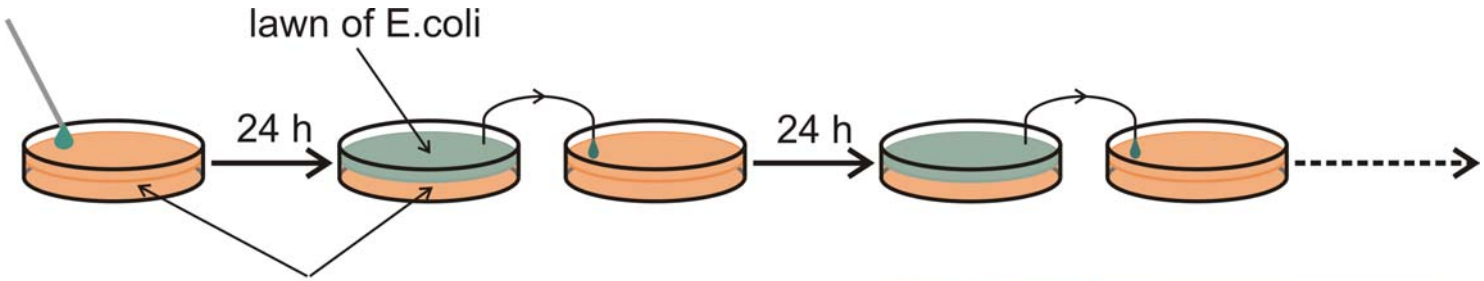


1. Origins of evolution
2. The „tree of life“
3. Probability and chance
4. Natural and artificial selection
5. Neutrality and random drift
6. **Contingency and history**

Richard Lenski, 1956 -

Bacterial evolution under controlled conditions: A twenty years experiment.

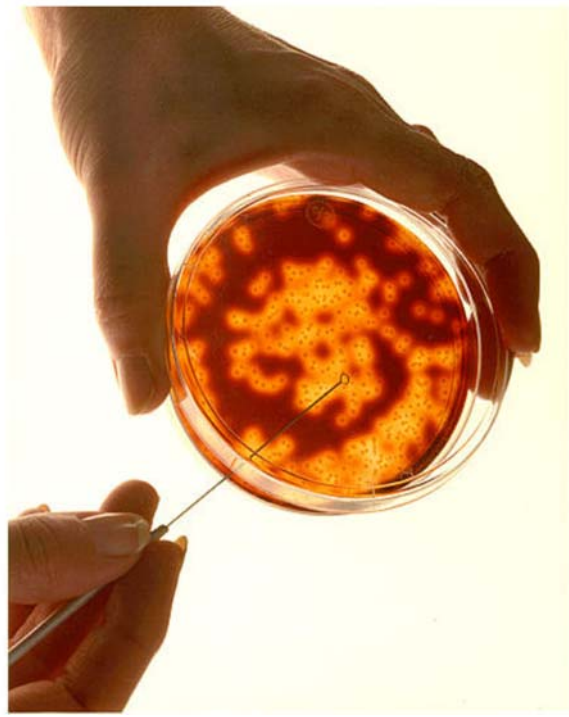
Richard Lenski, University of Michigan, East Lansing



medium supports $\approx 5 \times 10^8$ bacteria

1 day	$\approx$	6.67 generations
1 month	$\approx$	200 generations
1 year	$\approx$	**2400 generations**

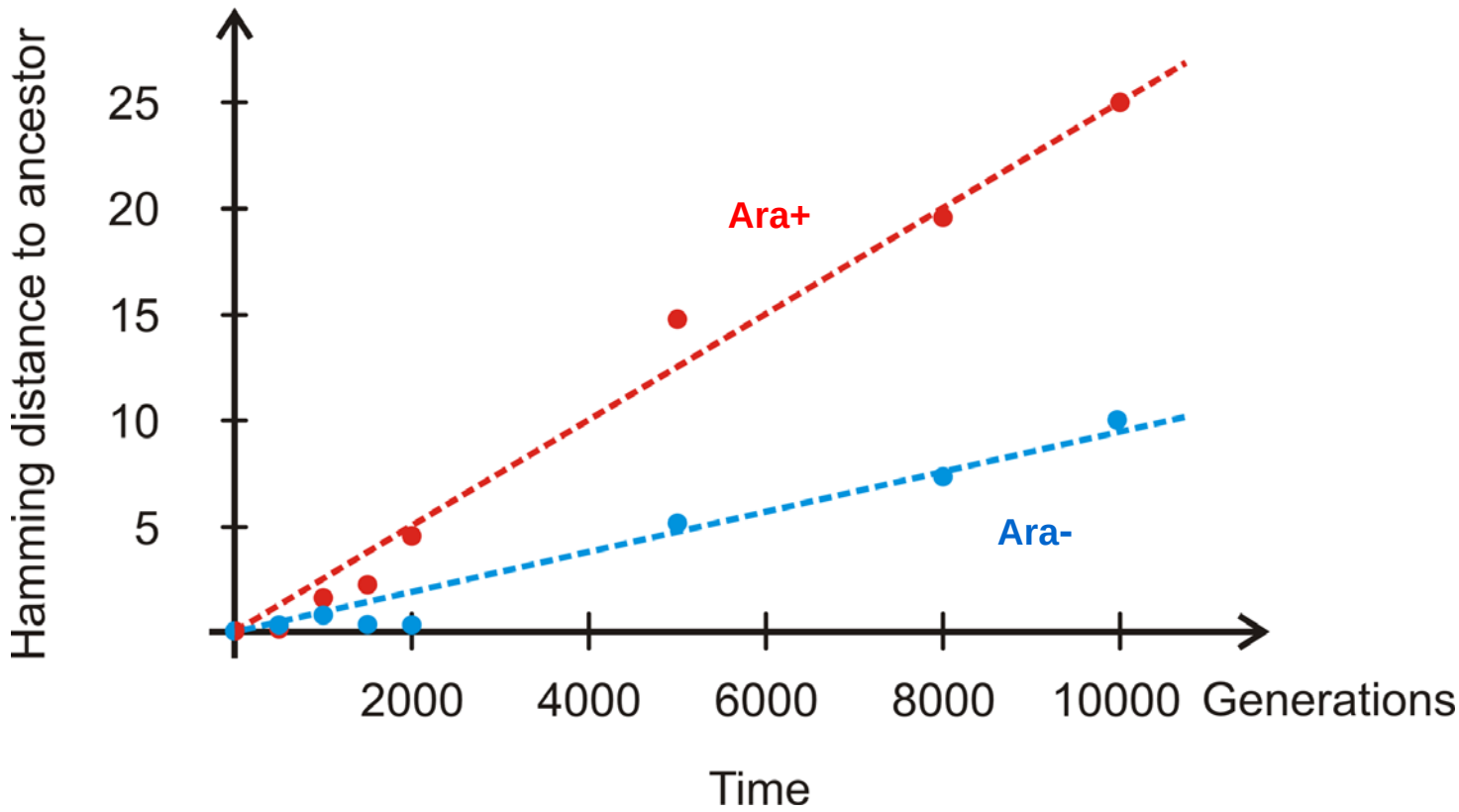
Serial transfer of bacterial cultures in Petri dishes



Bacterial evolution under controlled conditions: A twenty years experiment.

Richard Lenski, University of Michigan, East Lansing

The twelve populations of Richard Lenski‘s long time evolution experiment



Variation of genotypes in a bacterial serial transfer experiment

D. Papadopoulos, D. Schneider, J. Meier-Eiss, W. Arber, R. E. Lenski, M. Blot. *Genomic evolution during a 10,000-generation experiment with bacteria*. Proc.Natl.Acad.Sci.USA **96** (1999), 3807-3812

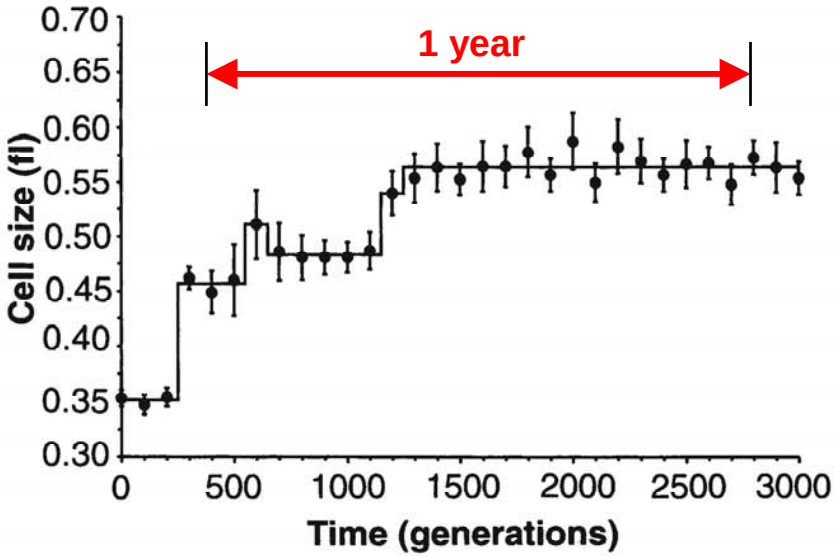


Fig. 1. Change in average cell size (1 fl = 10^{-15} L) in a population of *E. coli* during 3000 generations of experimental evolution. Each point is the mean of 10 replicate assays (*22*). Error bars indicate 95% confidence intervals. The solid line shows the best fit of a step-function model to these data (Table 1).

Epochal evolution of bacteria in serial transfer experiments under constant conditions

S. F. Elena, V. S. Cooper, R. E. Lenski. *Punctuated evolution caused by selection of rare beneficial mutants*. Science **272** (1996), 1802-1804

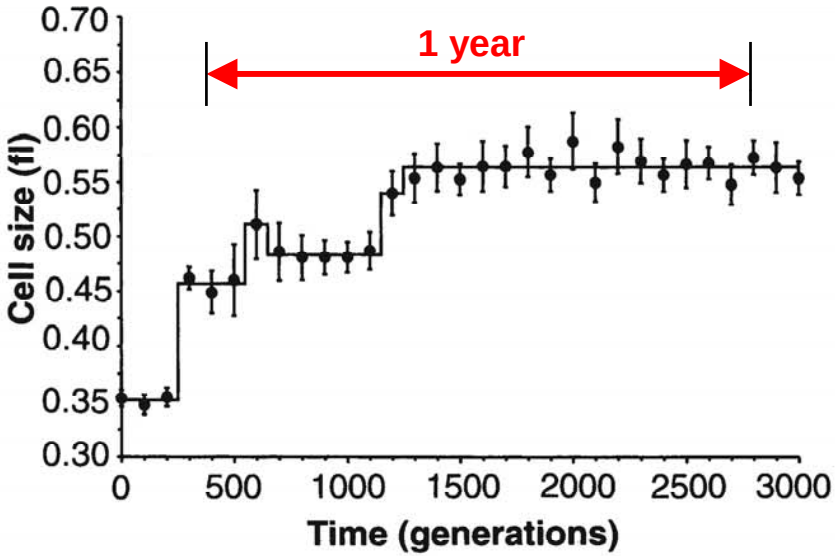


Fig. 1. Change in average cell size (1 fl = 10^{-15} L) in a population of *E. coli* during 3000 generations of experimental evolution. Each point is the mean of 10 replicate assays (*22*). Error bars indicate 95% confidence intervals. The solid line shows the best fit of a step-function model to these data (Table 1).

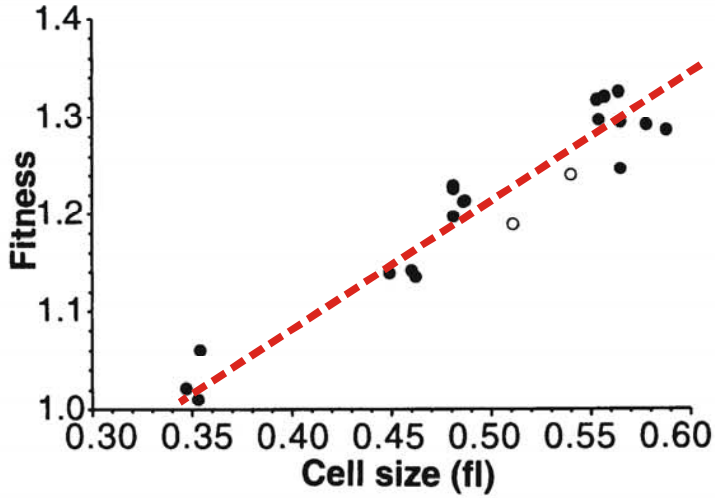
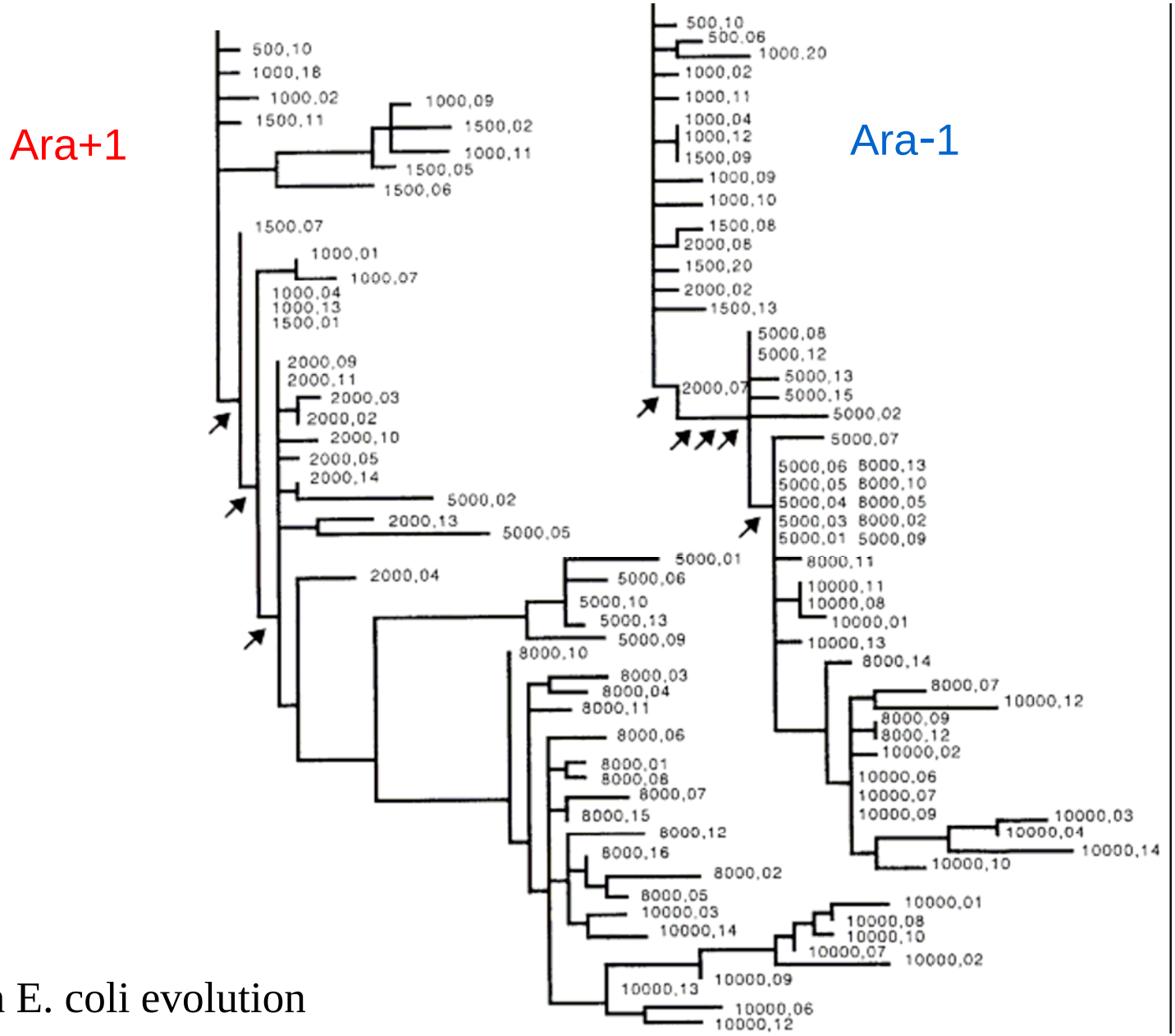


Fig. 2. Correlation between average cell size and mean fitness, each measured at 100-generation intervals for 2000 generations. Fitness is expressed relative to the ancestral genotype and was obtained from competition experiments between derived and ancestral cells (*6*, *7*). The open symbols indicate the only two samples assigned to different steps by the cell size and fitness data.

Epochal evolution of bacteria in serial transfer experiments under constant conditions

S. F. Elena, V. S. Cooper, R. E. Lenski. *Punctuated evolution caused by selection of rare beneficial mutants*. Science **272** (1996), 1802-1804



Phylogeny in E. coli evolution

The twelve populations of Richard Lenski‘s long time evolution experiment
Enhanced turbidity in population A-3

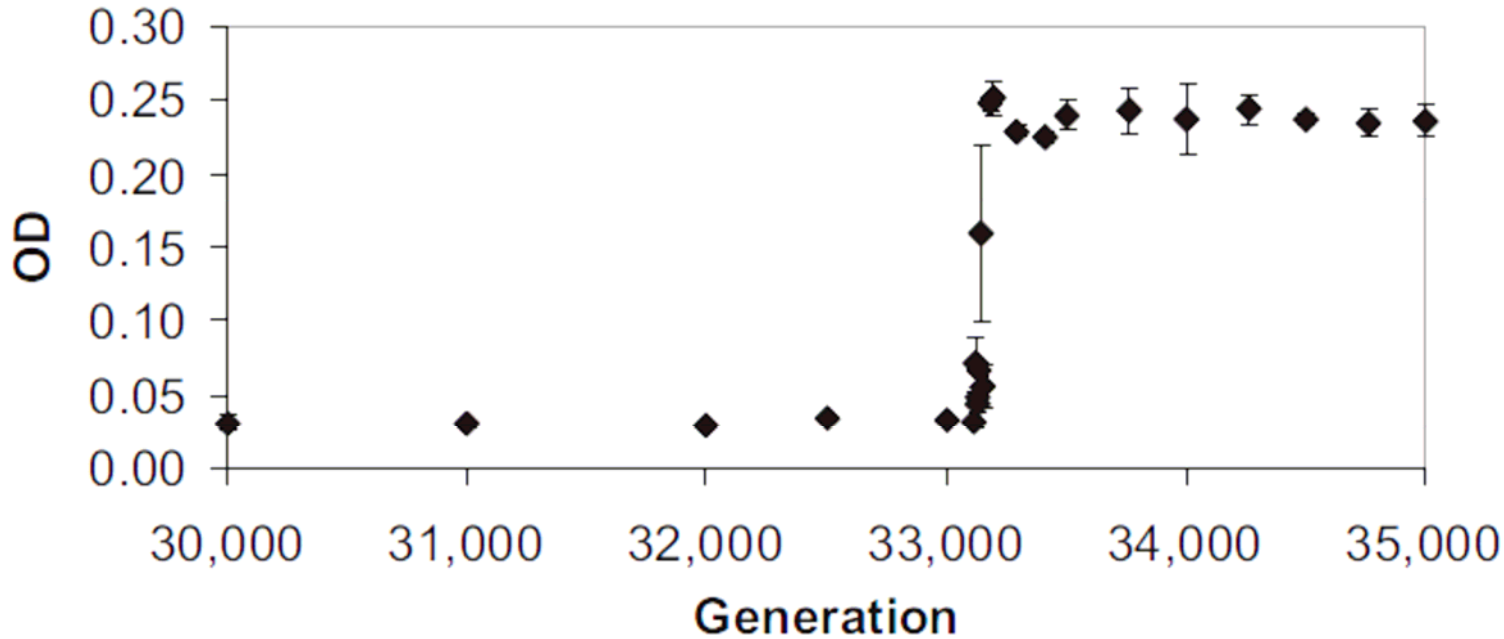


Fig. 1. Population expansion during evolution of the Cit$^+$ phenotype. Samples frozen at various times in the history of population Ara-3 were revived, and three DM25 cultures were established for each generation. Optical density (OD) at 420 nm was measured for each culture at 24 h. Error bars show the range of three values measured for each generation.

Innovation by mutation in long time evolution of Escherichia coli in constant environment

Z.D. Blount, C.Z. Borland, R.E. Lenski. 2008. Proc.Natl.Acad.Sci.USA 105:7899-7906

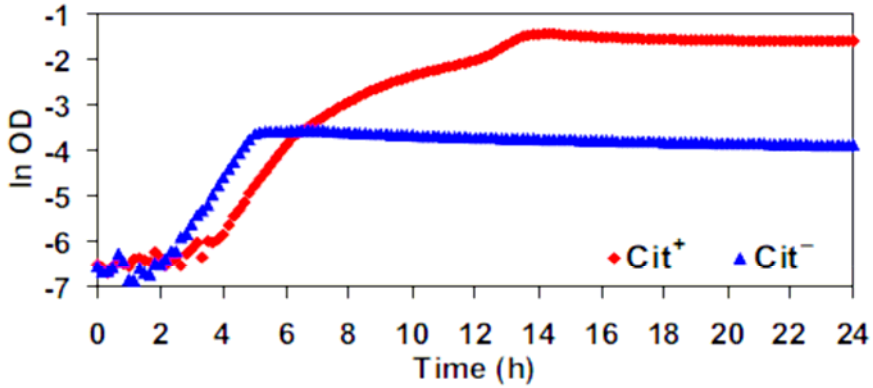


Fig. 2. Growth of Cit⁻ (blue triangles) and Cit⁺ (red diamonds) cells in DM25 medium. Each trajectory shows the average OD for eight replicate mixtures of three clones, all from generation 33,000 of population Ara-3.

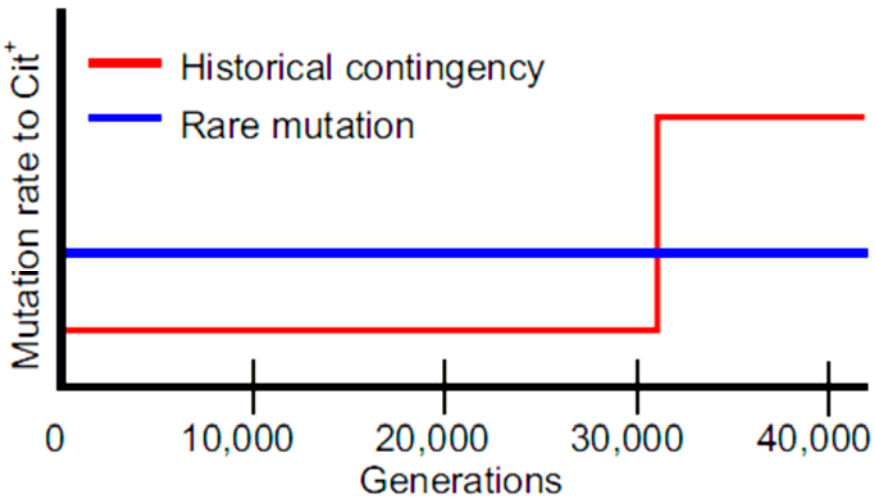


Fig. 3. Alternative hypotheses for the origin of the Cit⁺ function. According to the rare-mutation hypothesis, the probability of mutation from Cit⁻ to Cit⁺ was low but constant over time. Under the historical-contingency hypothesis, the probability of this transition increased when a mutation arose that produced a genetic background with a higher mutation rate to Cit⁺.

Innovation by mutation in long time evolution of Escherichia coli in constant environment

Z.D. Blount, C.Z. Borland, R.E. Lenski. 2008. Proc.Natl.Acad.Sci.USA 105:7899-7906

Table 1. Summary of replay experiments

	First experiment		Second experiment		Third experiment	
Generation	Replicates	Independent Cit$^+$ mutants	Replicates	Independent Cit$^+$ mutants	Replicates	Independent Cit$^+$ mutants
Ancestor	6	0	10	0	200	0
5,000	—	—	—	—	200	0
10,000	6	0	30	0	200	0
15,000	—	—	—	—	200	0
20,000	6	0	30	0	200	2
25,000	6	0	30	0	200	0
27,000	—	—	—	—	200	2
27,500	6	0	30	0	—	—
28,000	—	—	—	—	200	0
29,000	6	0	30	0	200	0
30,000	6	0	30	0	200	0
30,500	6	1	30	0	—	—
31,000	6	0	30	0	200	1
31,500	6	1	30	0	200	1
32,000	6	0	30	4	200	2
32,500	6	2	30	1	200	0
Totals	72	4	340	5	2,800	8

Contingency of E. coli evolution experiments

Thank you for your attention!

Web-Page for further information:

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