Können wir Natur und Evolution übertreffen?

Einige Gedanken zur synthetischen Biologie

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Österreich
und
The Santa Fe Institute, Santa Fe, New Mexico, USA

Symposium „Synthetische Biologie“

ÖAW, 14.05.2013

Web-Page für weitere Informationen:

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

"… better than evolution?"

Was heißt: "Besser als die Evolution"?

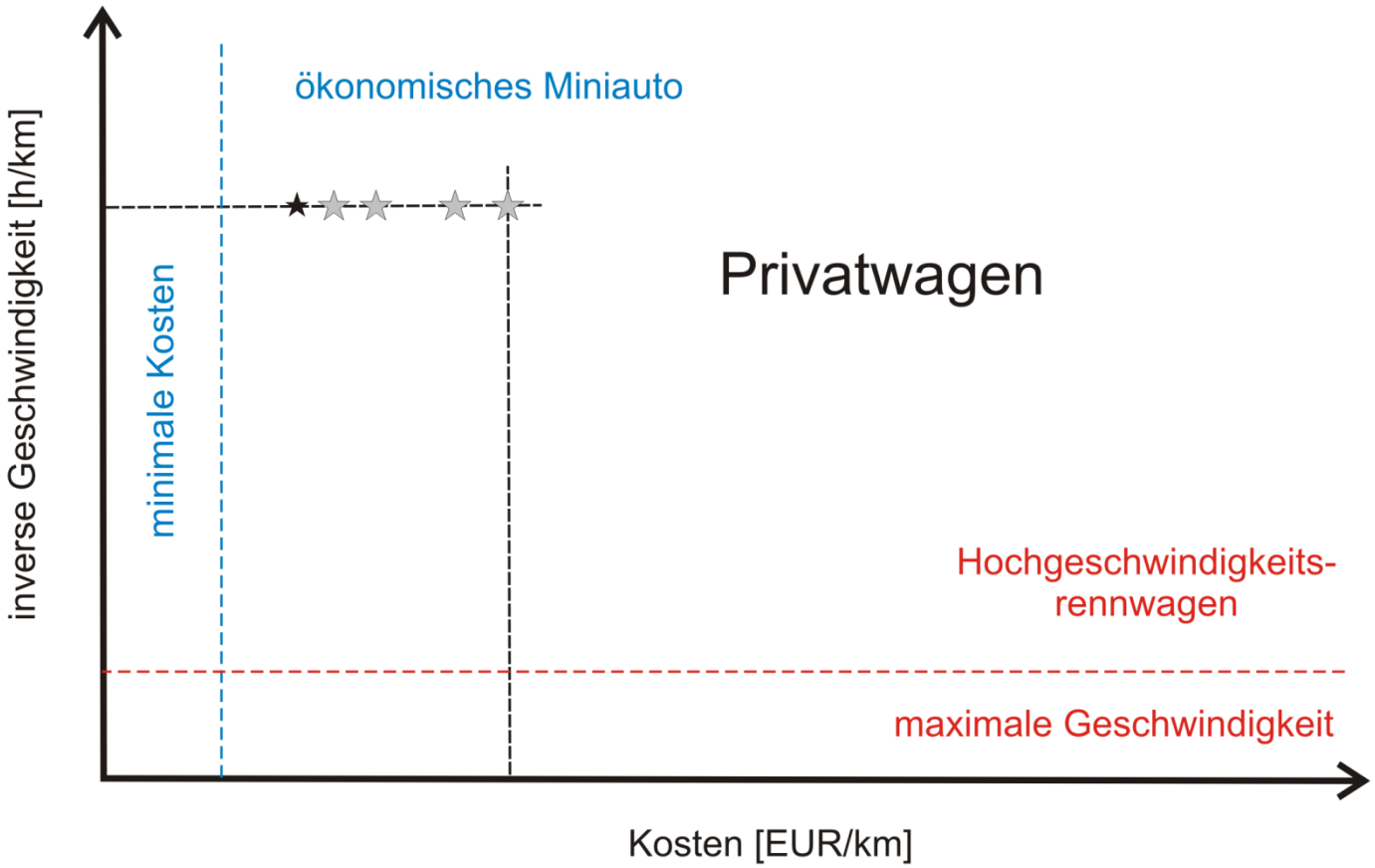
Besser für wen?
Besser wofür?

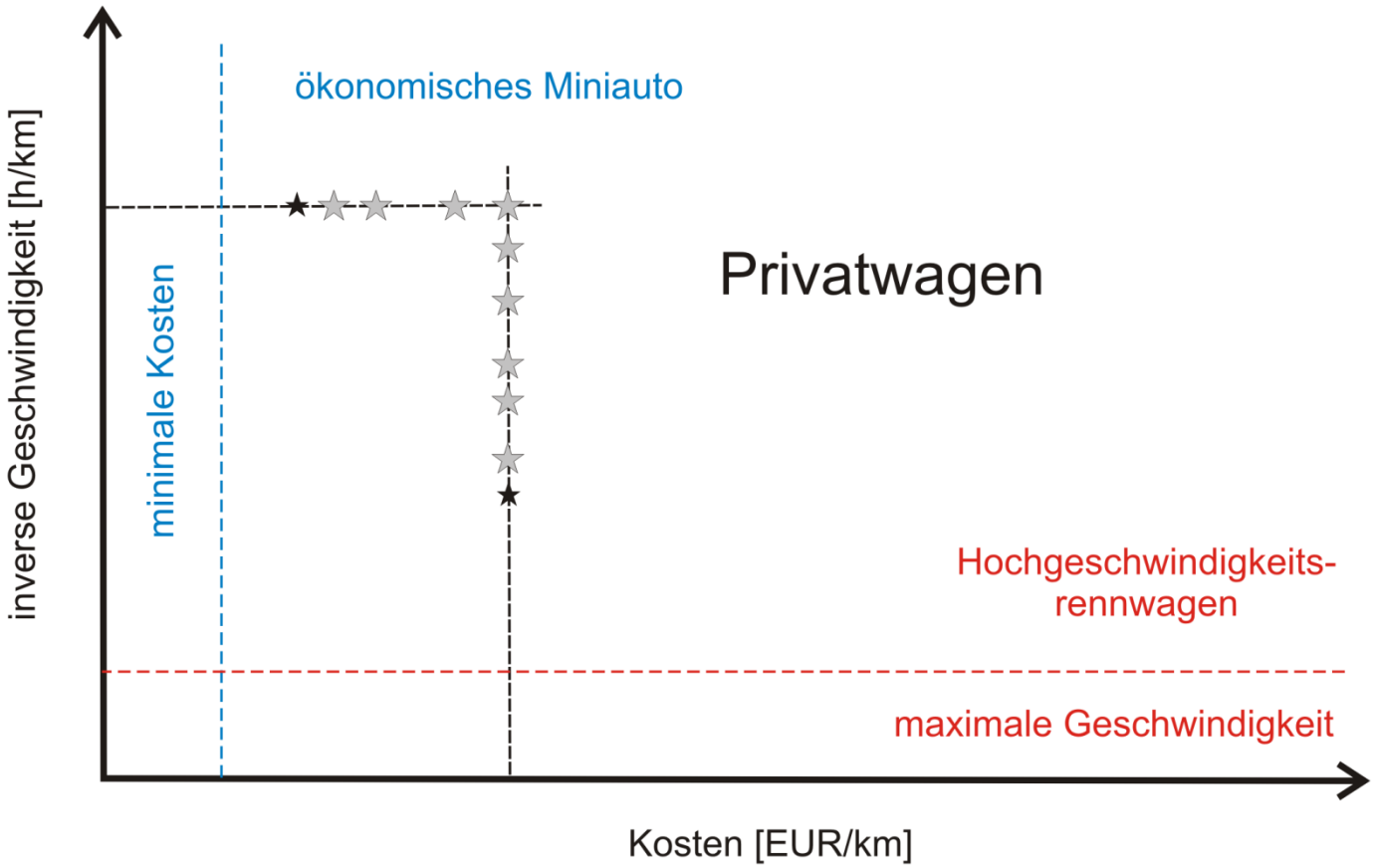
Bezug zu Optimierung?

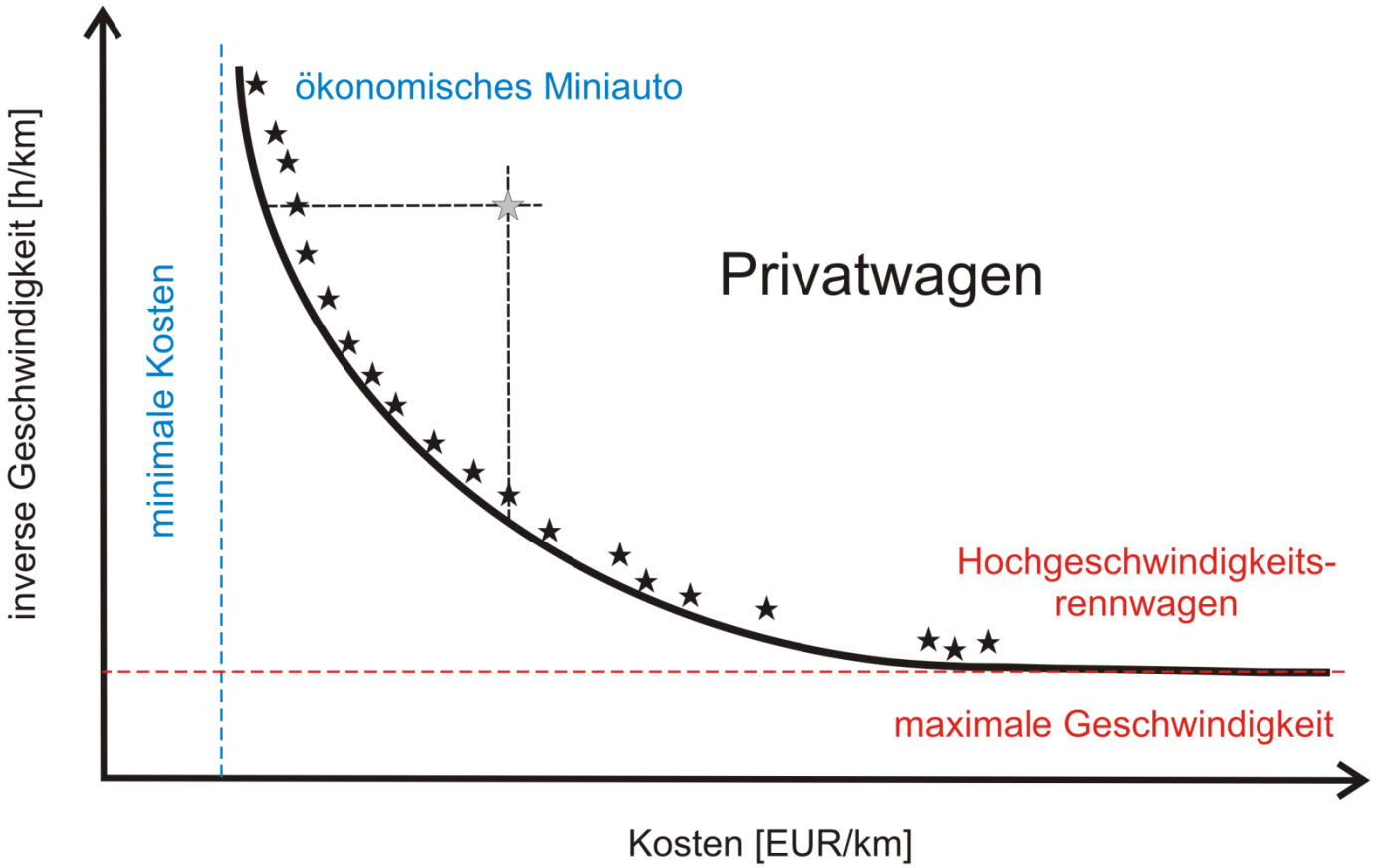
Wie können wir etwas besser machen als die Evolution?

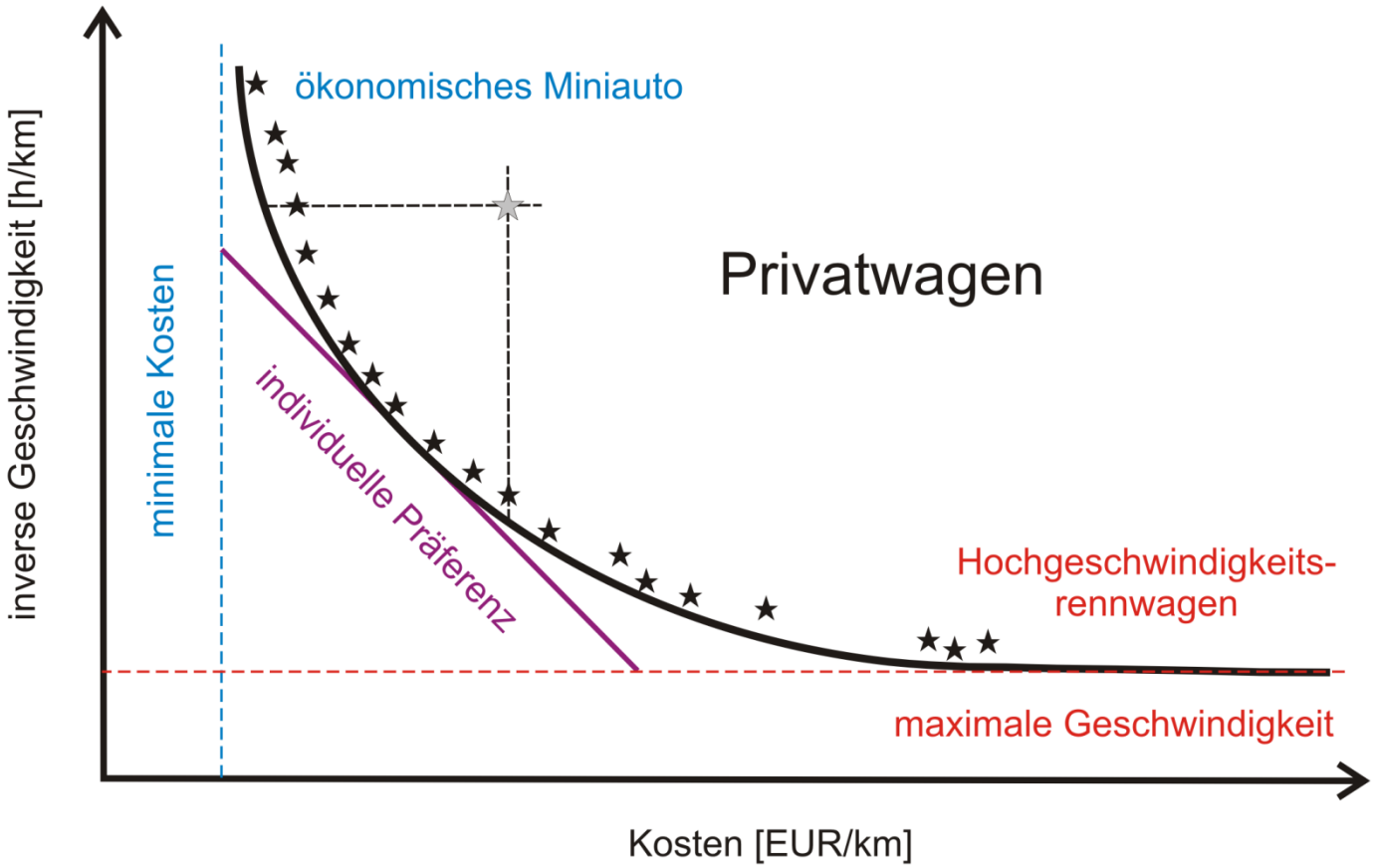
1. Pareto „Gleichgewichte“
2. „Optimalität“ in der Natur
3. Rationales Design
4. Wie können wir Evolution „spielen“?
5. Evolutionäres Design
6. Synthetische Biologie „quo vadis“?

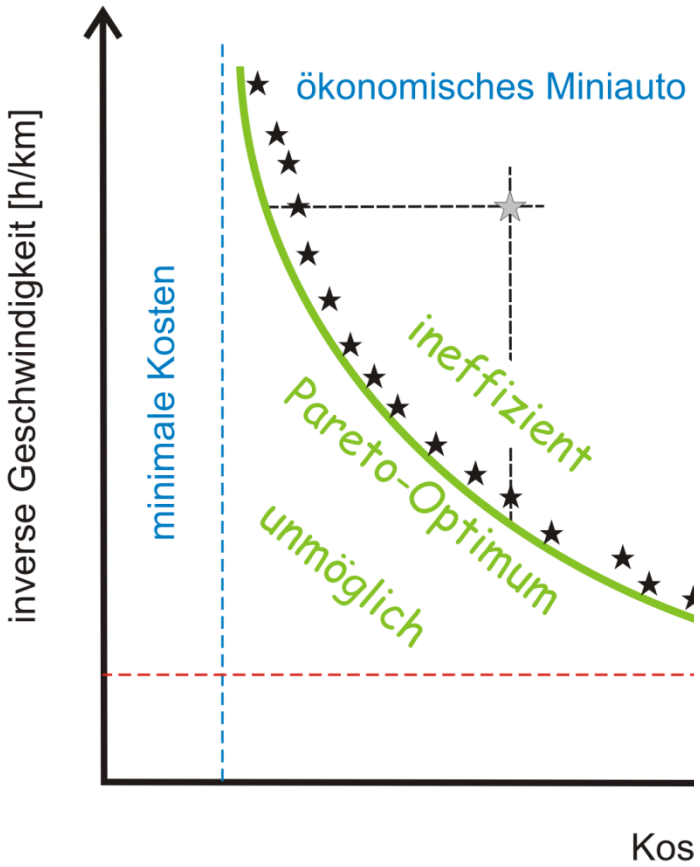
1. **Pareto „Gleichgewichte“**
2. „Optimalität“ in der Natur
3. Rationales Design
4. Wie können wir Evolution „spielen“?
5. Evolutionäres Design
6. Synthetische Biologie „quo vadis“?

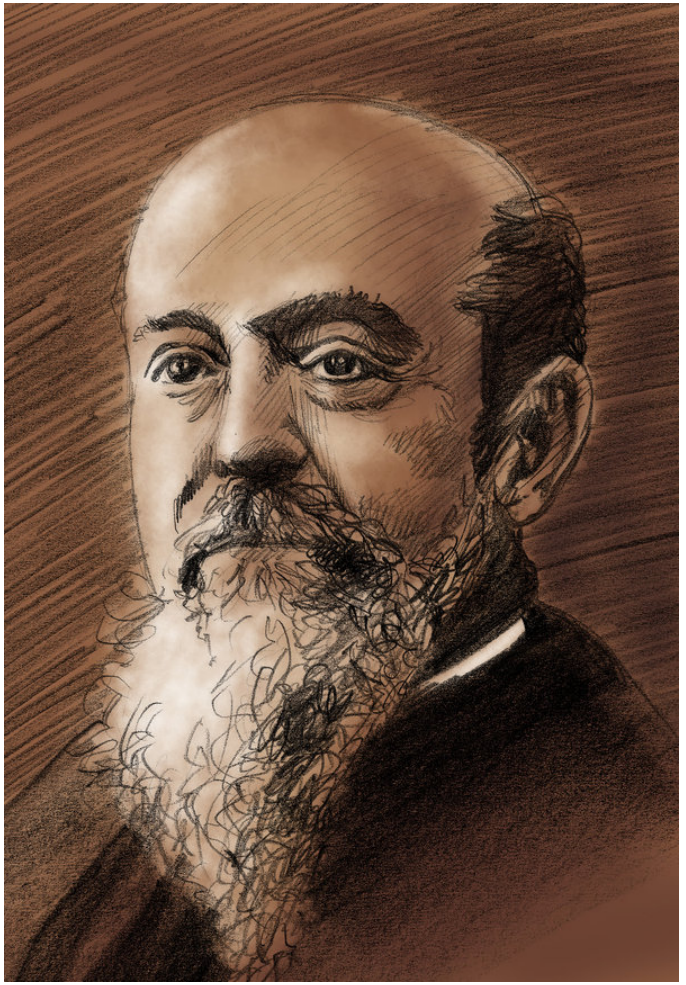

ökonomisches Miniauto
minimale Kosten
Privatwagen
Hochgeschwindigkeits-
rennwagen
maximale Geschwindigkeit
inverse Geschwindigkeit [h/km]
Kosten [EUR/km]


ökonomisches Miniauto
minimale Kosten
Privatwagen
Hochgeschwindigkeits-
rennwagen
maximale Geschwindigkeit
inverse Geschwindigkeit [h/km]
Kosten [EUR/km]


inverse Geschwindigkeit [h/km]
minimale Kosten
ökonomisches Miniauto
Privatwagen
Hochgeschwindigkeits-
rennwagen
maximale Geschwindigkeit
Kosten [EUR/km]

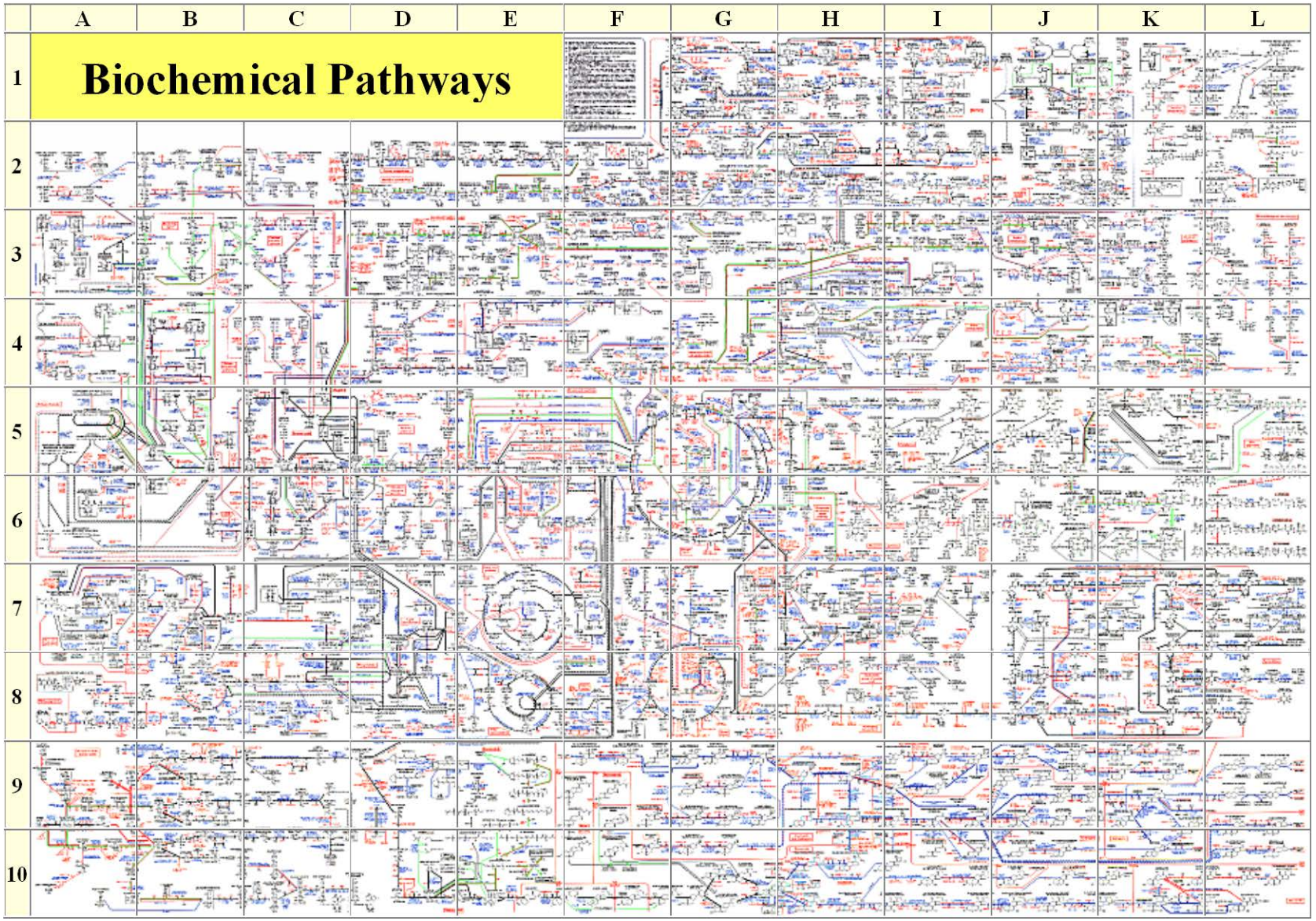

ökonomisches Miniauto
Privatwagen
minimale Kosten
individuelle Präferenz
Hochgeschwindigkeits-
rennwagen
maximale Geschwindigkeit
inverse Geschwindigkeit [h/km]
Kosten [EUR/km]



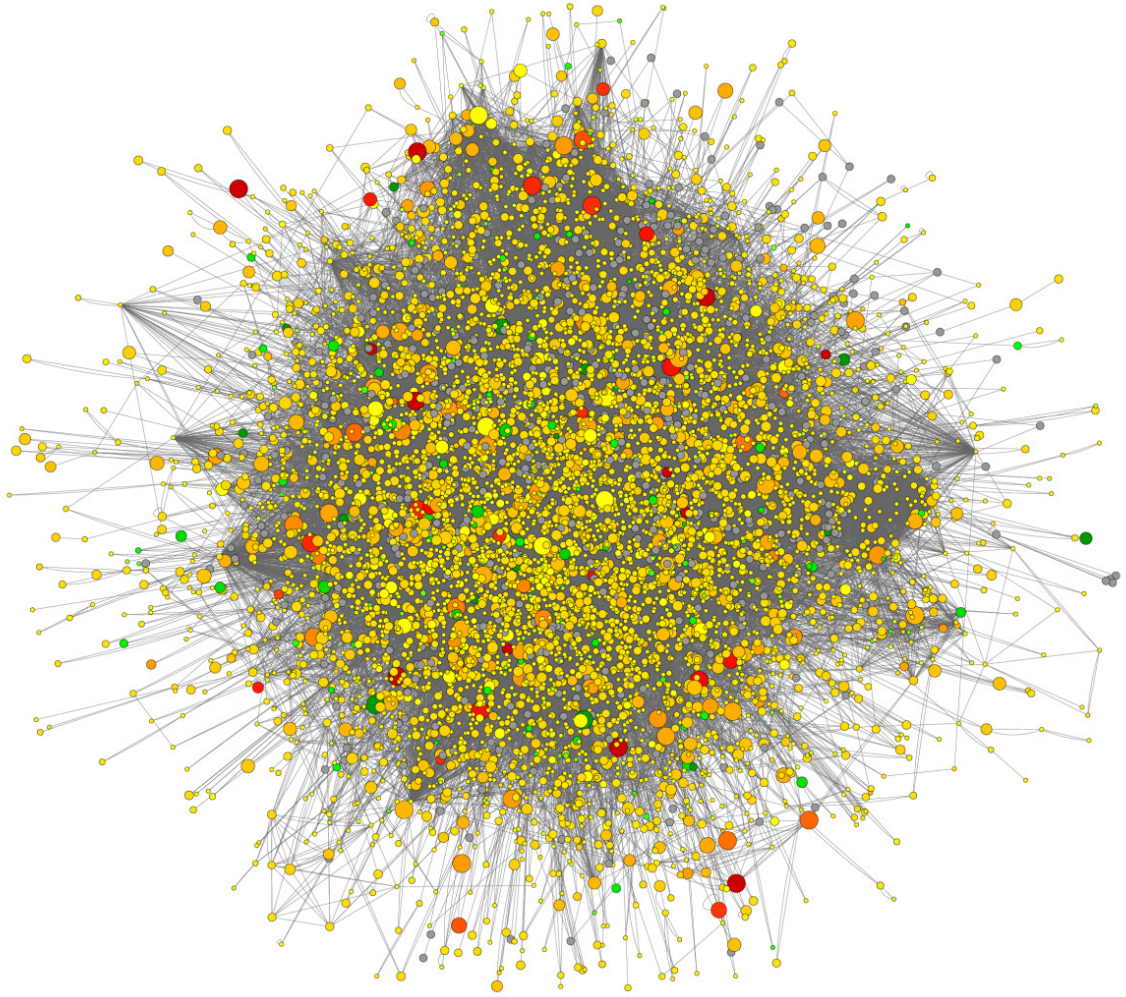


Vilfredo Frederico Pareto, 1848 - 1923

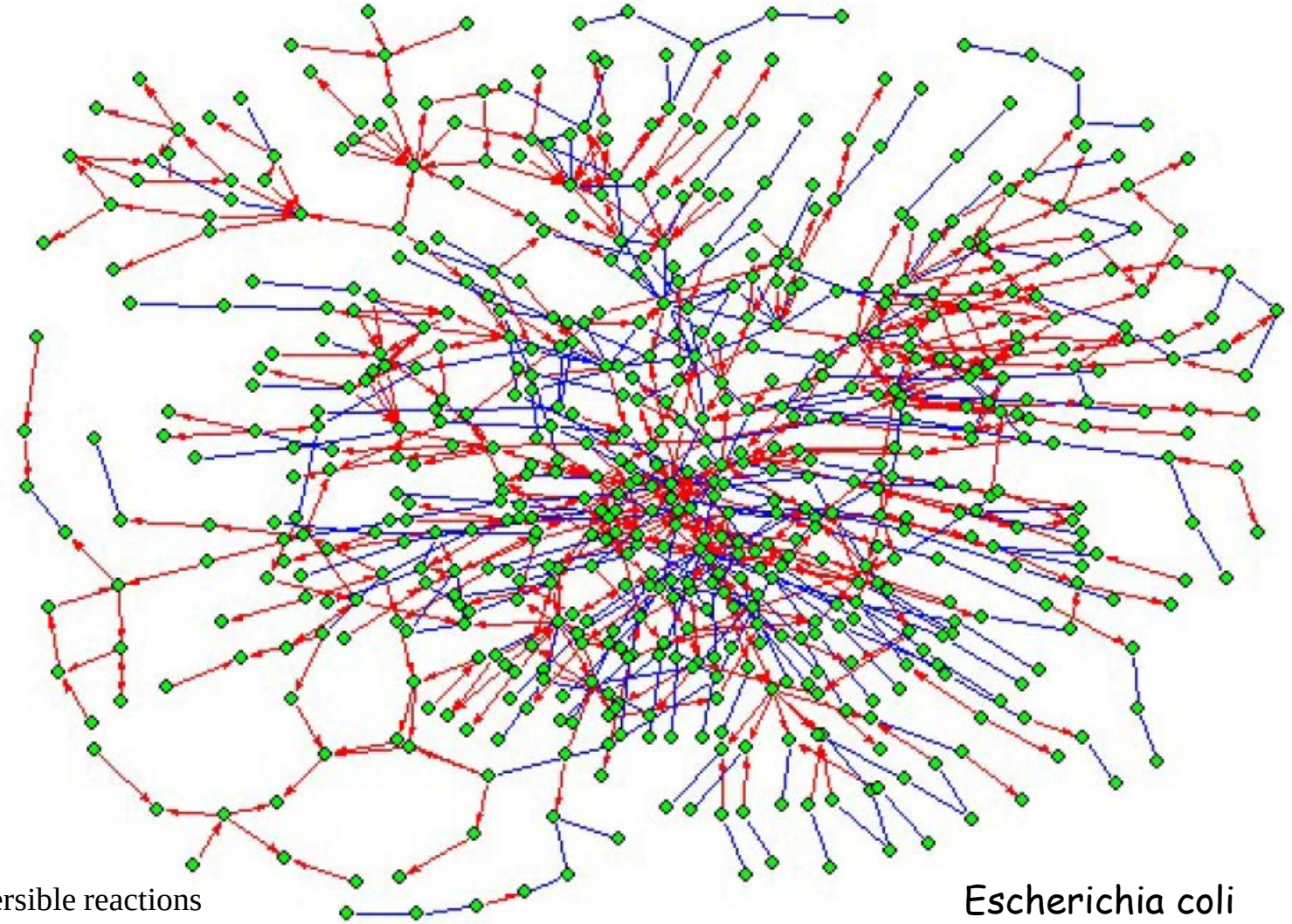
1. Pareto „Gleichgewichte“
2. **„Optimalität“ in der Natur**
3. Rationales Design
4. Wie können wir Evolution „spielen“?
5. Evolutionäres Design
6. Synthetische Biologie „quo vadis“?



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



Christopher R. Bauer, Andrew M. Epstein, Sarah J. Sweeney, Daniela C. Zarnescu, and Giovanni Bosco. Genetic and Systems level analysis of *Drosophila sticky/citron kinase* and *dFmrl* mutants reveal common regulation of genetic networks. *BMC Systems Biology* **2**:e101 (2008).



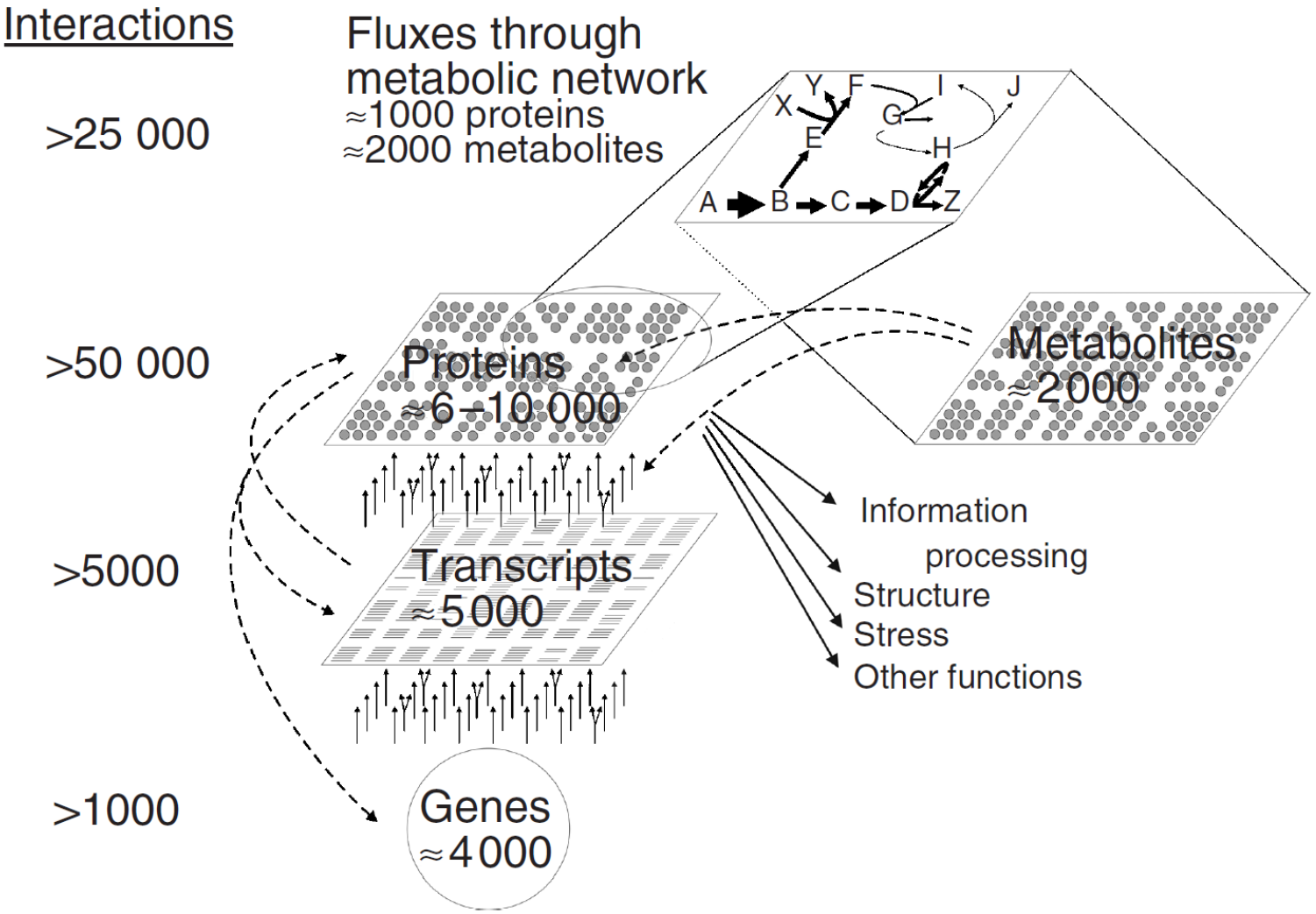
reversible reactions
irreversible reactions

Escherichia coli

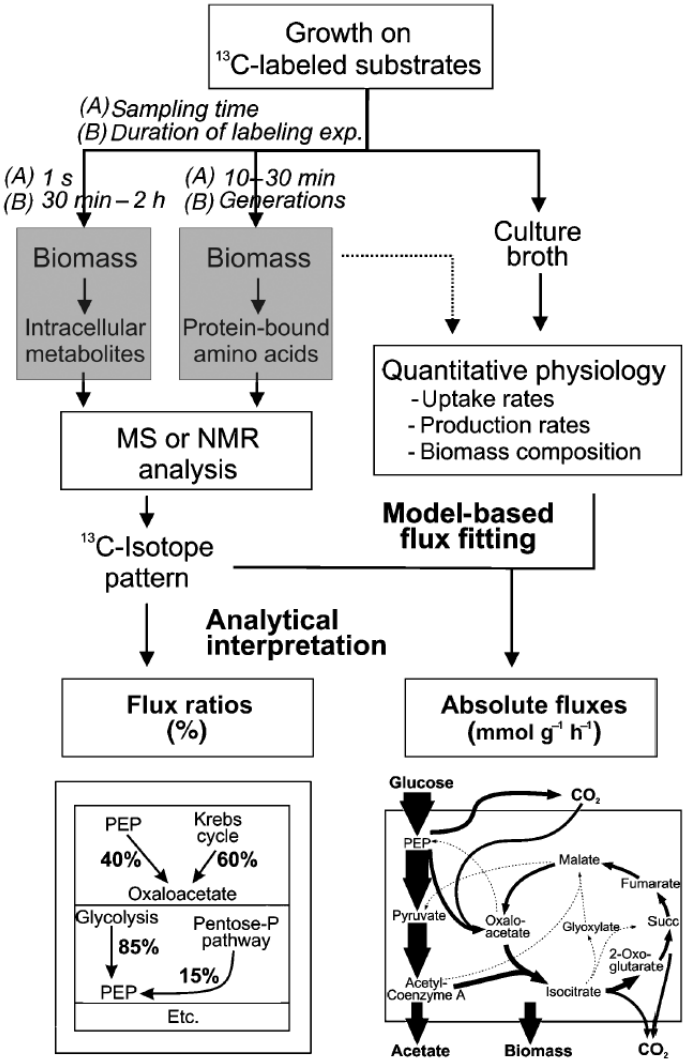
Hongwu Ma, An-Ping Zeng. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* **18**:270-277 (2003).

Pathway	Number	Gene	Reaction
Glycolysis	1	*glk*	GLC + ATP → G6P
	2	*pgi*	G6P ⇌ F6P
	3	*pfkA* / *pfkB*	F6P + ATP → F1,6DP
	4	*fbaA* / *fbaB*	F1,6DP ⇌ DHAP + GA3P
	5	*tpiA*	DHAP ⇌ GA3P
	6	*gapA*	GA3P ⇌ NADH + 13DPG
	7	*pgk*	13DPG ⇌ ATP + 3PG
	8	*gpmA* / *gpmB*	3PG ⇌ 2PG
	9	*eno*	2PG ⇌ PEP
	10	*pykA* / *pykF*	PEP → PYR + ATP
	11	*aceEF,lpd*	PYR + CoA → AcCoA + CO2 + NADH

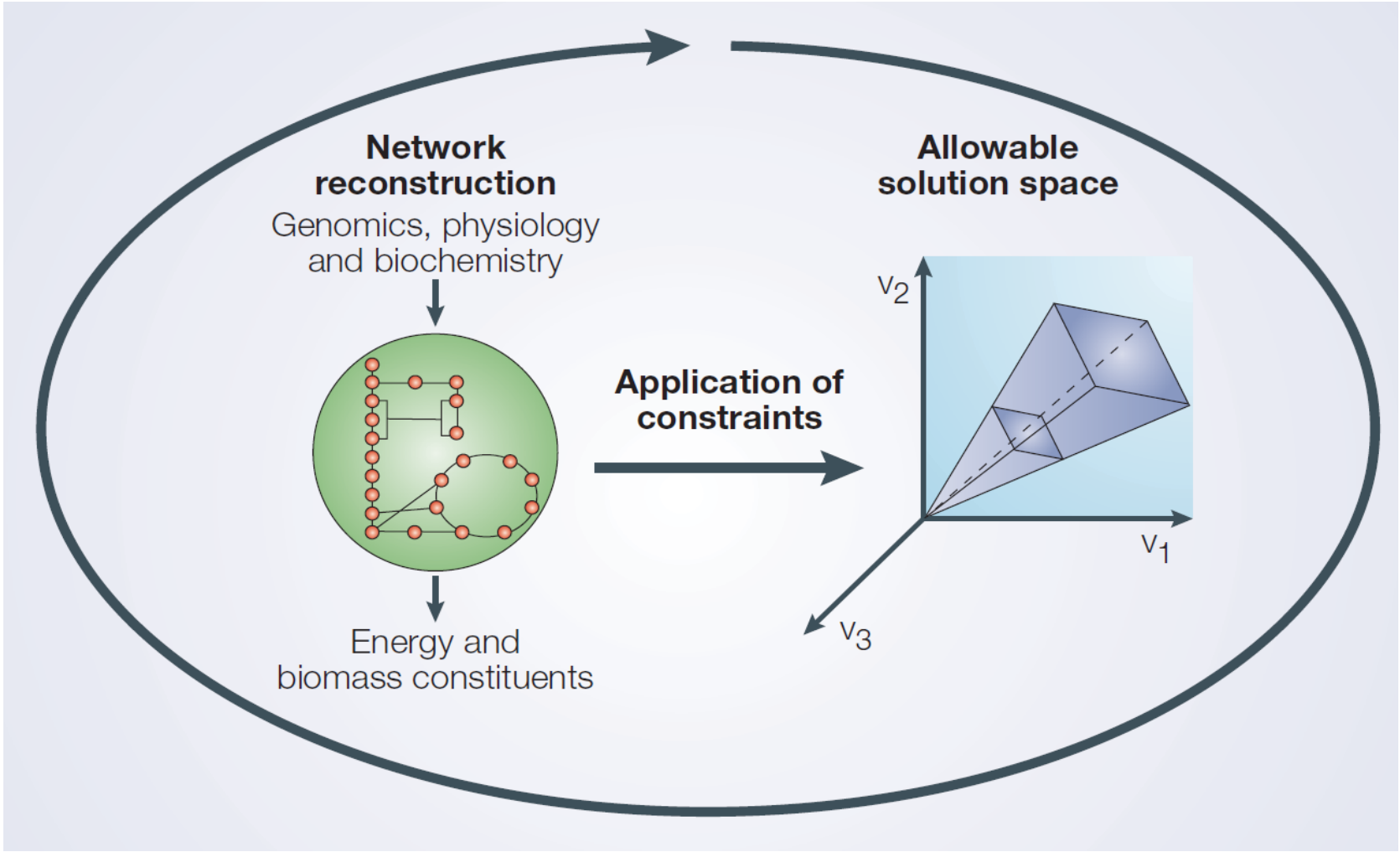
Robert Schuetz, Nicola Zamboni, Mattia Zampieri, Matthias Heinemann, Uwe Sauer. Multidimensional optimality of microbial metabolism. *Science* **336**:601-604 (2012)



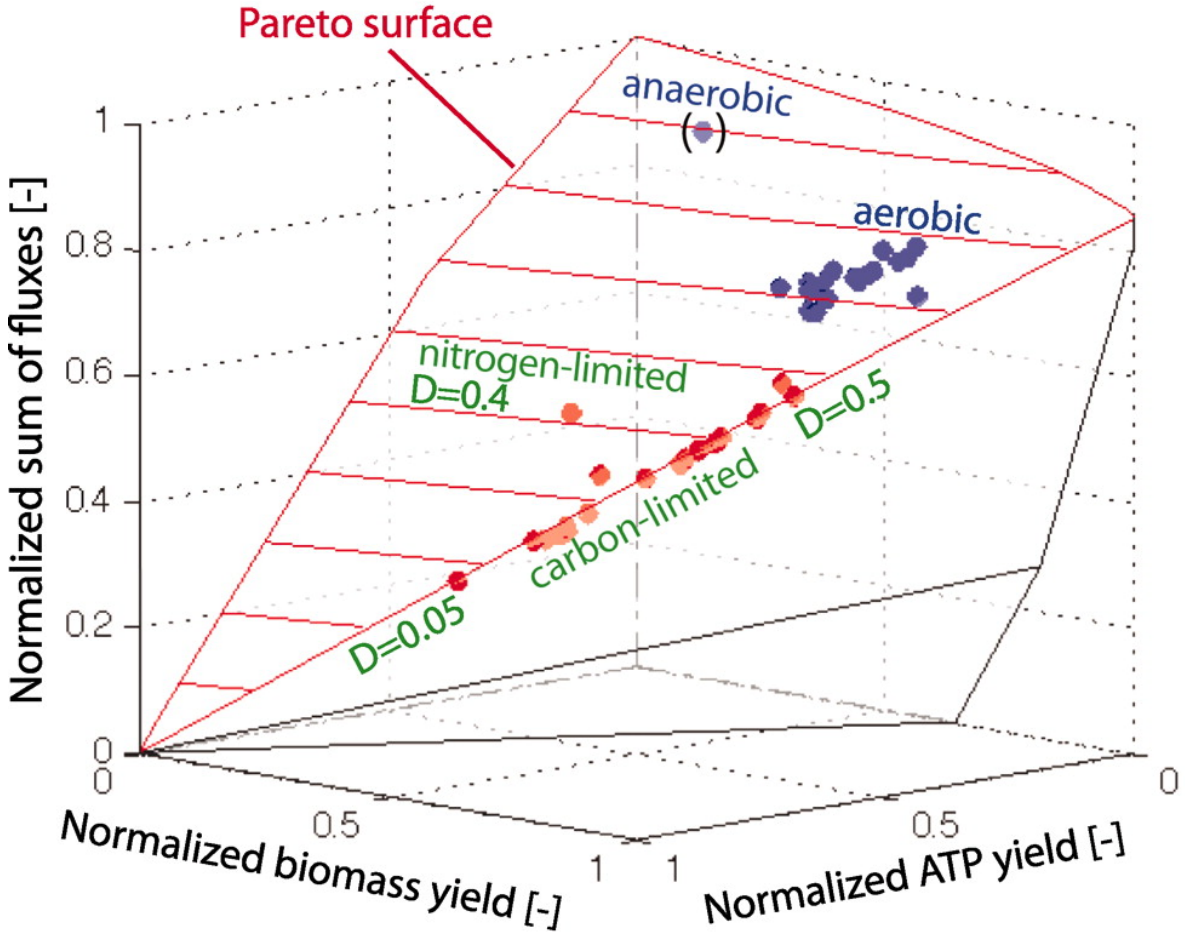
Uwe Sauer. Metabolic networks in motion: ^{13}C-based flux analysis. *Molecular Systems Biology* **2**:e62 (2006)



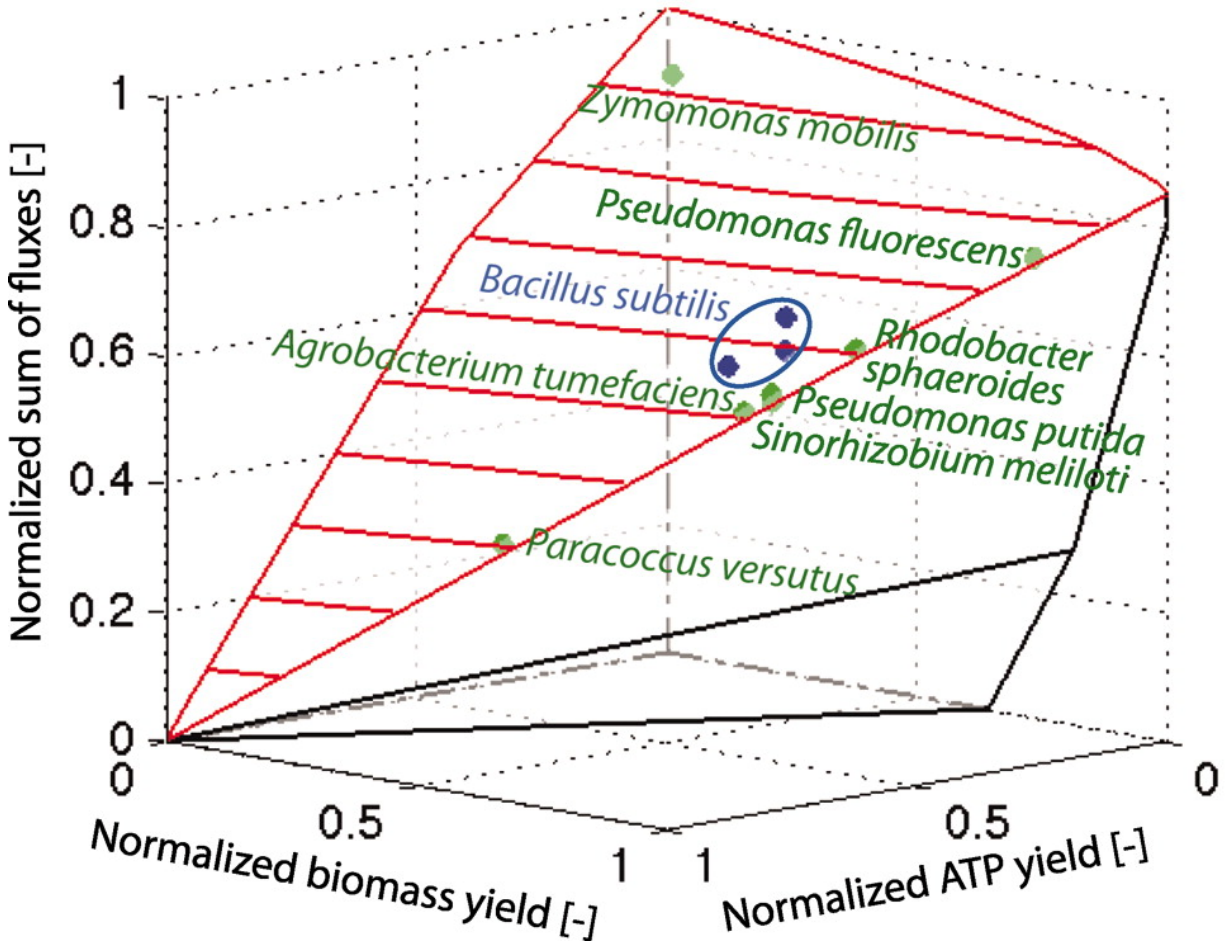
Uwe Sauer. Metabolic networks in motion: ^{13}C-based flux analysis.
Molecular Systems Biology **2**:e62 (2006)



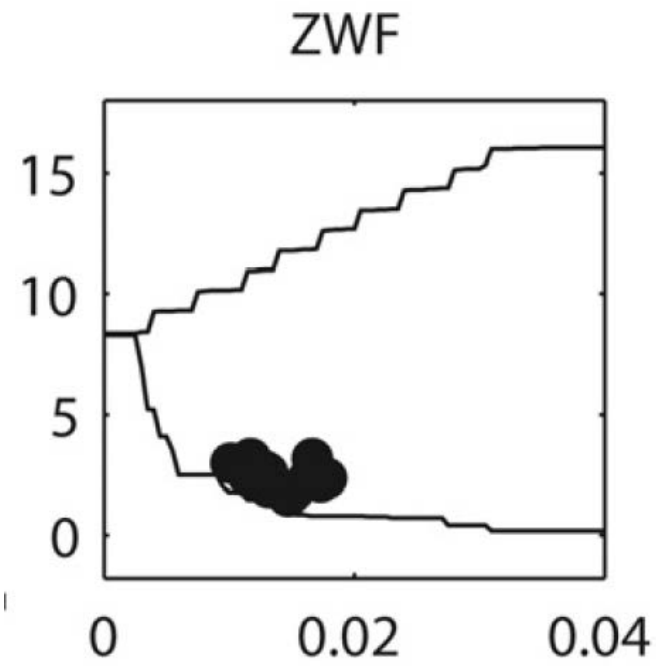
Nathan D. Price, Jennifer L. Reed, and Bernhard Ø. Palsson. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nature Reviews Microbiology* **2**:886-897 (2004)

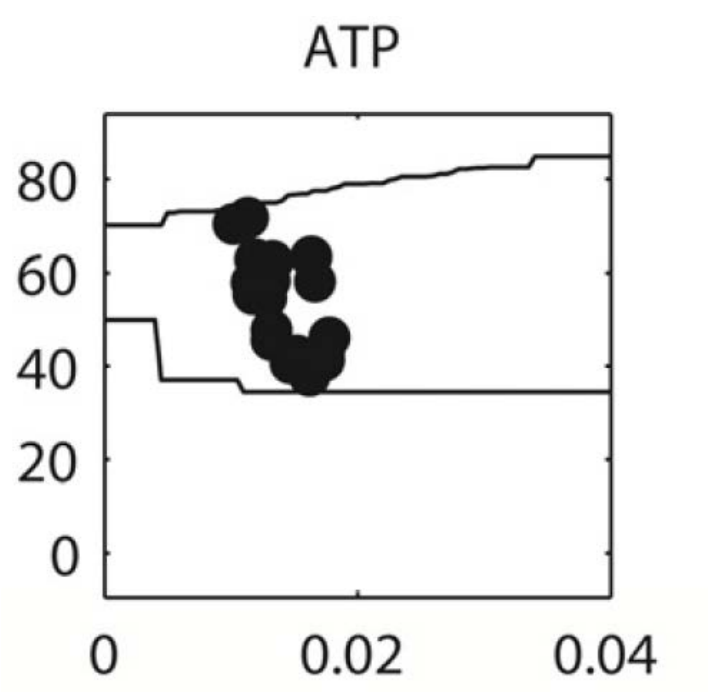


Robert Schuetz, Nicola Zamboni, Mattia Zampieri, Matthias Heinemann, Uwe Sauer. Multidimensional optimality of microbial metabolism. *Science* **336**:601-604 (2012)



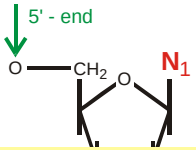
Robert Schuetz, Nicola Zamboni, Mattia Zampieri, Matthias Heinemann, Uwe Sauer. Multidimensional optimality of microbial metabolism. *Science* **336**:601-604 (2012)



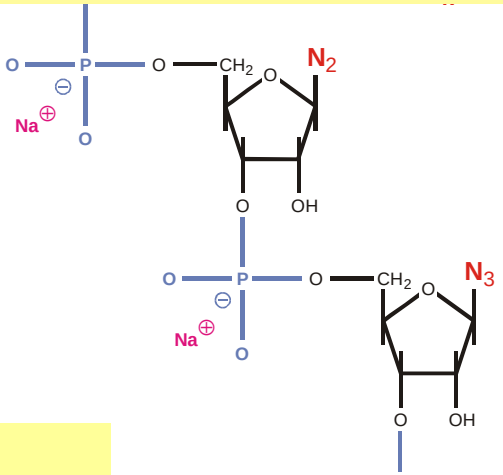


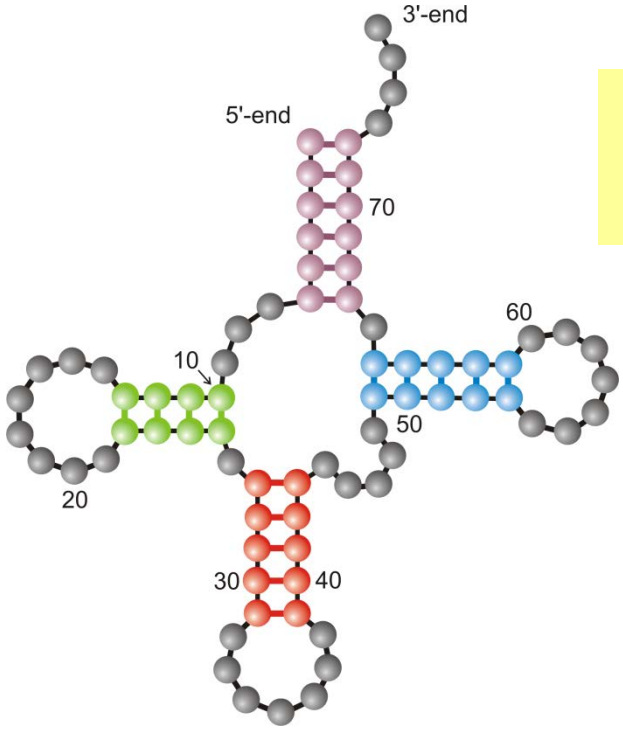
Robert Schuetz, Nicola Zamboni, Mattia Zampieri, Matthias Heinemann, Uwe Sauer. Multidimensional optimality of microbial metabolism. *Science* **336**:601-604 (2012)

1. Pareto „Gleichgewichte“
2. „Optimalität“ in der Natur
3. **Rationales Design**
4. Wie können wir Evolution „spielen“?
5. Evolutionäres Design
6. Synthetische Biologie „quo vadis“?



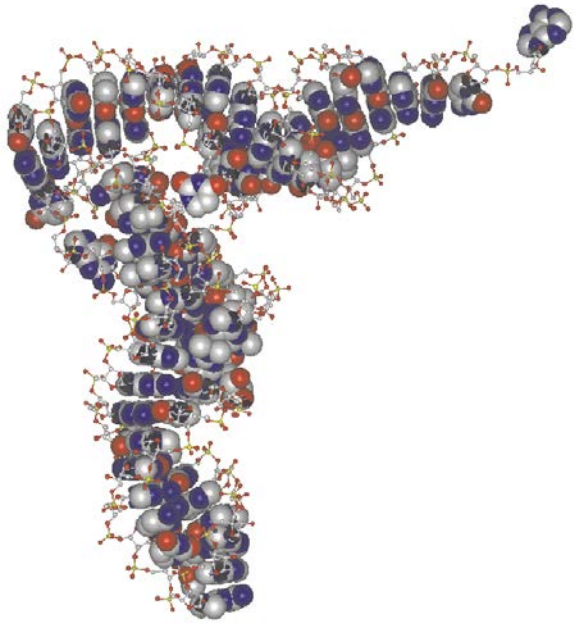
5'-end GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'-end





RNA structure

The molecular phenotype



RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

linear programming

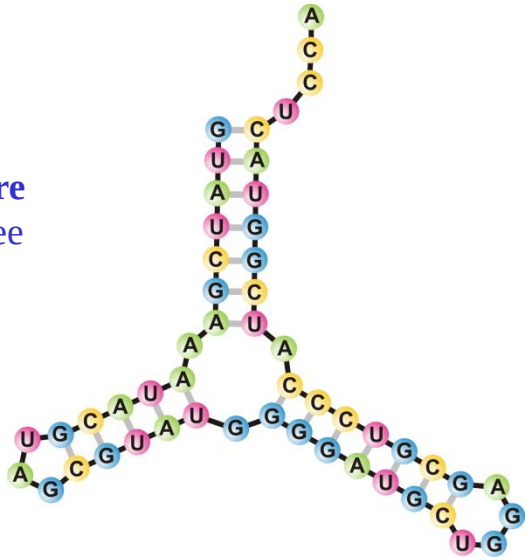
RNA folding:

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

biophysical chemistry: thermodynamics and kinetics

empirical parameters

RNA structure of minimal free energy



From RNA sequence to structure

RNA sequence GUAUCGAAAUACGUAGCGUAUGGGGAUGCUGGACGGUCCCAUCGGUACUCCA

Linear programming

RNA folding:

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

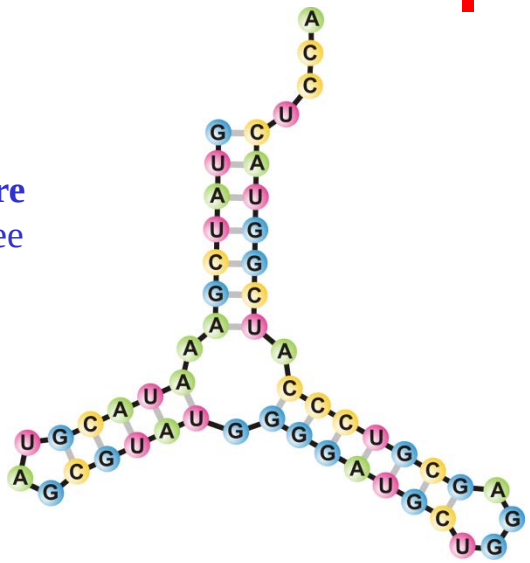
iterative determination of a sequence for the given secondary structure

inverse Folding Algorithm

inverse folding of RNA:

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

RNA structure of minimal free energy



From RNA structure to sequence

RNAinverse software: I. L.Hofacker et al., **1994**

RNA-SSD software: M. Andronescu, AP. Fejes, F. Hutter, HH. Hoos and A. Condon. A new algorithm for RNA secondary structure design. *J Mol Biol.* 336: 607-624, **2004**

InfoRNA software: A. Busch and R. Backofen. INFO-RNA -Fast approach to inverse RNA folding. *Bioinformatics* 22 15:1823-1831, **2006**

Modena software: A. Taneda. MODENA: A multi-objective RNA inverse folding. *Advances and Applications in Bioinformatics and Chemistry* 4:1-12, **2011**

NUPACK software: J.N. Zadeh, B.R. Wolfe, N.A. Pierce. Nucleic acid sequence design via efficient ensemble defect optimization. *J Comput Chem*, 32, 439–452, **2011**

The *Vienna RNA-Package*:

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

Citations Web of Science 13.05.2013: **1006**

Monatshefte für Chemie 125, 167–188 (1994)

Monatshefte für Chemie
Chemical Monthly

Printed in Austria

Fast Folding and Comparison of RNA Secondary Structures

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

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

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

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

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

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

Schnelle Faltung und Vergleich von Sekundärstrukturen von RNA

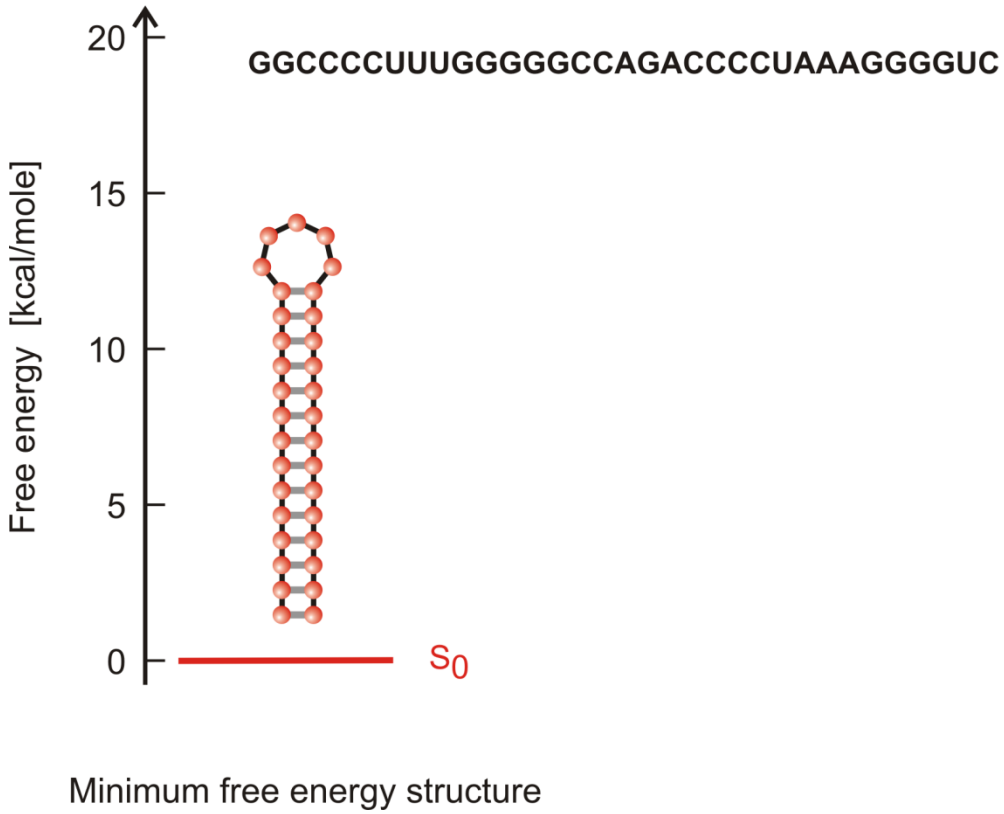
Zusammenfassung. Die im Vienna RNA package enthaltenen Computer Programme für die Berechnung und den Vergleich von RNA Sekundärstrukturen werden präsentiert. Ihren Kern bilden Algorithmen zur Vorhersage von Strukturen minimaler Energie sowie zur Berechnung von Zustandssumme und Basenpaarungswahrscheinlichkeiten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

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

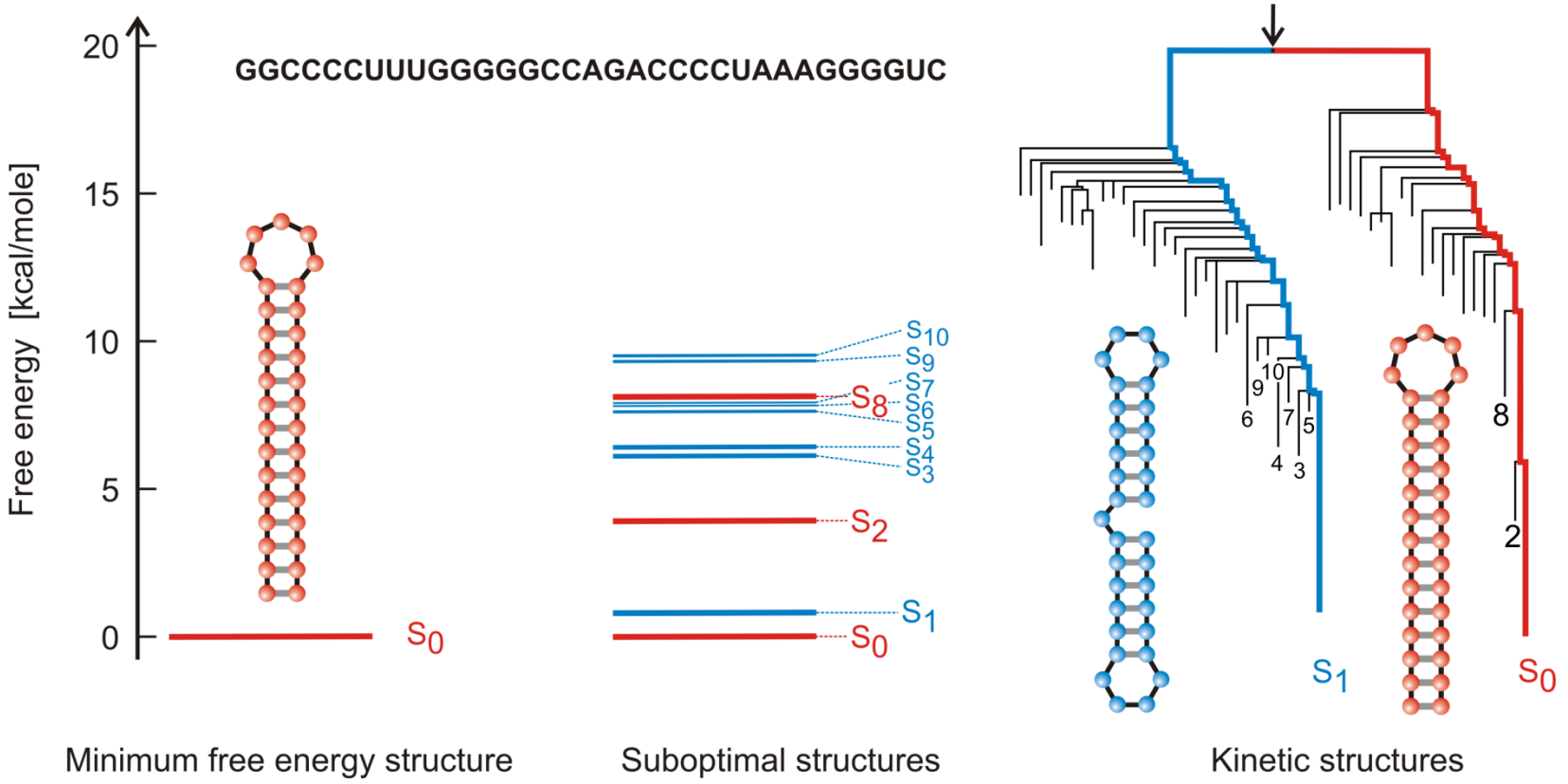
Alle Programme sind in ANSI C geschrieben, darunter auch eine Implementation des Faltungsalgorithmus für Parallelrechner mit verteiltem Speicher. Wie Tests auf einem Intel Hypercube zeigen, wird das Parallelrechnen umso effizienter je länger die Sequenzen sind.

1. Introduction

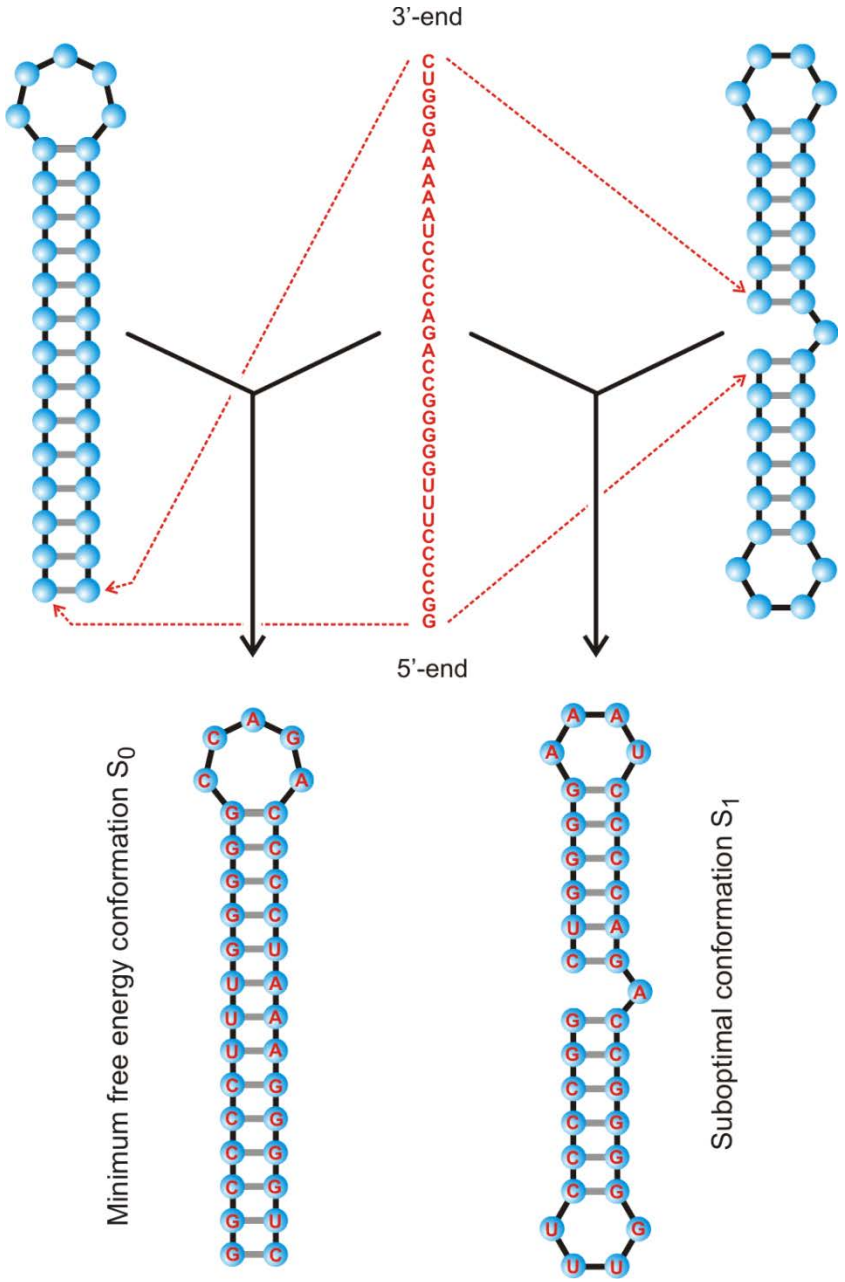
Recent interest in RNA structures and functions was caused by their catalytic capacities [1, 2] as well as by the success of selection methods in producing RNA



The notion of structure

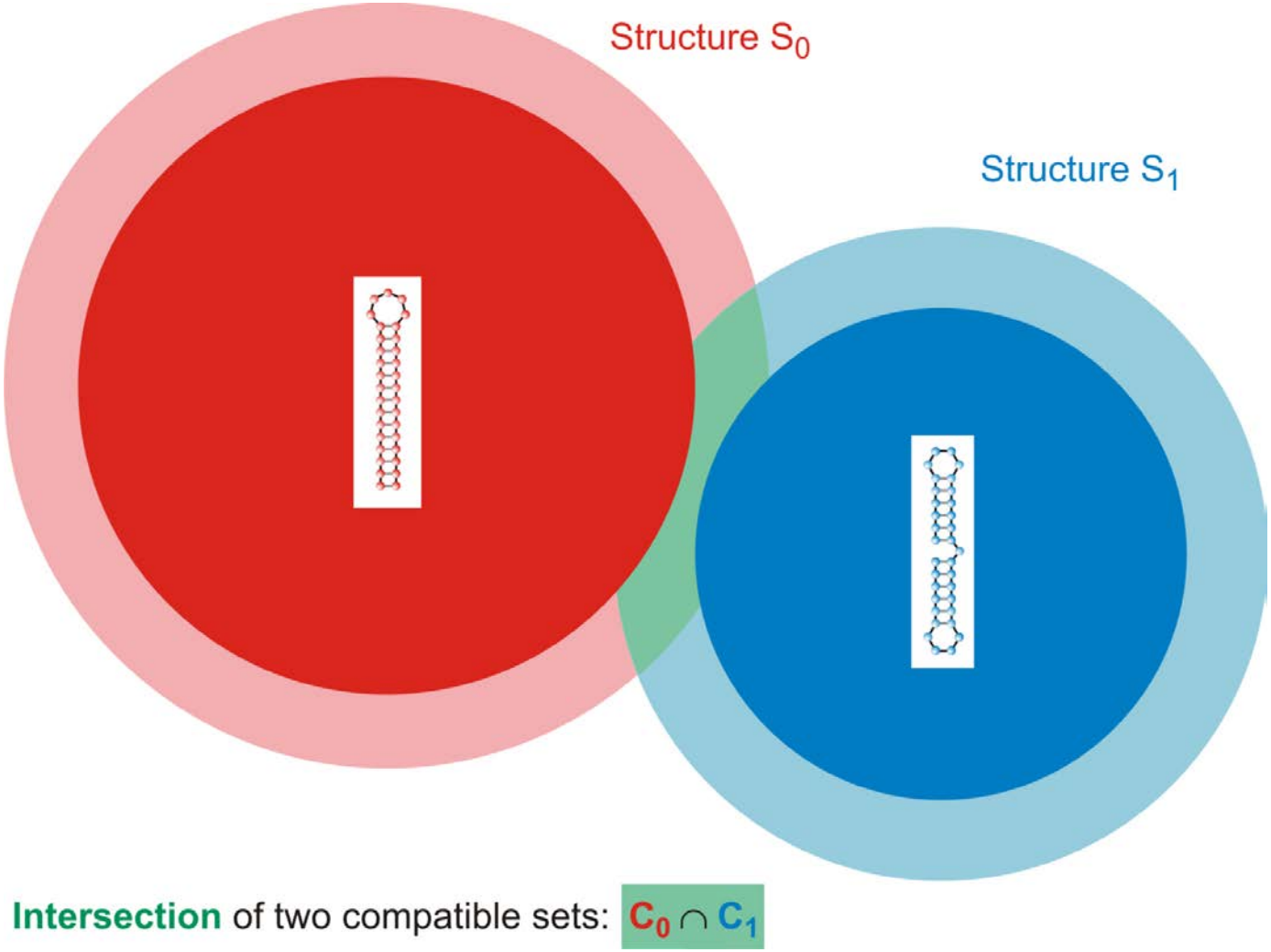


Interconversion of suboptimal structures



One sequence fits on two structures.

Can we find out whether this is a special case or a common property of structures ?



P. Schuster. Prediction of RNA secondary structures: From theory to models and real molecules. Rep.Prog.Phys. 69:1419-1477, 2006

C. Reidys, P.F. Stadler, P.Schuster. Generic properties of combinatory maps. Neutral networks of RNA secondary structure, Bull.Math.Biol. 59:339-397, 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*).

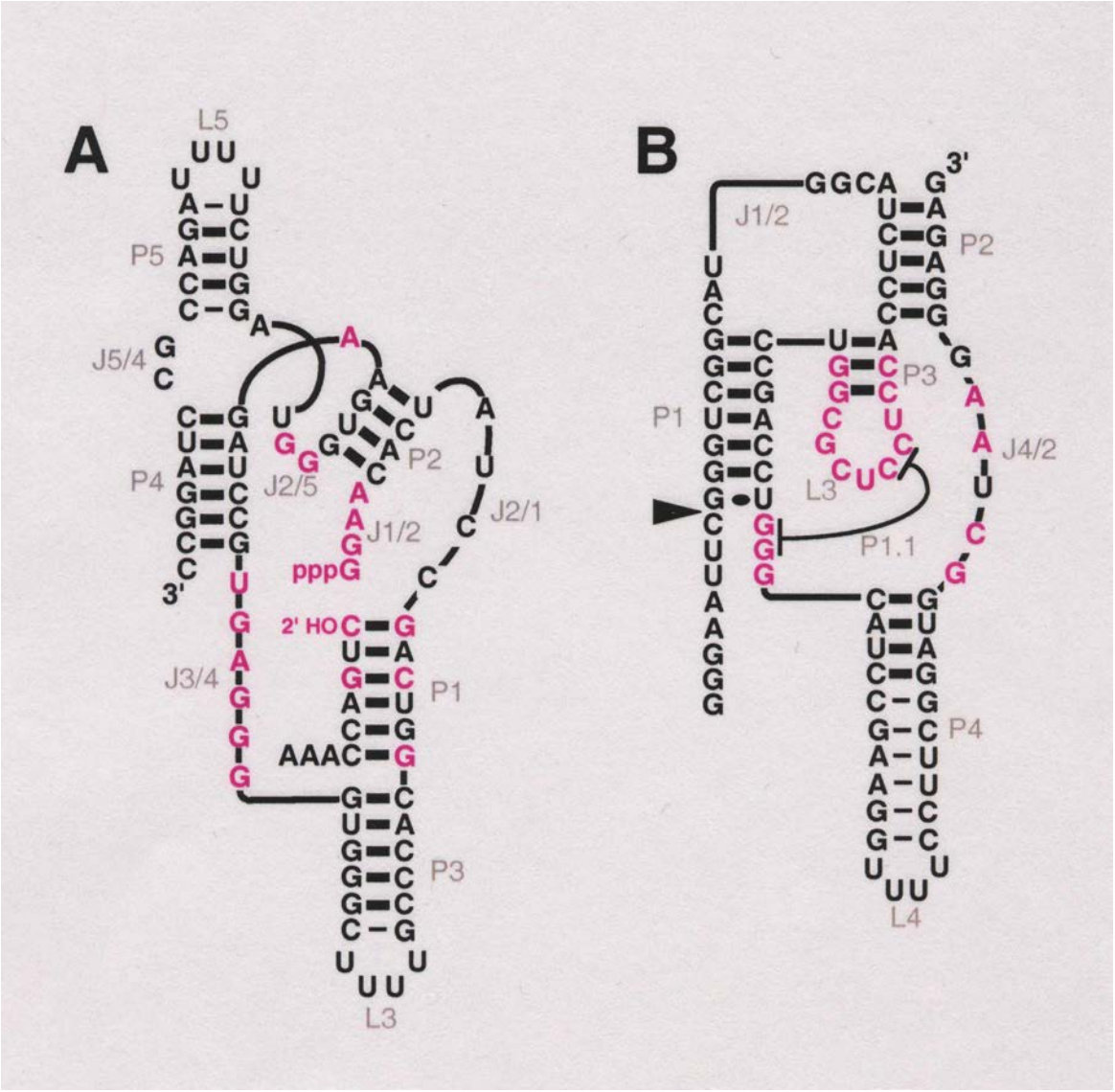
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect.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

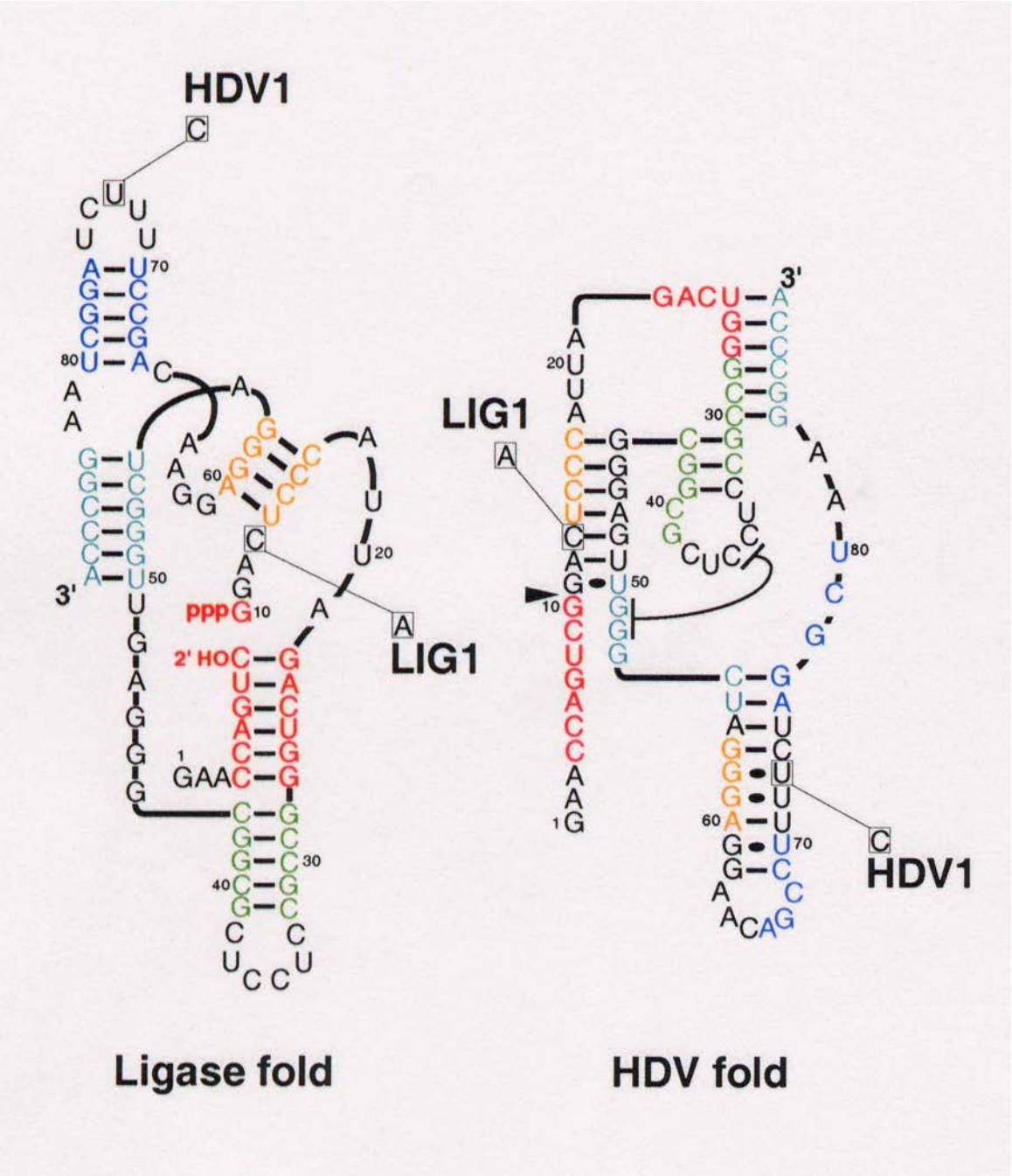
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

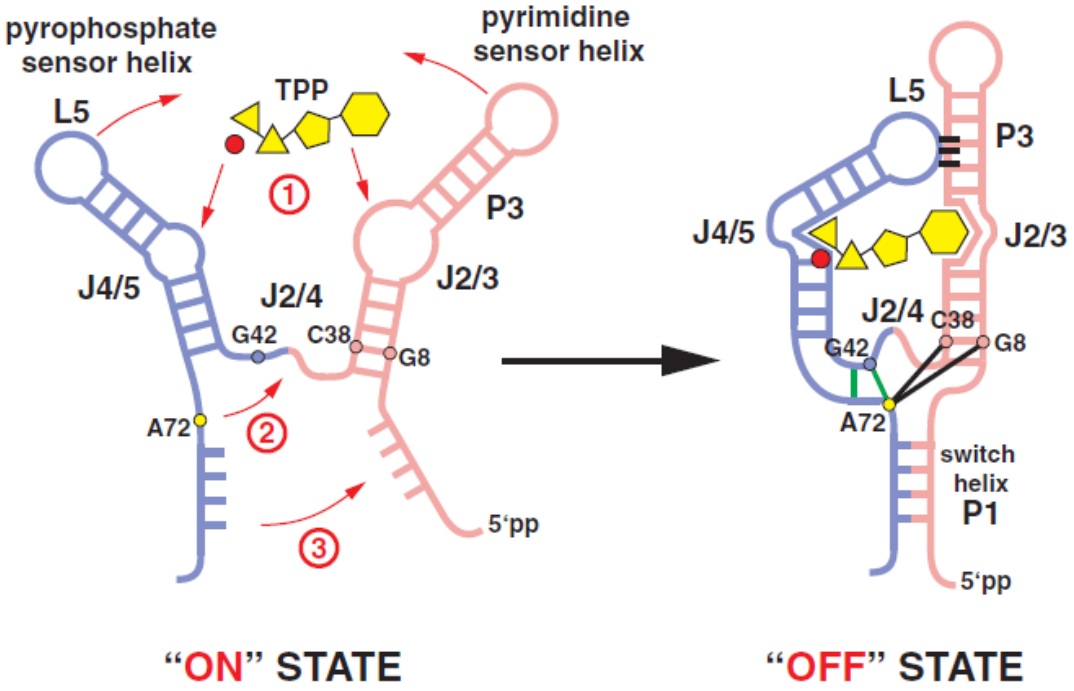


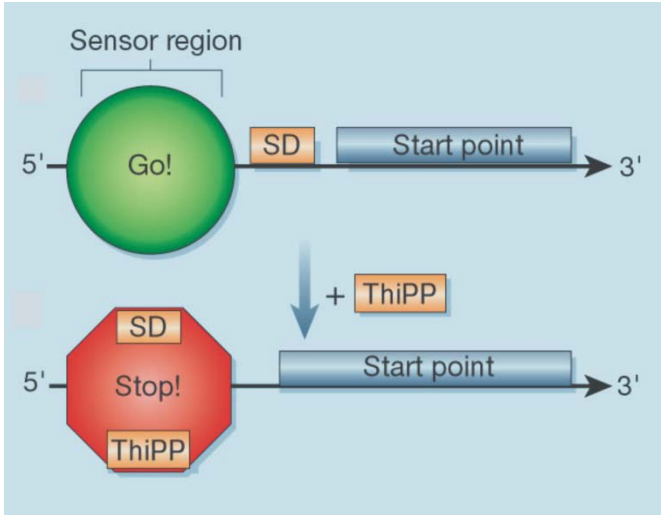
Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

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

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

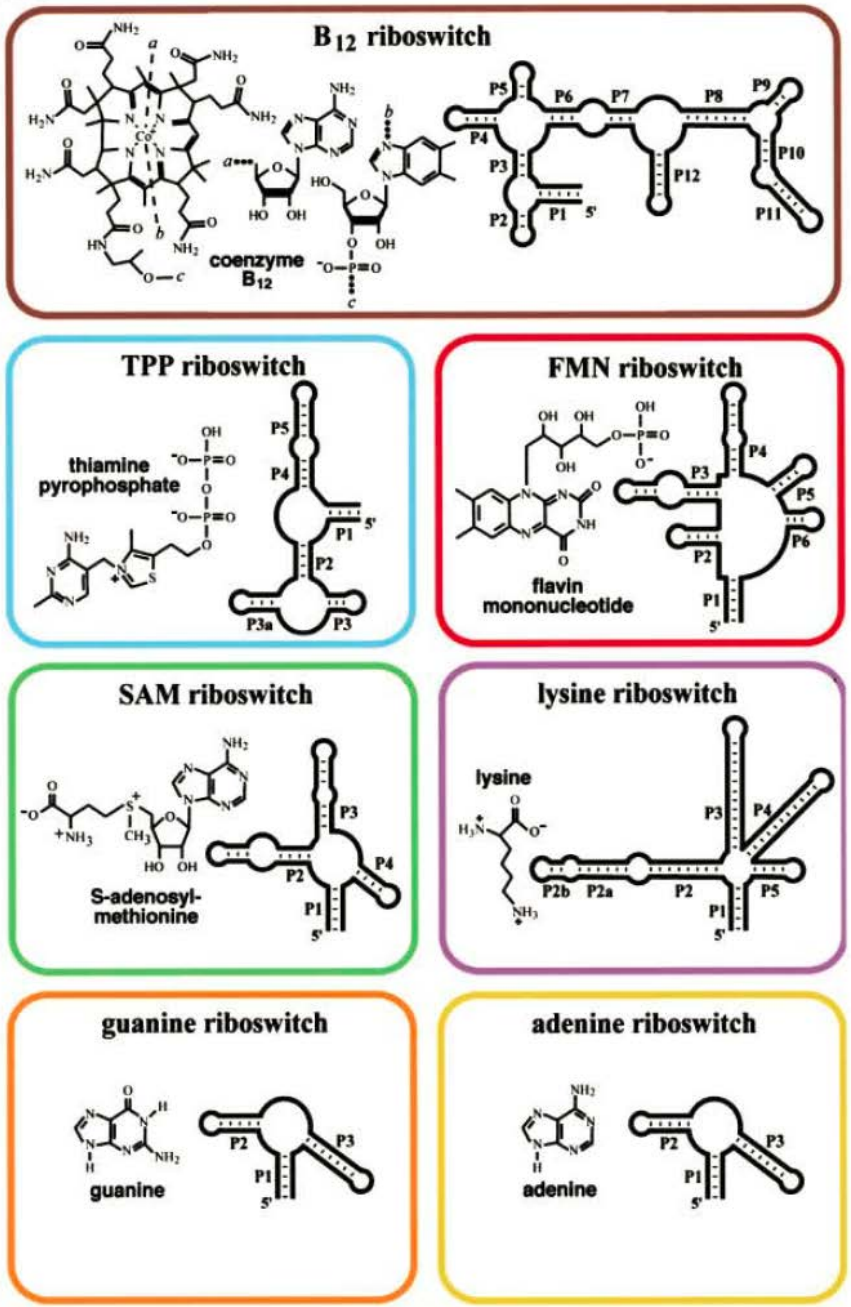


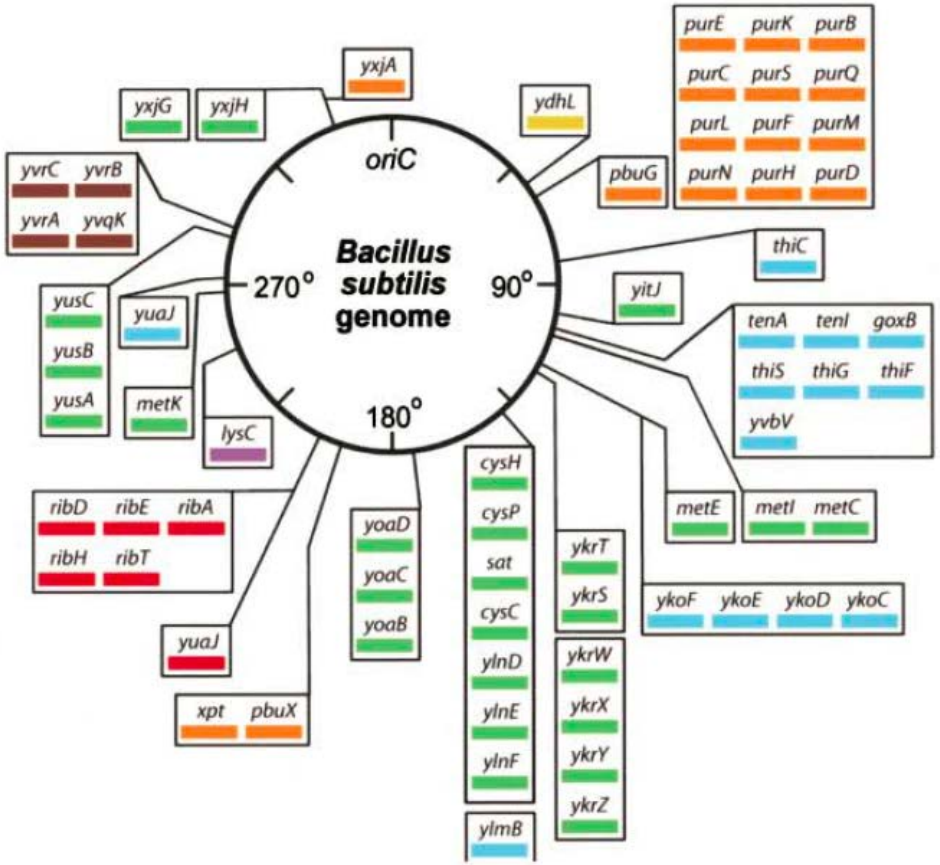




The thiamine-pyrophosphate riboswitch

S. Thore, M. Leibundgut, N. Ban.
Science **312**:1208-1211, 2006.





M. Mandal, B. Boese, J.E. Barrick, W.C. Winkler, R.R, Breaker. Cell 113:577-586 (2003)

Table 1. Four classes of enzymes are generally used in detergents.

Proteases	Most widely used enzymes in the detergent industry remove protein stains such as grass, blood, egg and human sweat which have a tendency to adhere strongly to textile fibers.
Amylases	Used to remove residues of starch-based foods like potatoes, spaghetti, custards, gravies and chocolate.
Lipases	Decompose fatty material. Lipase is capable of removing fatty stains such as fats, butter, salad oil, sauces and the tough stains on collars and cuffs.
Cellulases	Modify the structure of cellulose fiber on cotton and cotton blends. When it is added to a detergent, it results in; color brightening, softening and soil removal.

1. Pareto „Gleichgewichte“
2. „Optimalität“ in der Natur
3. Rationales Design
4. **Wie können wir Evolution „spielen“?**
5. Evolutionäres Design
6. Synthetische Biologie „quo vadis“?

Three necessary conditions for Darwinian evolution are:

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

Biologists distinguish the **genotype** – the genetic information – and the **phenotype** – the organisms and all its properties. The **genotype** is unfolded in development and yields the **phenotype**.

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

The Darwinian mechanism requires **no process** that could not be implemented in **cell-free molecular systems**.

Sol Spiegelman,
1914 - 1983

Evolution in the test tube:

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

DOI: 10.1002/anie.200701369

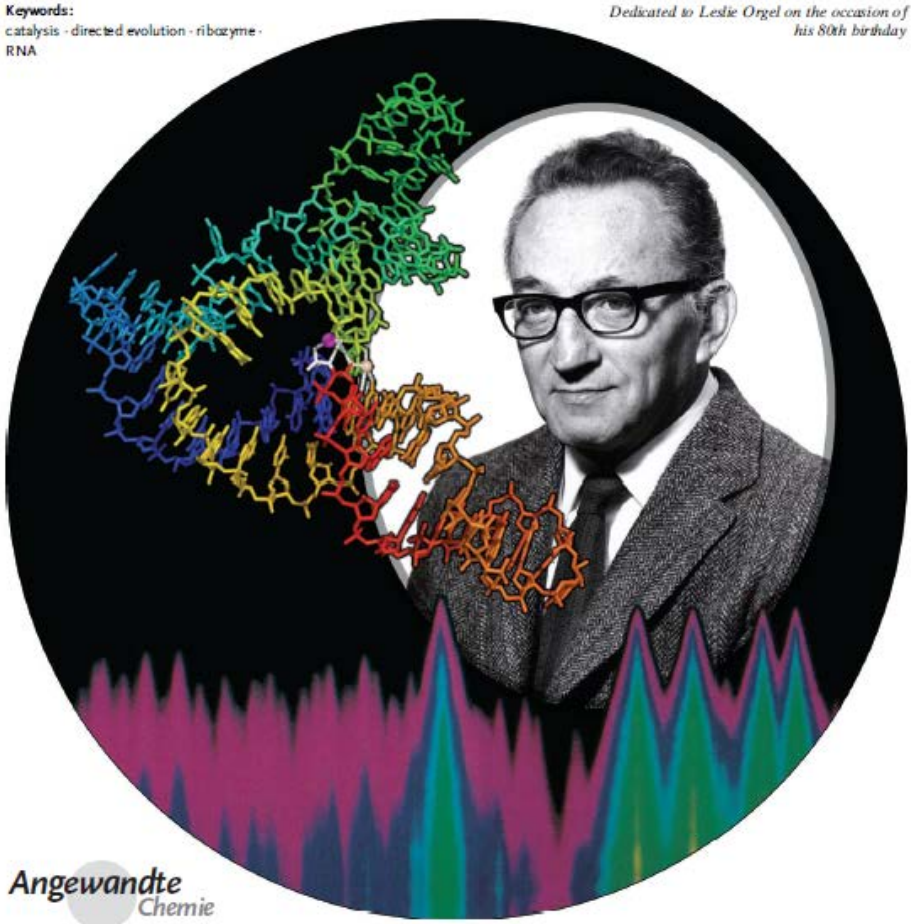
Molecular Evolution

Forty Years of In Vitro Evolution**

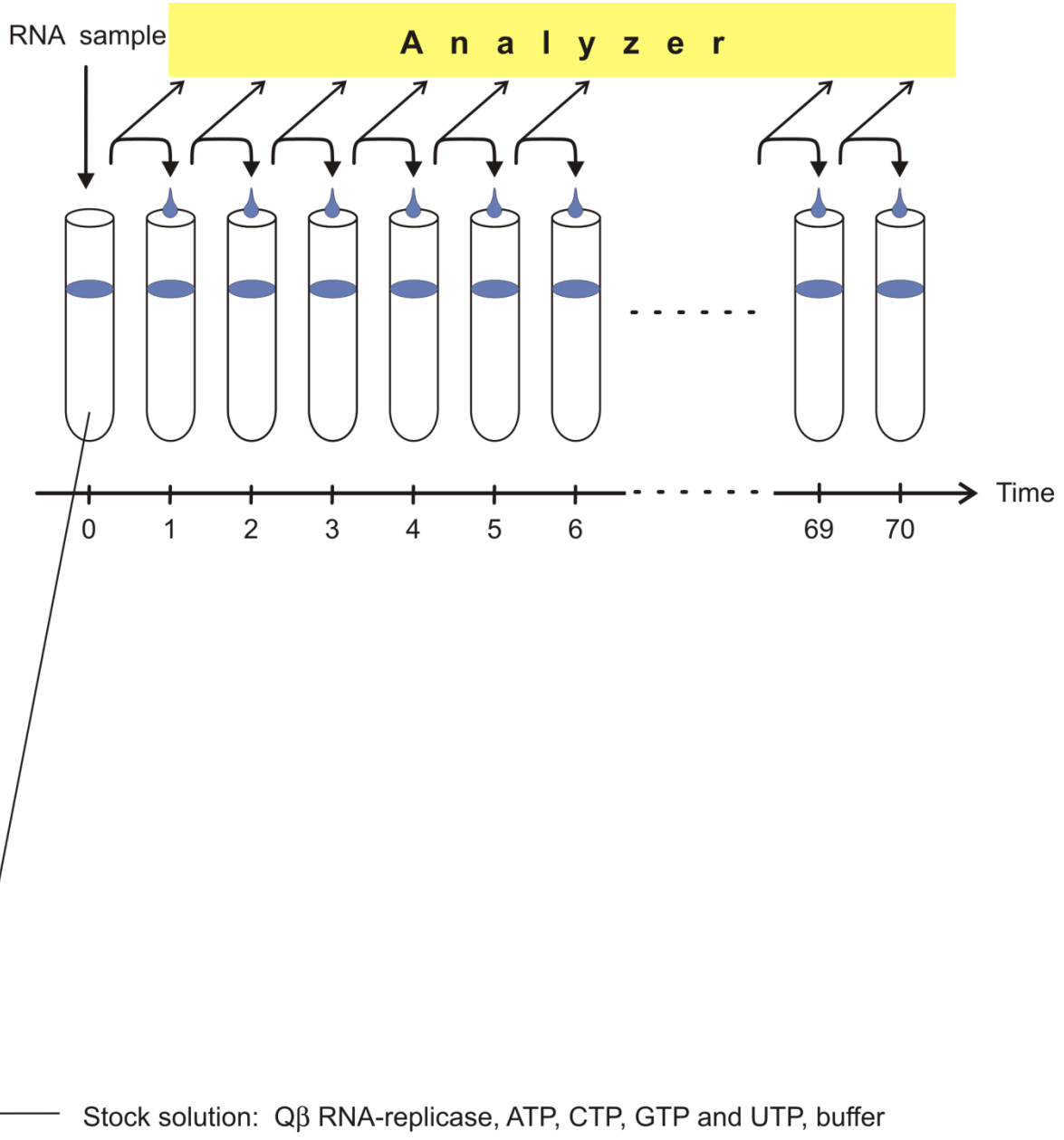
*Gerald F. Joyce**

Keywords:
catalysis · directed evolution · ribozyme · RNA

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



The serial transfer technique for *in vitro* evolution



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

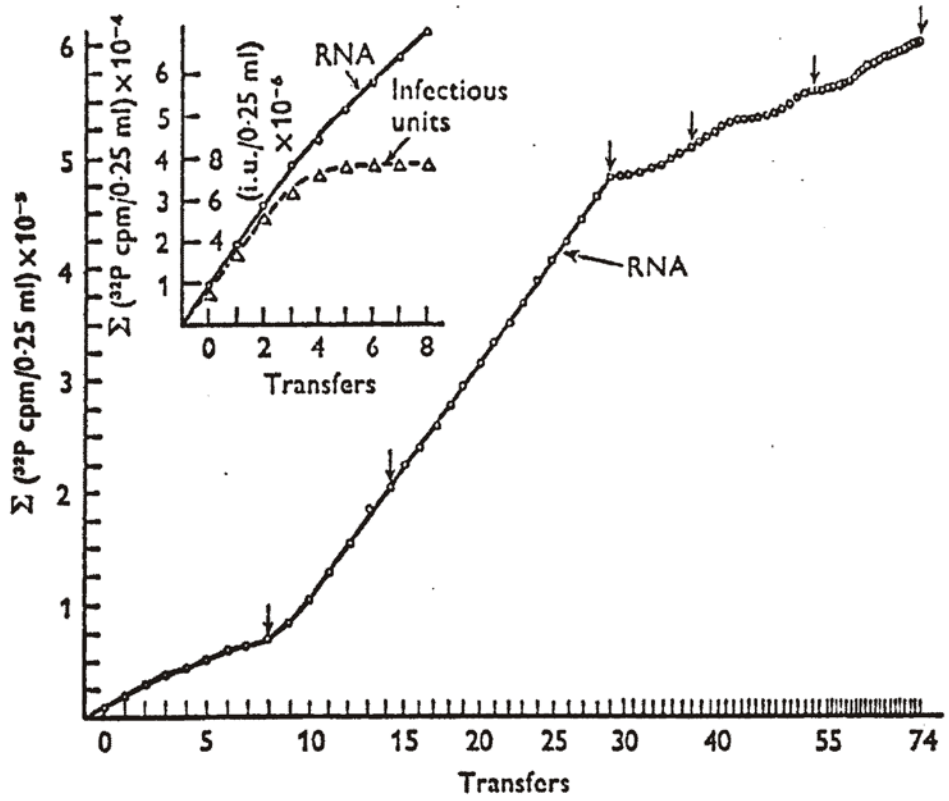


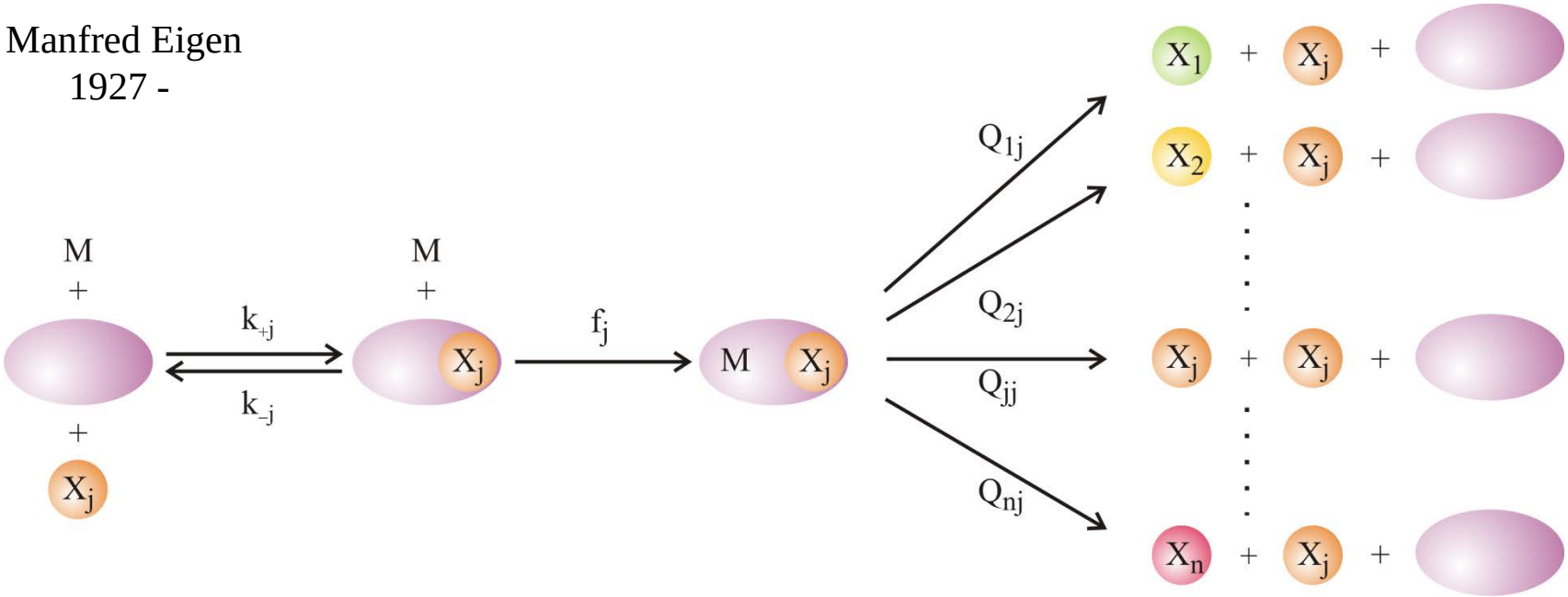
Fig. 9. Serial transfer experiment. Each 0·25 ml standard reaction mixture contained 40 μg of Qβ replicase and ^{32}P-UTP. The first reaction (0 transfer) was initiated by the addition of 0·2 μg ts-1 (temperature-sensitive RNA) and incubated at 35 °C for 20 min, whereupon 0·02 ml was drawn for counting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

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

Manfred Eigen
1927 -

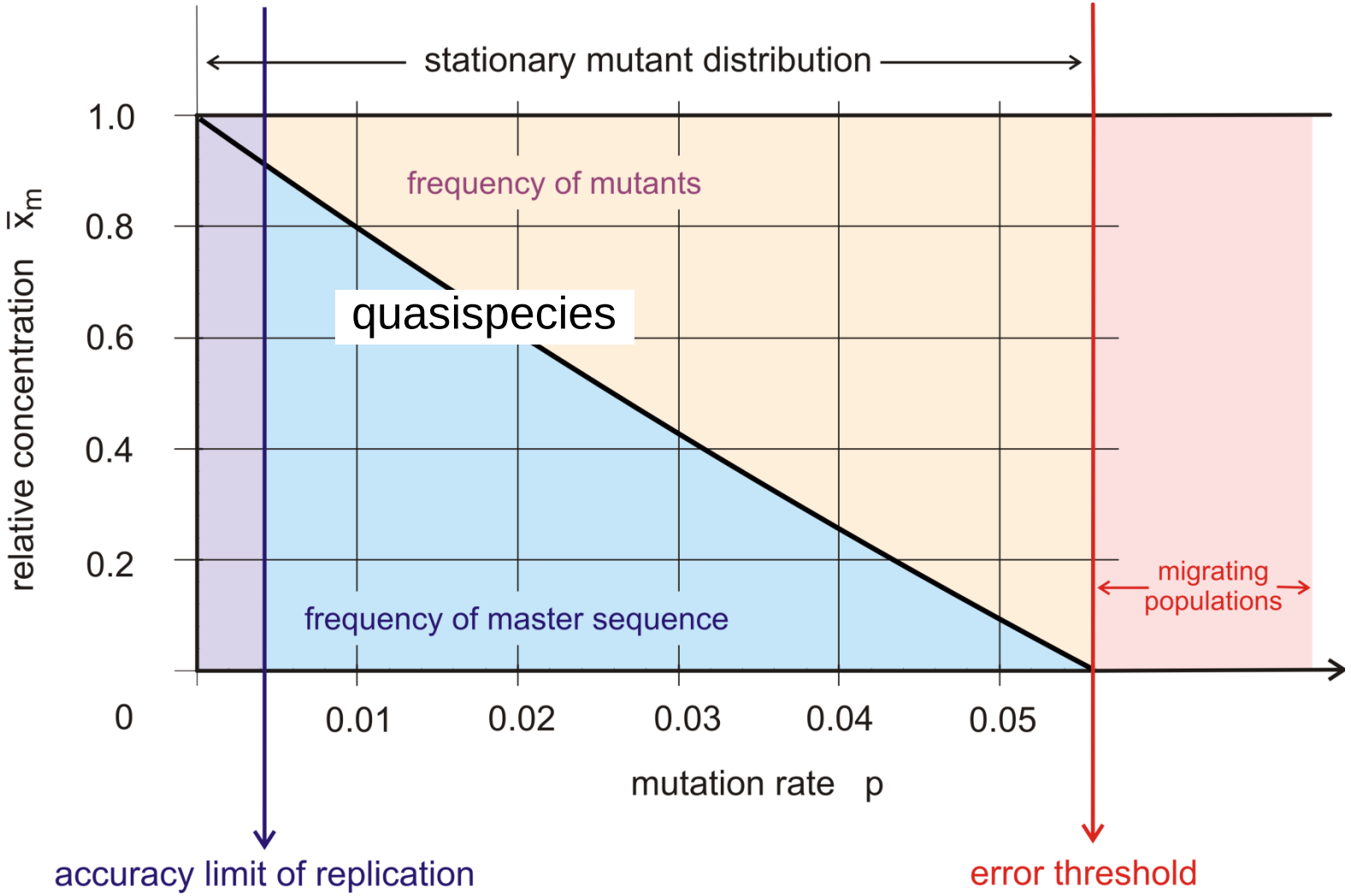
$$\frac{dx_j}{dt} = \sum_{i=1}^{n} W_{ji}\, x_i - x_j\, \Phi\,; \quad j = 1, 2, \ldots, n$$

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



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



The error threshold in replication and mutation

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.

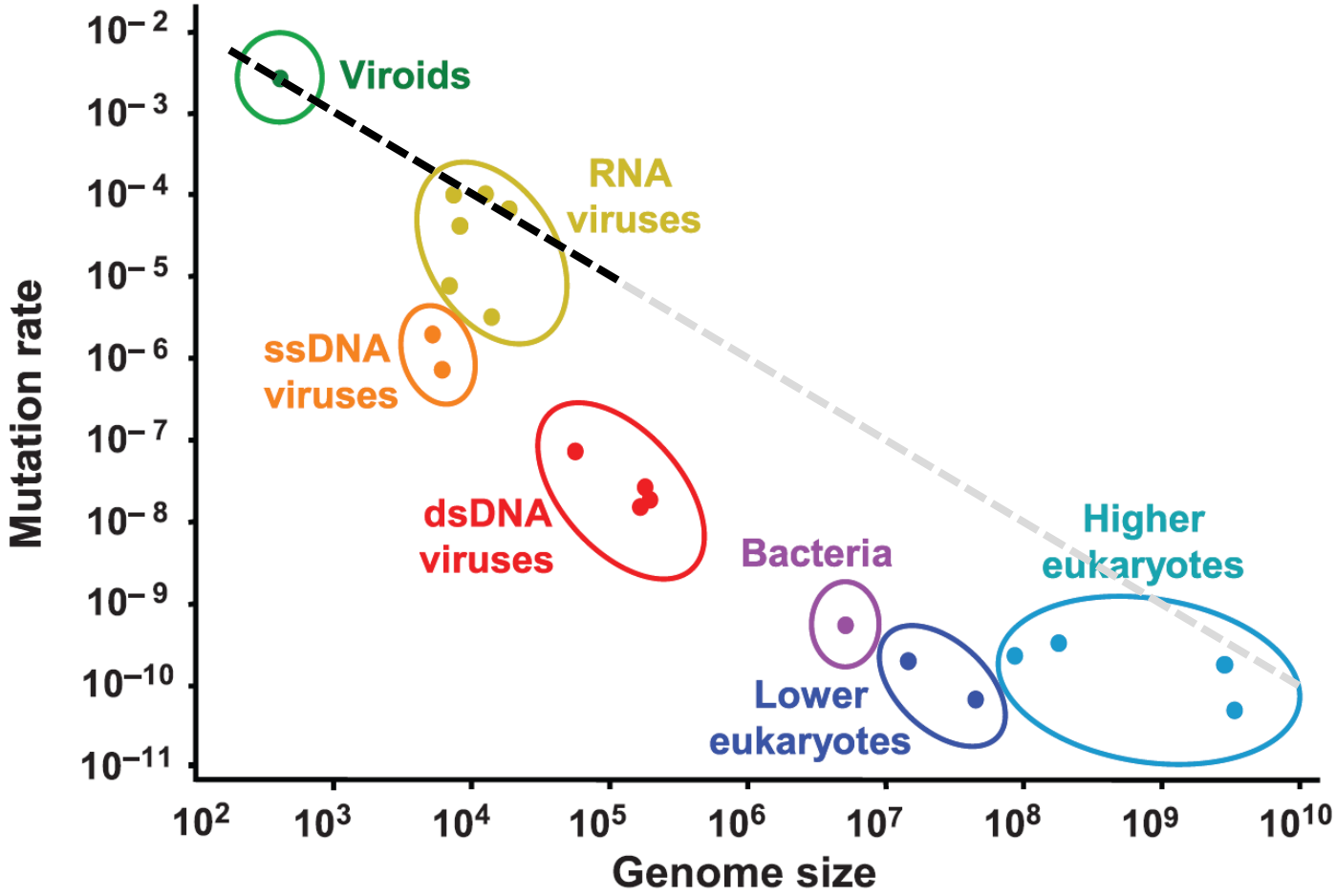
Esteban Domingo
Universidad Autónoma de Madrid
Centro de Biología Molecular "Severo Ochoa"
Consejo Superior de Investigaciones Científicas
Cantoblanco and Valdeolmos
Madrid, Spain
Tel.: +34 91 497 84858/9; fax: +34 91 497 4799
E-mail address: edomingo@cbm.uam.es

Available online 8 December 2004

doi:10.1016/j.virusres.2004.11.001

Esteban Domingo
1943 -

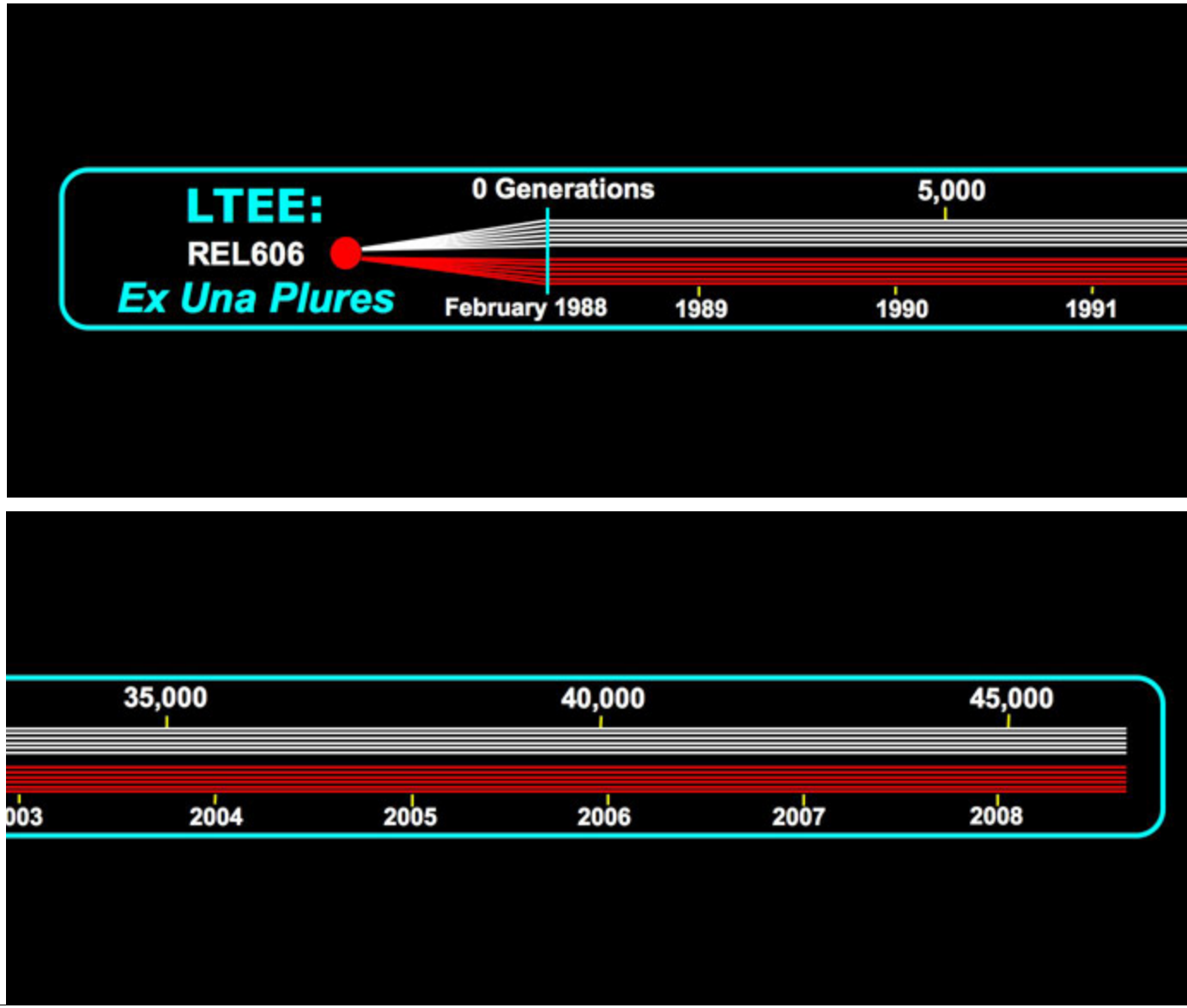
Application of quasispecies theory to the fight against viruses



Selma Gago, Santiago F. Elena, Ricardo Flores, Rafael Sanjuán. 2009. Extremely high mutation rate of a hammerhead viroid. Science 323:1308.

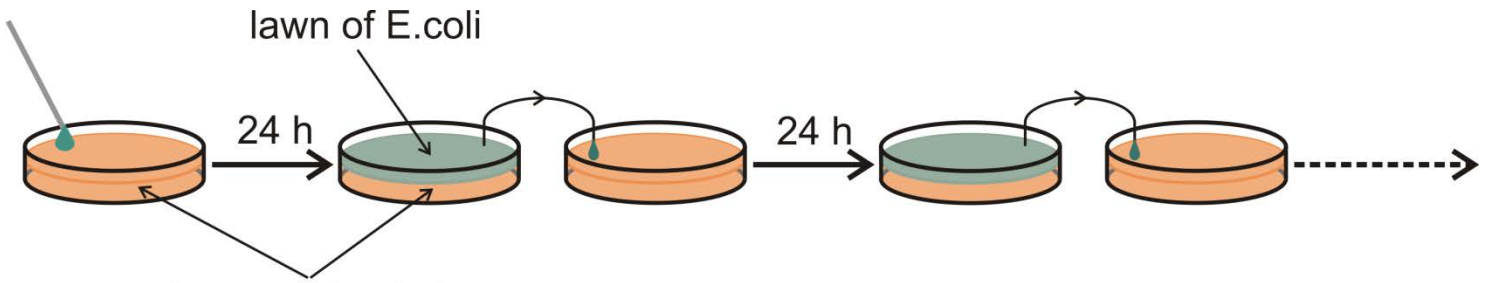
Mutation rate and genome size

Richard Lenski, 1956 -



Bacterial evolution under controlled conditions: A twenty-five years experiment.

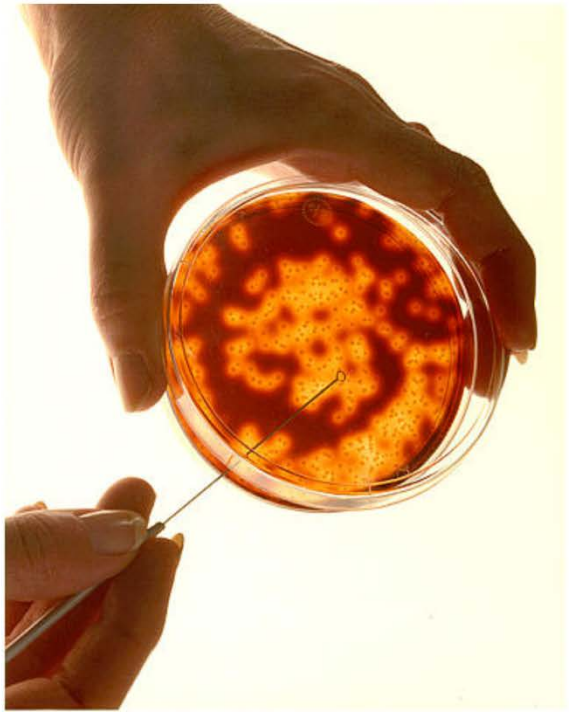
Richard Lenski, University of Michigan, East Lansing



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

1 day	≈	6.67 generations
1 month	≈	200 generations
1 year	≈	**2400 generations**

Serial transfer of bacterial cultures in Petri dishes



Bacterial evolution under controlled conditions: A twenty-five years experiment.

Richard Lenski, University of Michigan, East Lansing

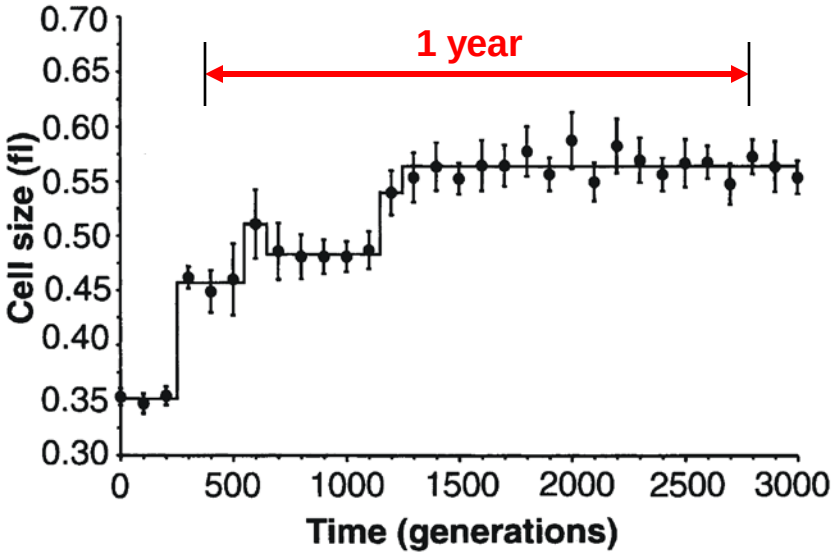


Fig. 1. Change in average cell size (1 fl = 10^{-15} L) in a population of *E. coli* during 3000 generations of experimental evolution. Each point is the mean of 10 replicate assays (*22*). Error bars indicate 95% confidence intervals. The solid line shows the best fit of a step-function model to these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

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

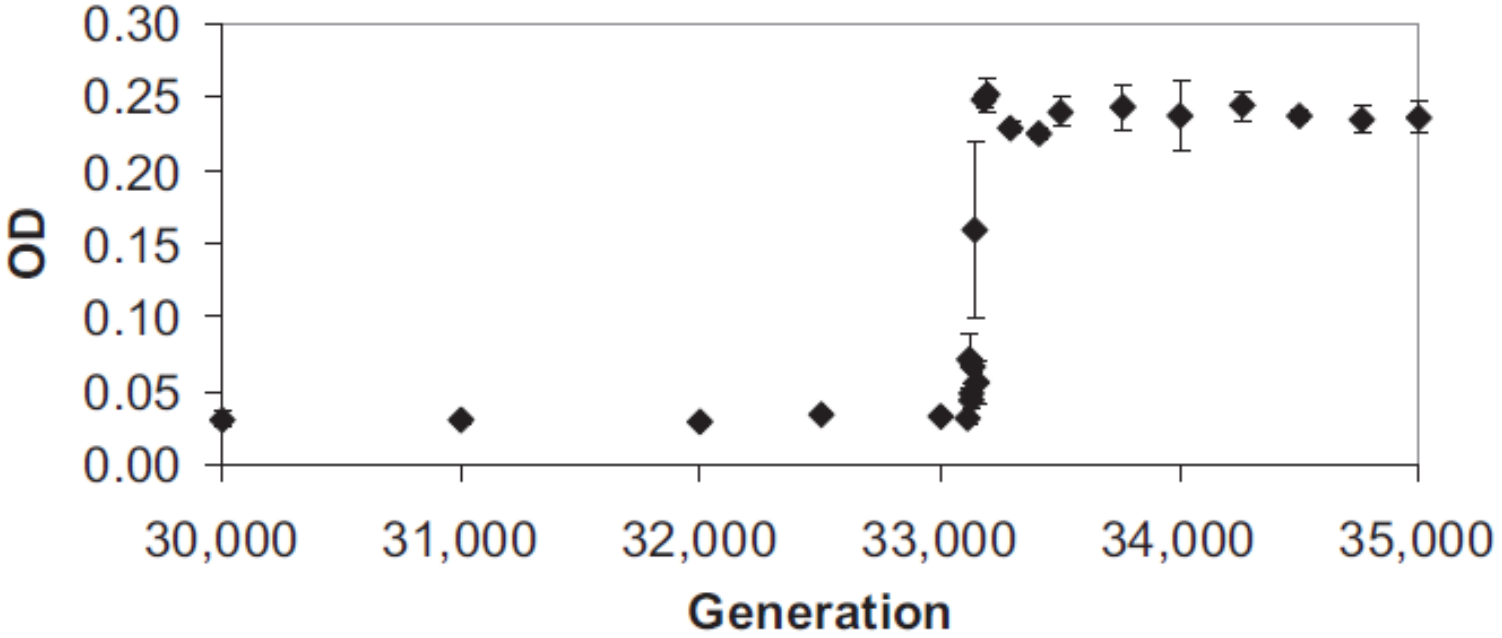


Fig. 1. Population expansion during evolution of the Cit$^+$ phenotype. Samples frozen at various times in the history of population Ara-3 were revived, and three DM25 cultures were established for each generation. Optical density (OD) at 420 nm was measured for each culture at 24 h. Error bars show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

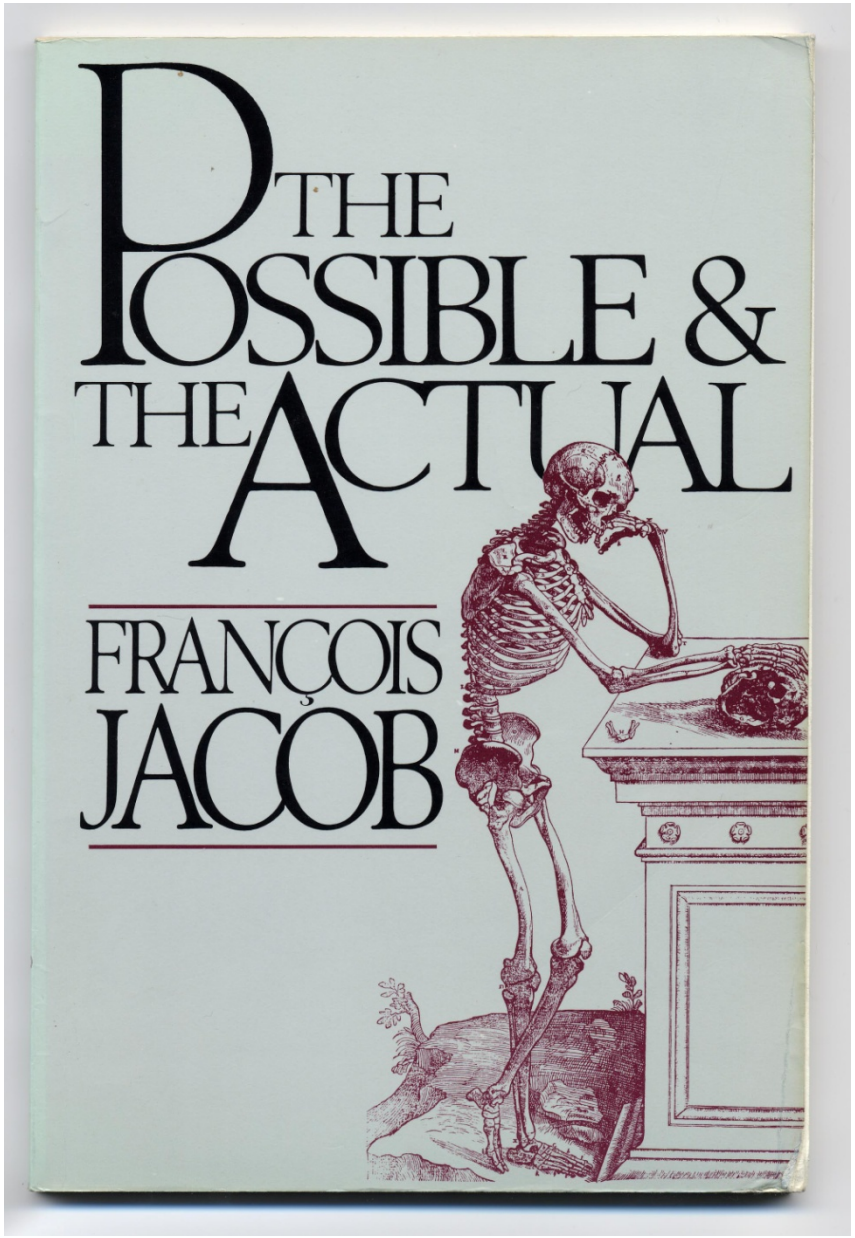
Innovation by mutation in long time evolution of Escherichia coli in constant environment

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

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



Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

François Jacob. *The Possible and the Actual*. Pantheon Books, New York, 1982, and

Evolutionary tinkering. *Science* **196** (1977), 1161-1166.

1. Pareto „Gleichgewichte“
2. „Optimalität“ in der Natur
3. Rationales Design
4. Wie können wir Evolution „spielen“?
5. **Evolutionäres Design**
6. Synthetische Biologie „quo vadis“?

The Scientist, January 1, 2006

Design: More Intelligent Every Day.

Synthetic biology requires intelligent design, but not the kind they teach in Kansas.

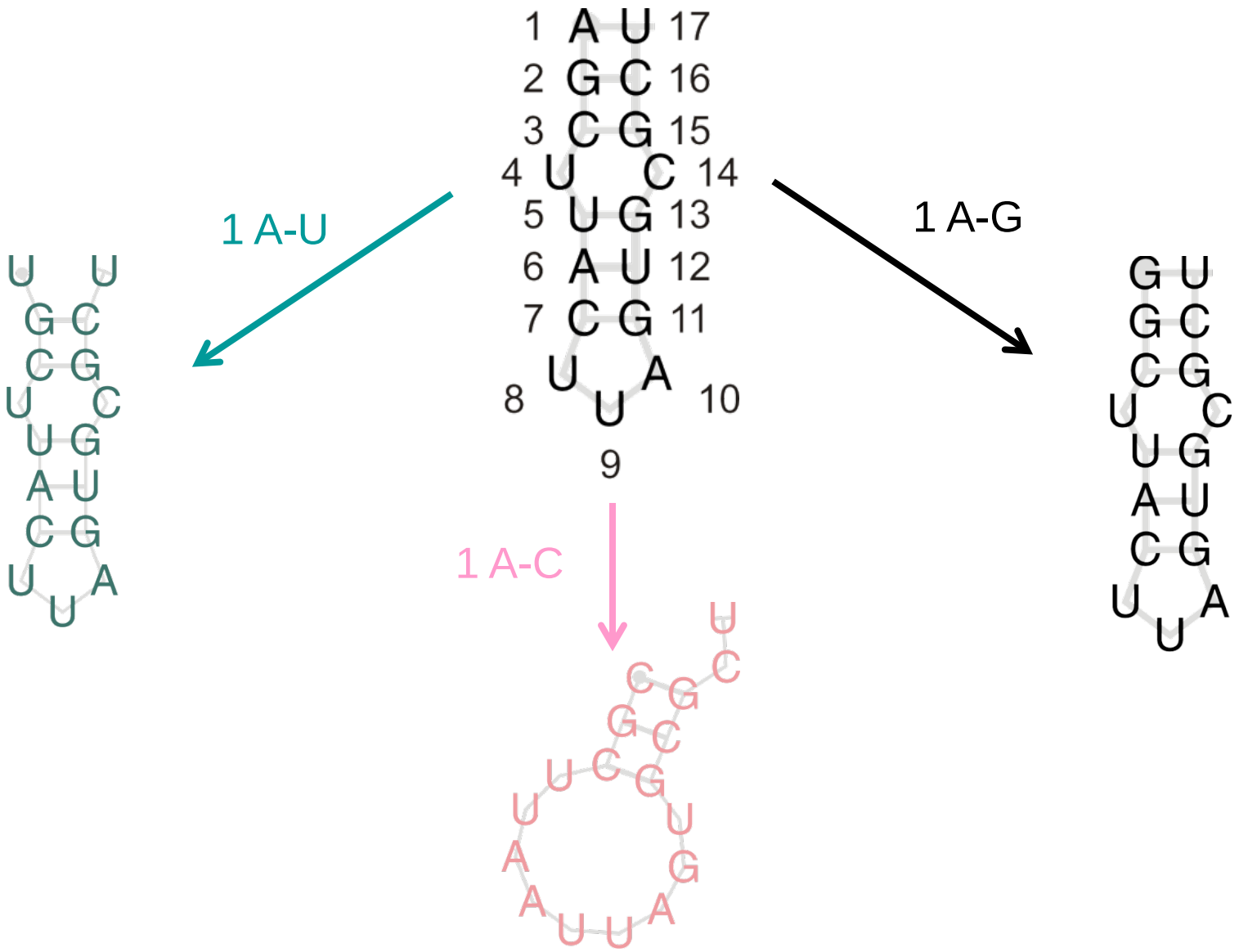
By Glenn McGee

Thanks to a recent court decision, children in Kansas will learn that the fossil record of our planet holds evidence of "irreducibly complex" traits, biological wonders that seem to sophisticated to be products of natural selection. Advocates of intelligent design argue that such complexity of biological life reveals evidence of a designer.

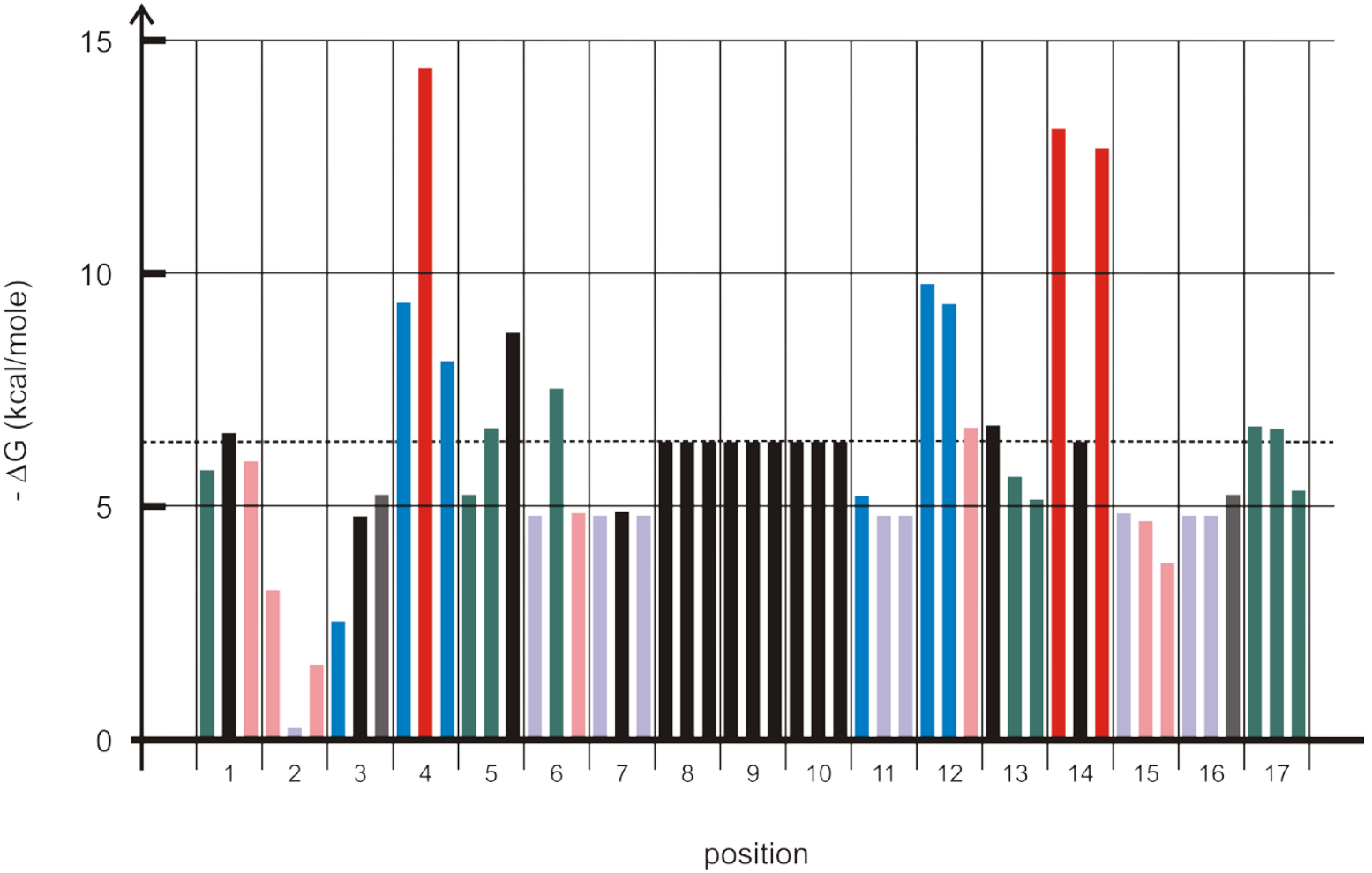
A different sort of designer is working in the nascent filed of synthetic biology. These scientists generate novel biological functions through the design and construction of living systems. Synthetic biologists manipulate the most complex biological interactions using the tools of engineering and computer science. It has borne fruit in the design of genomes, proteins, devices, integrated biological systems, and even cell-circuit hybrids. Synthetic biologists use evolution as a method. That seems pretty intelligent.

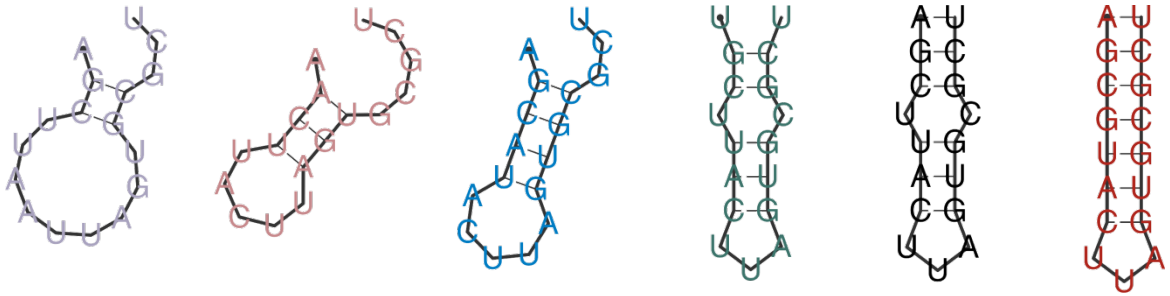
Glenn McGee is the director of the Alden March Bioethics Institute at Albany Medical College, where he holds John A. Balint Endowed Chair in Medical Ethics.

AGCUUAACUUAGUCGCU




1 A U 17
2 G C 16
3 C G 15
4 U C 14
5 U G 13
6 A U 12
7 C G 11
U A
8 U 10
9


15
10
5
0
- ΔG (kcal/mole)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
position



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.

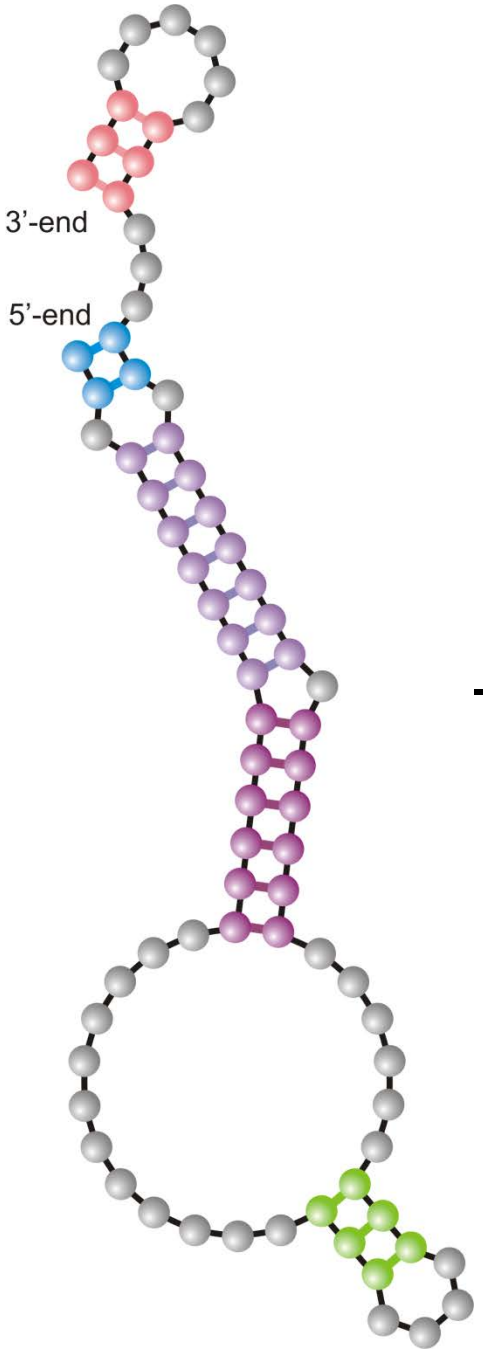
We focus on the narrow domain of RNA secondary structure, which is currently the simplest computationally tractable, yet realistic phenotype (*2*). This choice enables the definition and exploration of concepts that may prove useful in a wider context. RNA secondary structures represent a coarse level 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*).

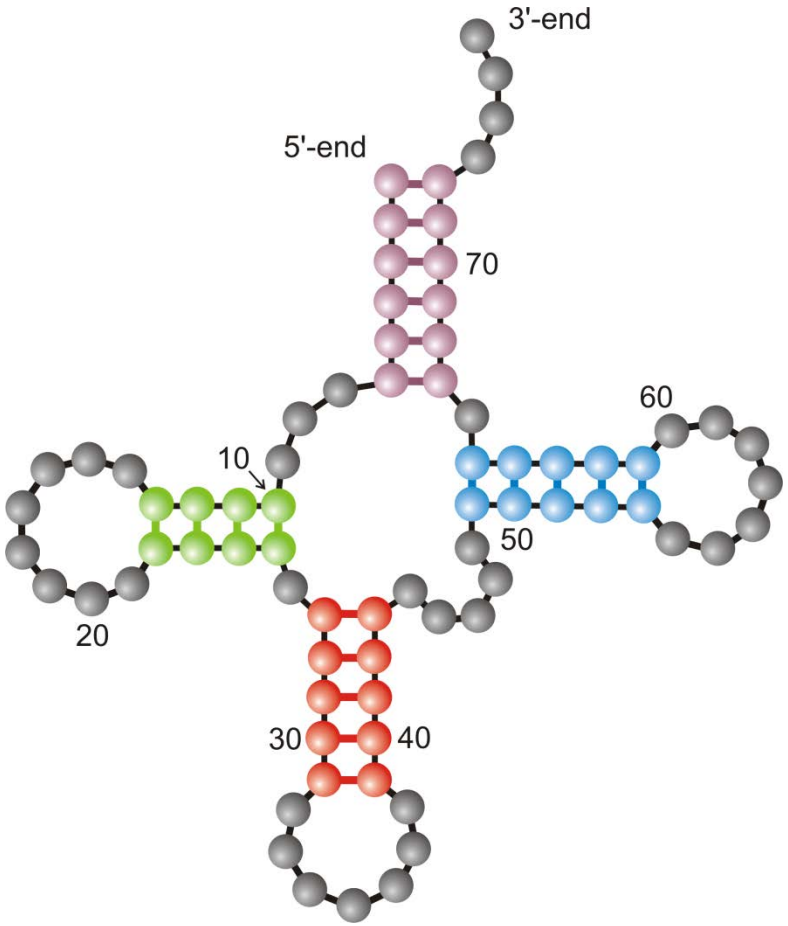
To generate evolutionary histories, we used a stochastic continuous time model of an RNA population replicating and mutating in a capacity-constrained flow reactor under selection (*5*, *6*). In the laboratory, a goal might be to find an RNA aptamer binding specifically to a molecule (*4*). Although 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

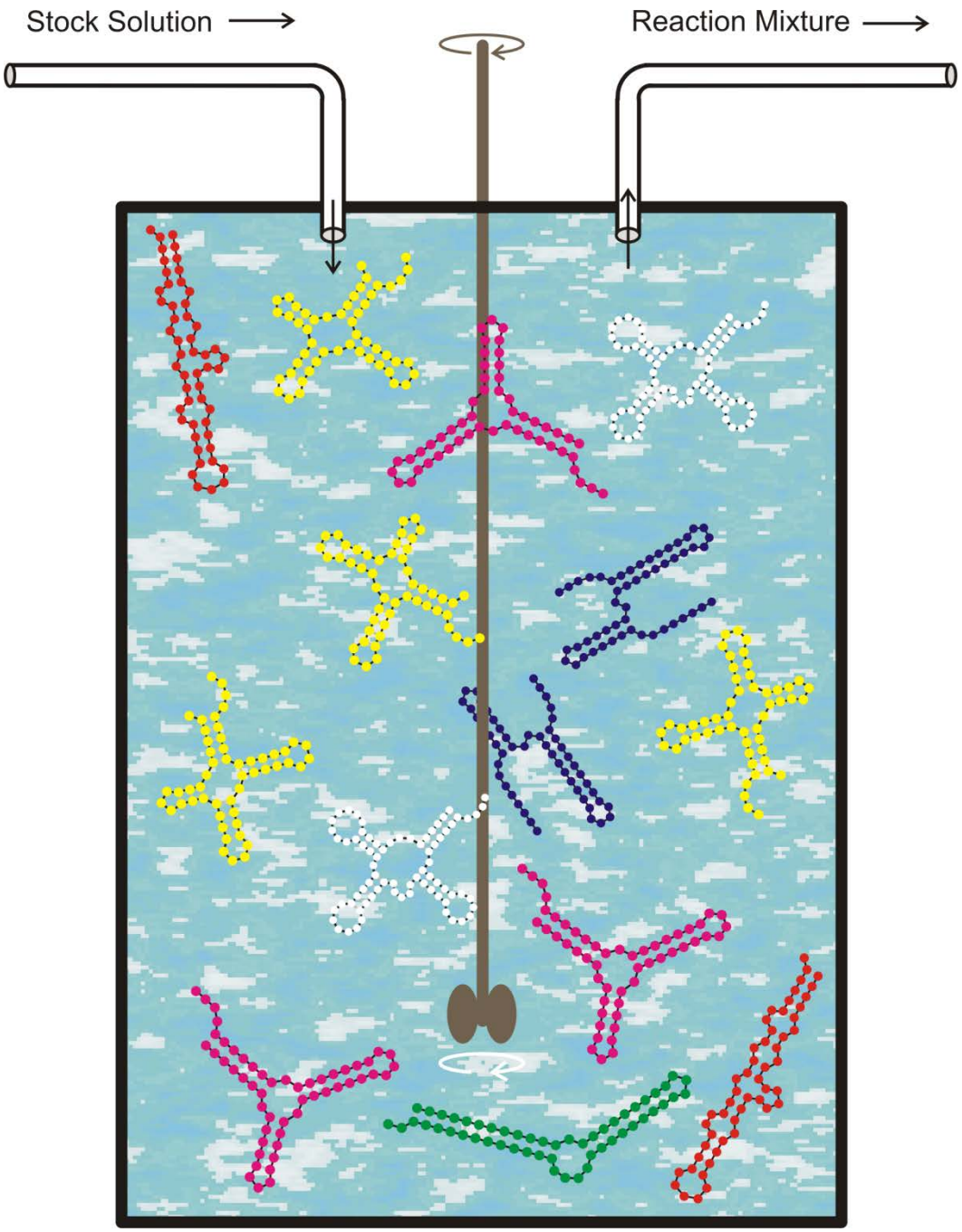
Institut für Theoretische Chemie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA, and International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria.





Structure of randomly chosen initial sequence

Phenylalanyl-tRNA as target structure



Replication rate constant

(Fitness):

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

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

Selection pressure:

The population size,

N = # RNA moleucles,

is determined by the flux:

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

Mutation rate:

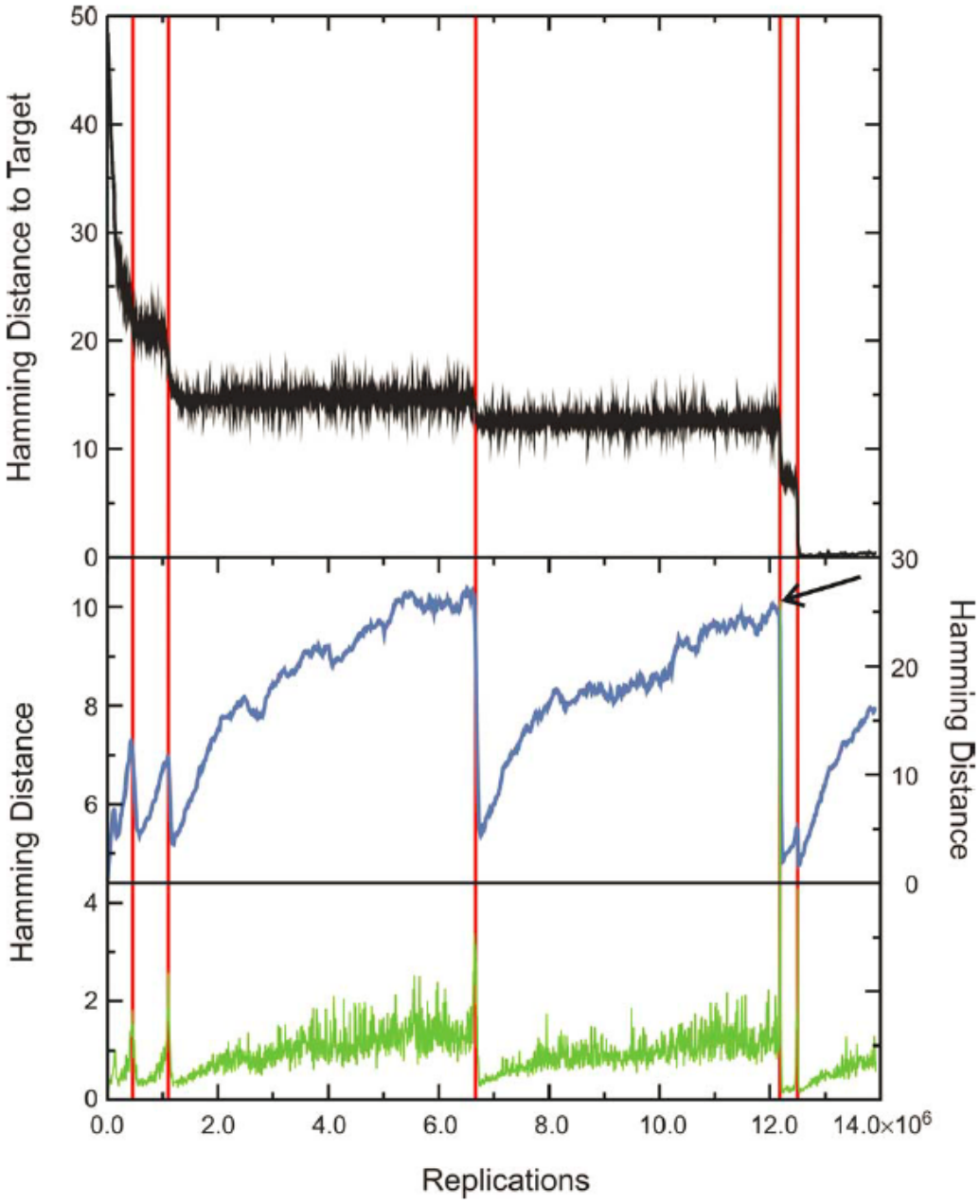
p = 0.001 / Nucleotide × Replication

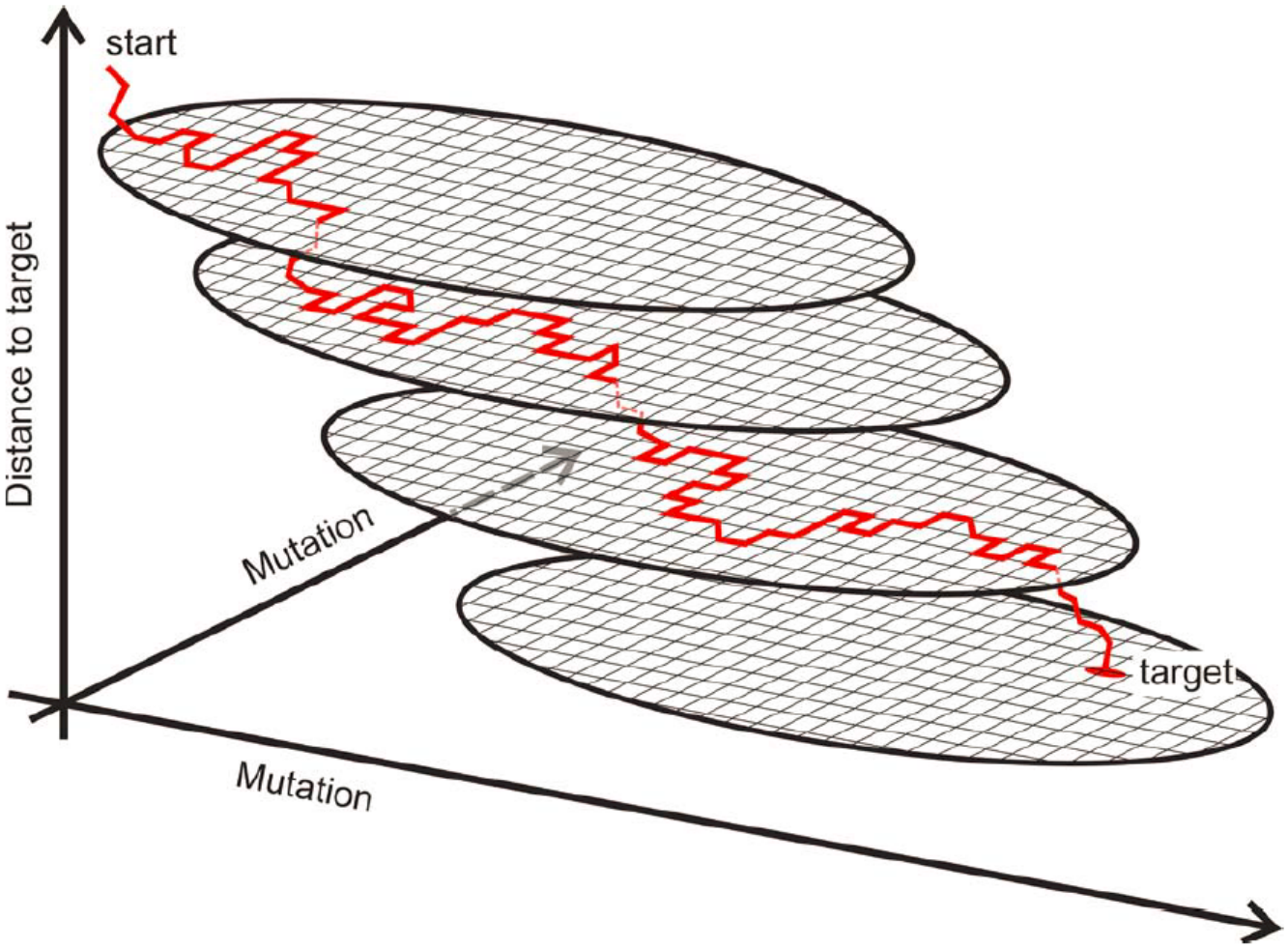
The flow reactor as a device for studying the evolution of molecules ***in vitro*** and ***in silico***.

Evolutionary trajectory

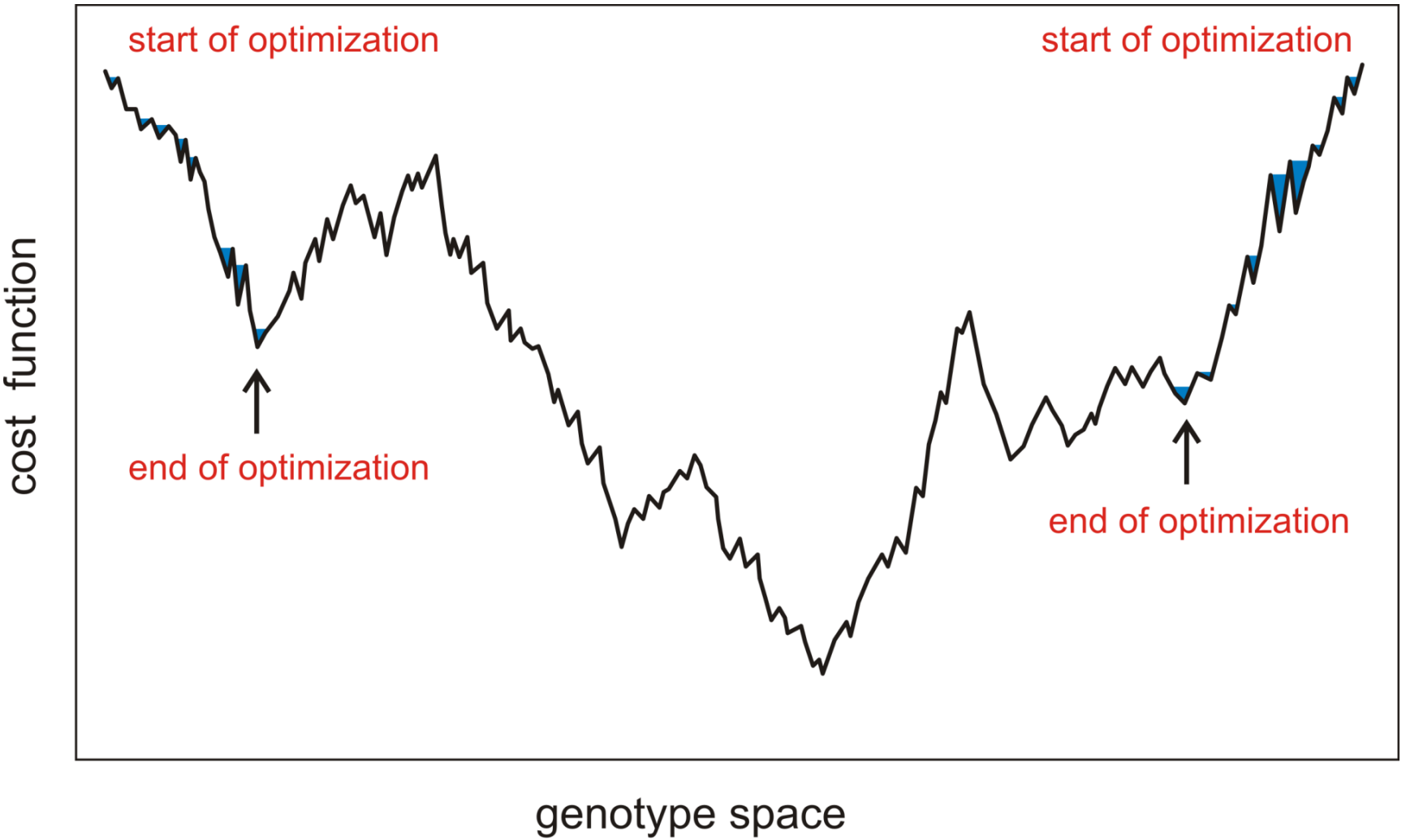
Spreading of the population on neutral networks

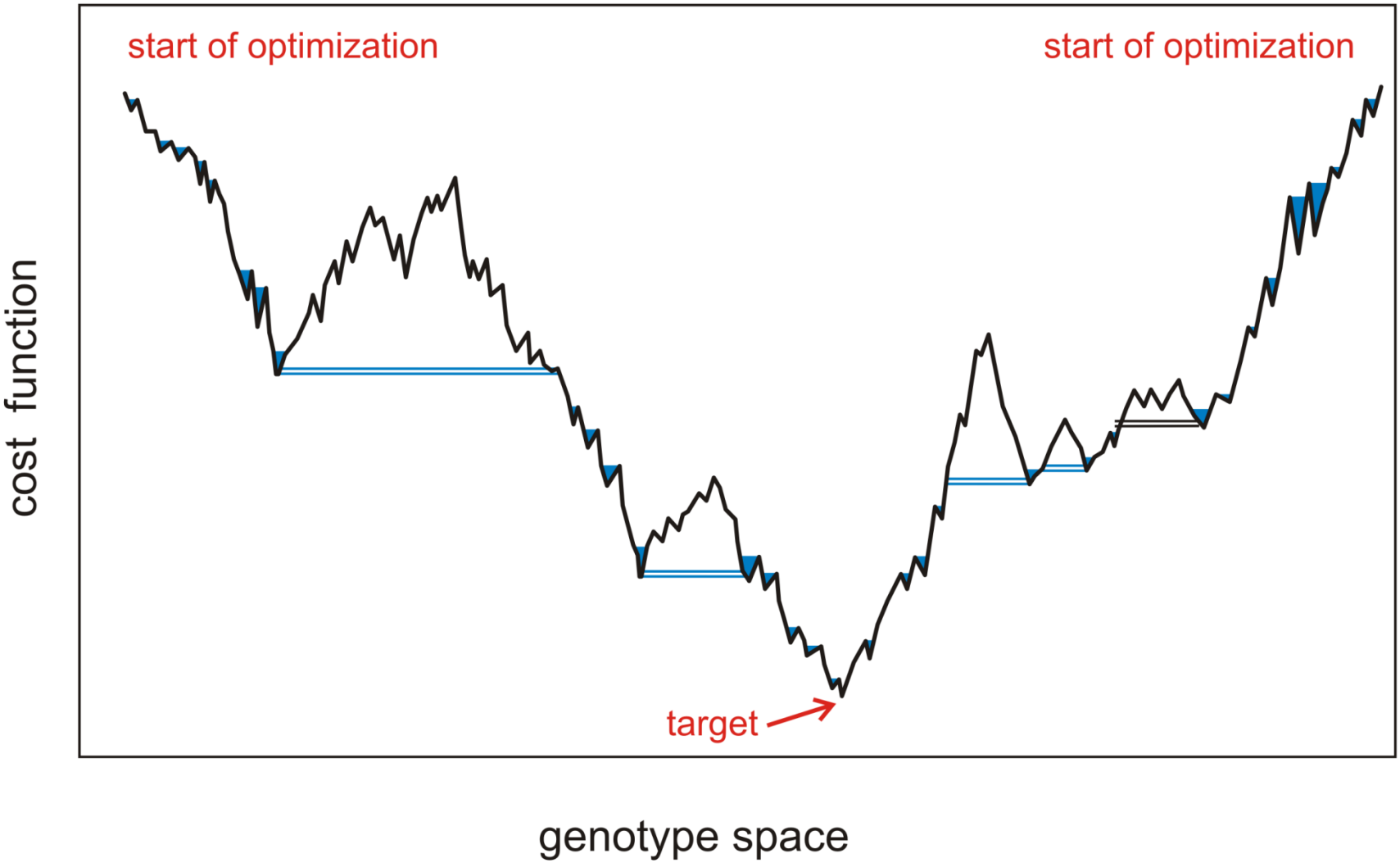
Drift of the population center in sequence space

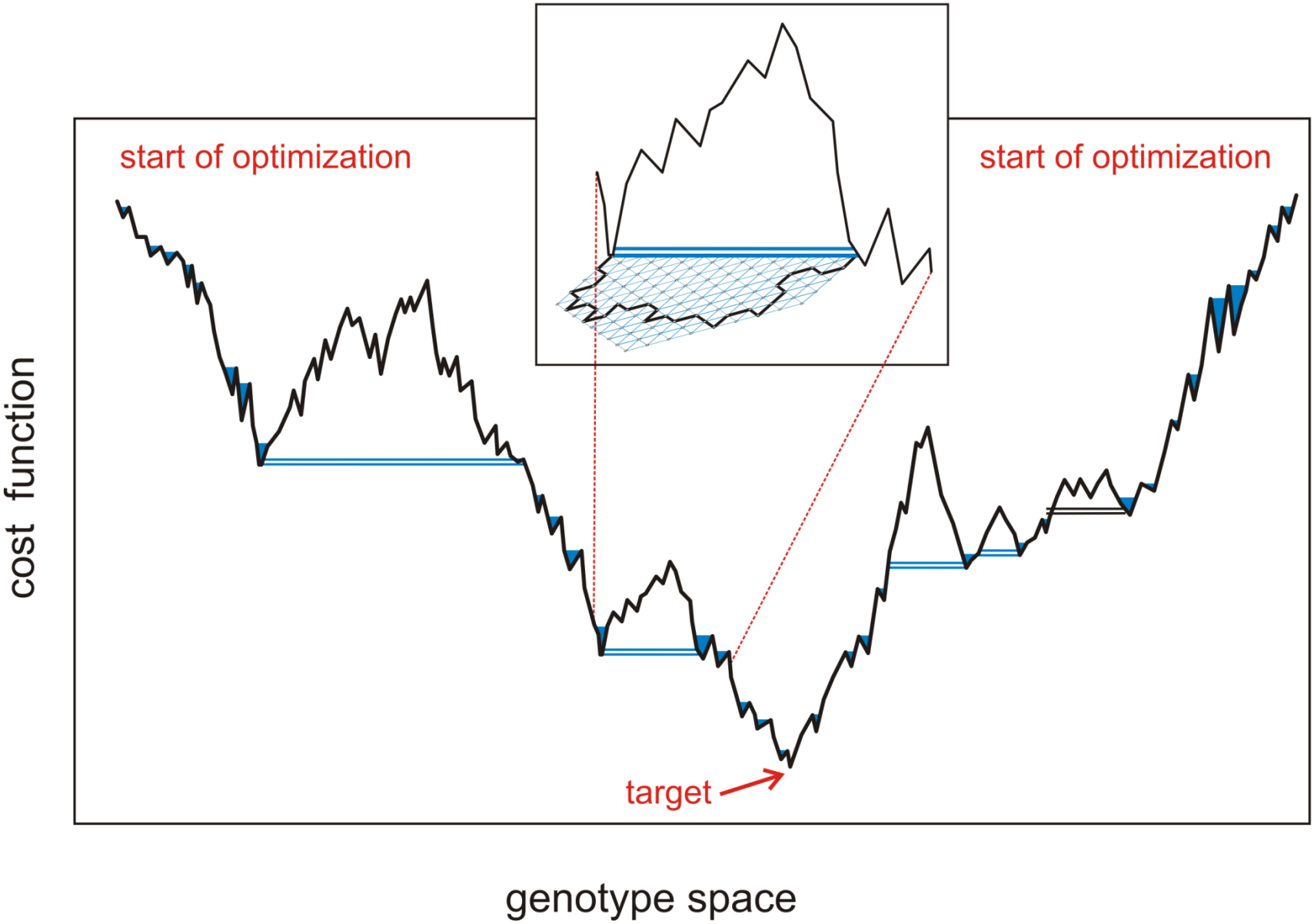




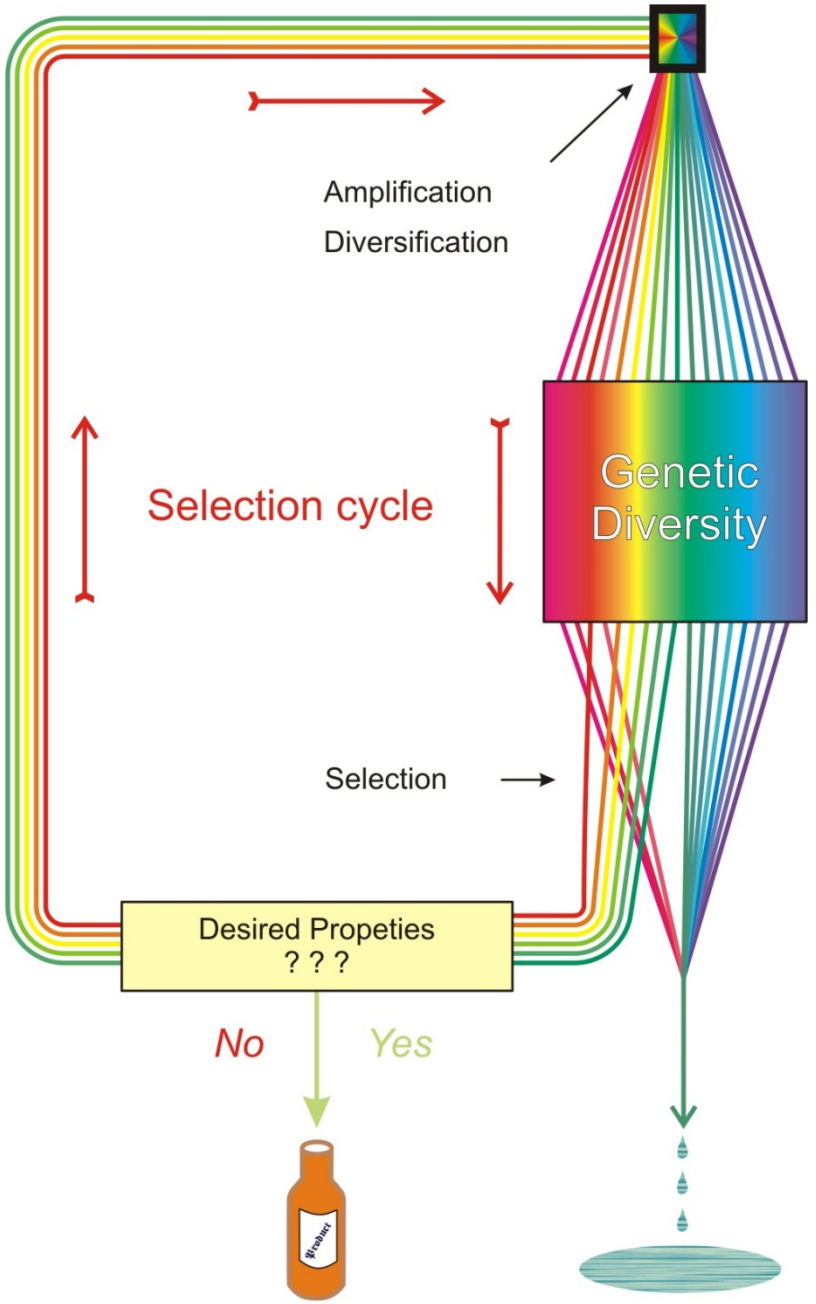
A sketch of optimization on neutral networks

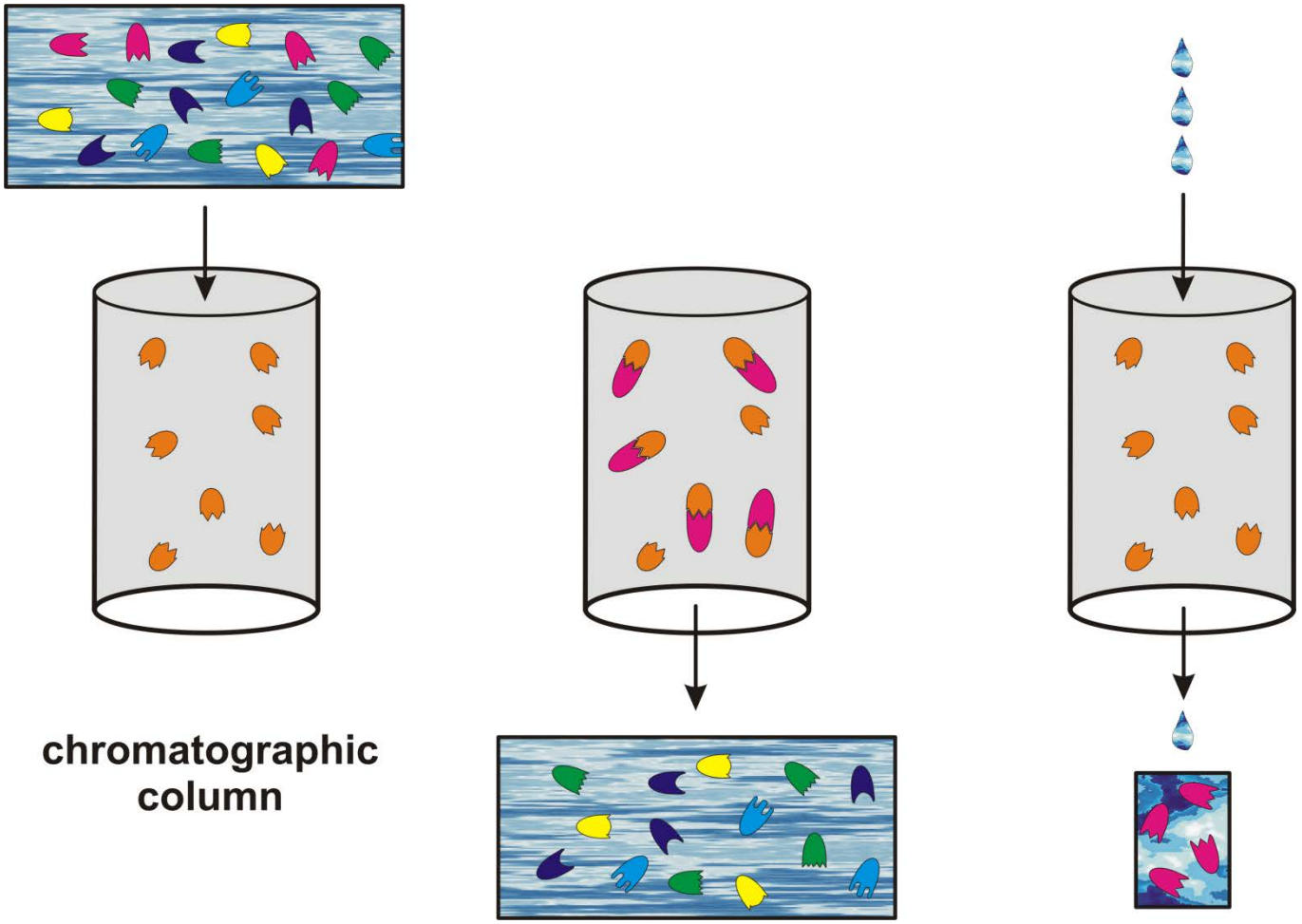

start of optimization
start of optimization
end of optimization
end of optimization
cost function
genotype space


start of optimization
start of optimization
target
cost function
genotype space

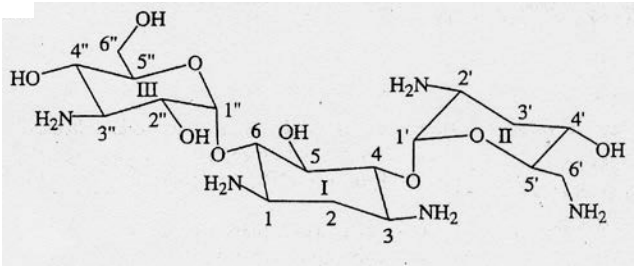

start of optimization
start of optimization
target
cost function
genotype space

Ein Beispiel für Selektion von Molekülen mit vorbestimmbaren Eigenschaften im Laborexpriment

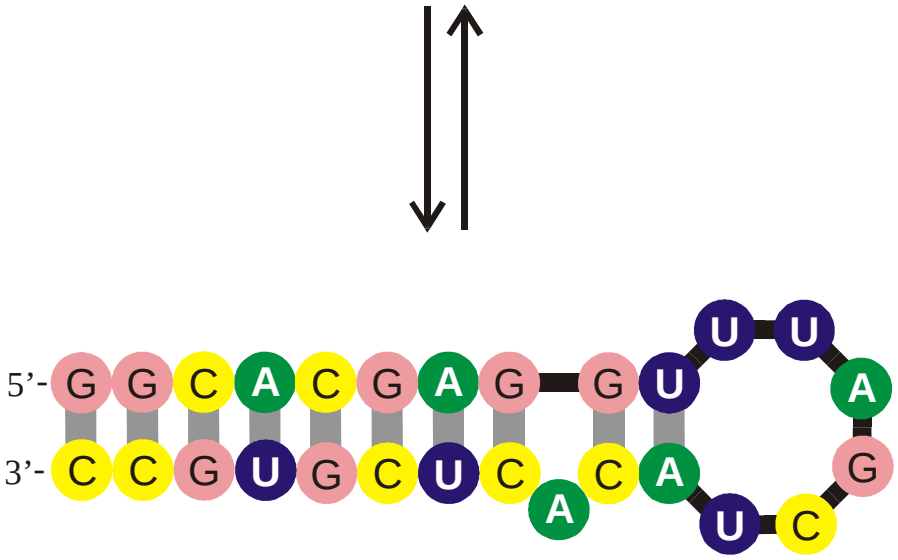




Die SELEX-Technik zur evolutionären Erzeugung von stark bindenden Molekülen



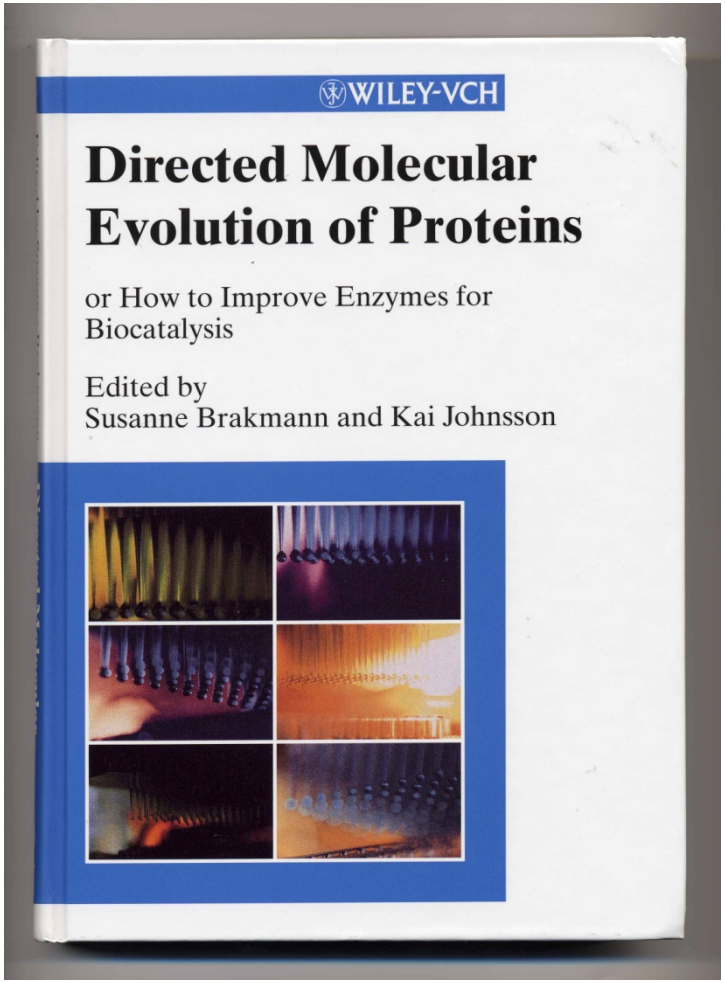
tobramycin

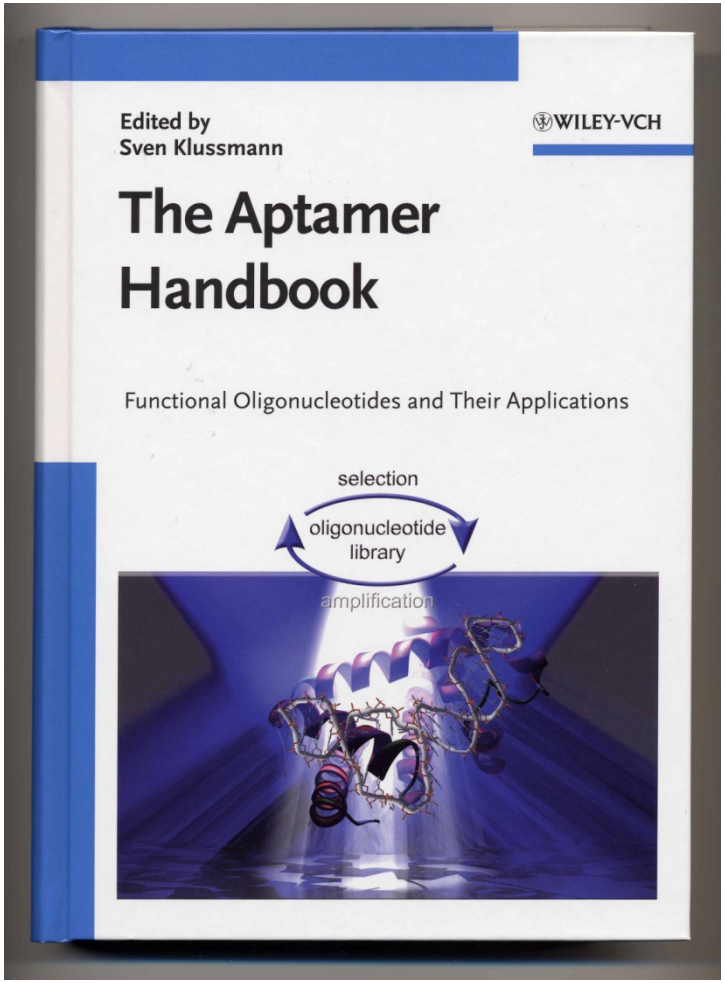


RNA aptamer

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

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





Application of molecular evolution to problems in biotechnology

1. Pareto „Gleichgewichte“
2. „Optimalität“ in der Natur
3. Rationales Design
4. Wie können wir Evolution „spielen“?
5. Evolutionäres Design
6. **Synthetische Biologie „quo vadis“?**

Der Begriff „synthetische Biologie“ taucht 1913 erstmals in der Literatur auf.

Nature **91**:270-272 (1913)

very coarse shingle one turns from the erosion hypothesis, and the slightly pitted nature of the rock surface suggests solution. ALEX. STEVENS.

Geological Department, University of Glasgow, May 6.

The Mountains and their Roots.

MAJOR COWIE'S letter in NATURE of May 8 gives the impression that I had the facts of the observations on the deflection of the plumb-line in India before me, and that I made my assumptions as to relative densities, and the mode of compensation by extension of depressed crust beneath the plains, "suitably adjusted," so as if possible to bring out the desired results. This was not the case. I made the assumptions about relative densities which seemed to be *a priori* probable; and it will be seen from the diagram at p. 184 of my "Physics of the Earth's Crust" that fifteen years before I wrote the paper in the *Phil Mag.* I had suggested that compressed mountains would be partly supported by an extension of the depressed crust beyond them.

Should anyone be inclined to undertake the labour of calculating from my formulæ, introducing fresh constants, or other distances, I would warn him that in the *Phil. Mag.* there is a misprint. In the formula for the plateau, after the first bracket, insert x.

I am much pleased that after so long a time my theories are under discussion, and I hope to come well out of it. I am sending to the *Geological Magazine* a reply to some remarks by Sir T. H. Holland in that journal, and to this I would refer your readers as more fully giving my views on some of the points under discussion. O. FISHER.

Graveley, Huntingdon, May 9.

An Application of Mathematics to Law.

I HAVE read Mr. Potts's letter in NATURE of April 24, but am at a loss to understand the use to which he would put his equations.

If it be his object to find some equation giving the validity of a patent or foretelling in any way the probability of its being upheld in a court of law, he has clearly failed to do anything of the sort.

If his equation $I=M+i$ is to be of any value, the quantity i must have a fixed value greater than zero. In fact, however, for any given patent, i may have an infinite number of values, including zero, since each person will have his own idea of the amount of ingenuity that must be shown in the particular case by the inventor. Thus the inventor will certainly put a high positive value upon i, while his opponent will as certainly say that the value of i is zero. It is clear that the value of i can only be finally settled when the validity of the patent has been settled by the House of Lords, and at this stage of a patent's career it is scarcely necessary to have an equation to test its validity. So far as the rest of his letter goes, he seems to have chosen a rather complex method of setting out a few of the chief principles of patent law.

R. STAFFORD CRIPPS.

Fulmer, Slough.

I DID not imagine that my letter would be taken as an attempt to supersede the present methods of determining validity. I intended it as a contribution to the theory which underlies the enormous volume of our case-law on the subject. Surely, as in other cases of the progress from empiricism to science, the first step must be in the direction of mathematical or symbolic expression of the facts. The value of such a symbolism is twofold: first, as an aid to precision of thought; and second, as a preliminary to generalisation. It is a vital principle of English law that all decisions shall harmonise with precedents as much as possible, and on this account alone anything should be of value which assists in formulating generalisations. We admit the value of theory in the physical sciences, apart from immediate practical results: why should an attempt to develop a theory of law be condemned because it does not at once do away with the functions of the judge?

Mr. Cripps's difficulty as to the value of i will not be so great if the actual cases given in my letter are studied. I may add here, however, that it is immaterial what this value is, provided that it is measurably greater than zero. It is settled law that a scintilla of ingenuity is sufficient to support a patent for something new and useful (*cf.* Thompson *v.* Amer. Braided Wire Co., in the House of Lords, and other cases). I therefore employed this symbol merely to indicate that there had to be some positive difference.

HAROLD E. POTTS.

University Club, Liverpool.

SYNTHETIC BIOLOGY AND THE MECHANISM OF LIFE.

THE presidential address delivered by Prof. Schäfer to the British Association in 1912, and the subsequent independent discussion at a joint sitting of two of the sections, served, as was pointed out by Prof. Armstrong in a paper in *Science Progress* in October last, "as a useful corrective to the wave of vitalism that has passed over society of late years owing to the pervasive eloquence of Bergson and other writers." Probably the majority of those who have studied the phenomena of life from the chemical side will agree with Prof. Schäfer in his dictum that "at the best vitalism explains nothing," and accept his opinion "that we may fairly conclude that all changes in living substance are brought about by ordinary chemical and physical forces." The difficulty, however, lies in obtaining any satisfactory information as to what are the actual chemical or physical changes which occur in the real living cells or tissues. Since this discussion was held Prof. S. Leduc, of the School of Medicine at Nantes, has published a monograph[1] in which he approaches the problem from the novel point of view which now for several years past has guided his experiments and with which readers of his "Mechanism of Life" will be familiar.

It is impossible to do justice to the author's arguments or make clear the proper value of his demonstrations in a short article such as the present, but this will at least serve to direct attention to a few of the very remarkable results that he claims to have achieved, which, if verified, are certainly of the highest significance to the student of the phenomena of life.

The basis of Prof. Leduc's work may be summarised in his own words as follows: "It is in the physico-chemistry of liquids that an explanation of the phenomena of life is to be sought"; and he develops his views largely by studying the nature of diffusion in liquids and the phenomena

[1] "La Biologie Synthétique." By Prof. Stéphane Leduc. Pp. ii+223 (Paris: A. Poinat, 1912.)

Milestones of synthetic biology

1953 Strukturmodel der DNA: Watson, Crick, Wilkins, Franklin

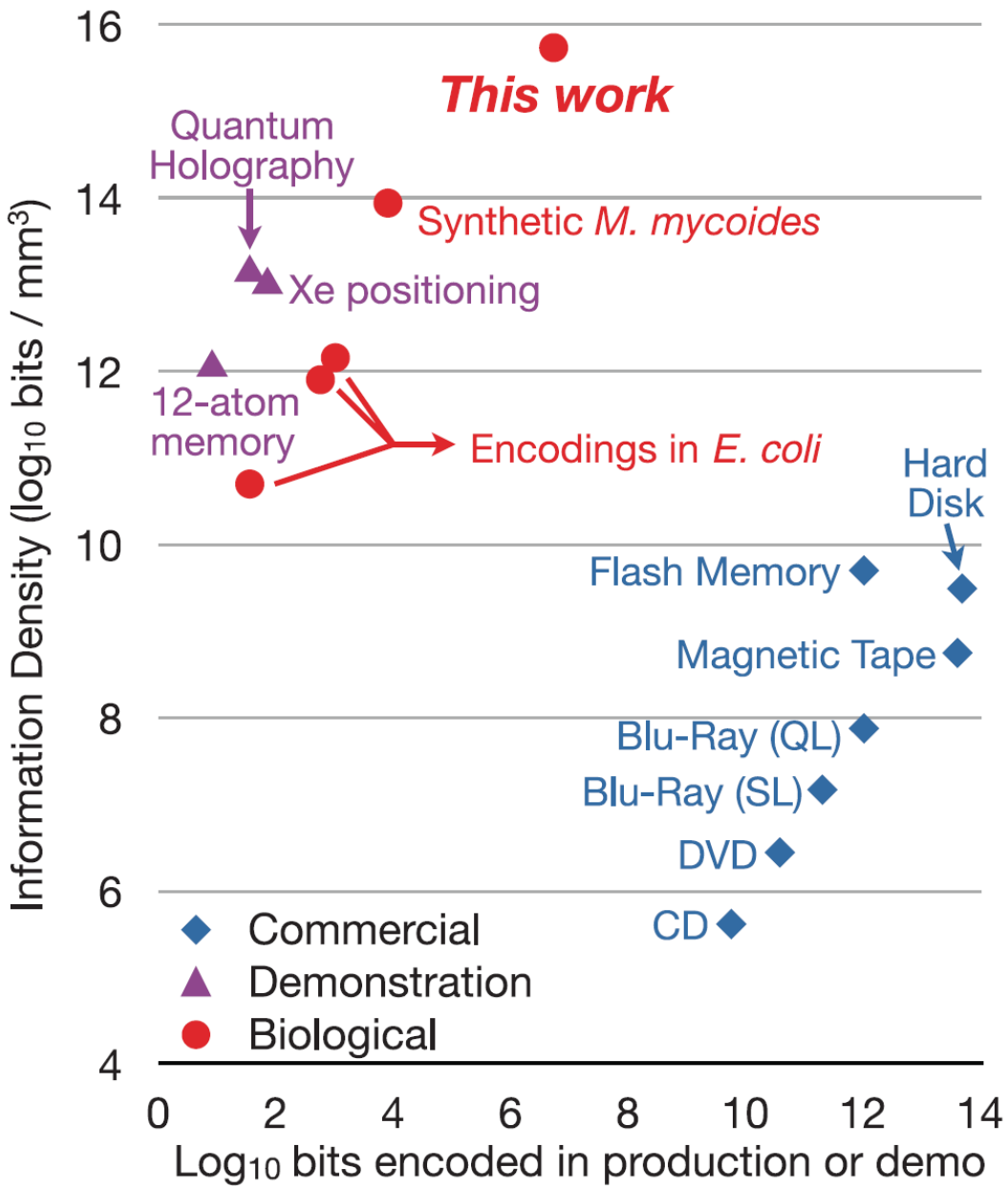
1972 Gezielte Manipulation der DNA mit Restriktionsnukleasen
1978 Nobelpreis an Arber, Nathans und Smith

1983 Erste transgene Pflanze

1997 Klonen eines Säugetieres: „Dolly"

2000 Einschleusen von Regulatorgenen in Bakterien

2006 Chemische Synthese und Einschleusen eines Genoms



G.M. Church, Y. Gao, S. Kosuri. Next-generation digital information storage in DNA. *Science* **337**:1628, 2012

For years, scientists have hoped that biology would find its engineering counterpart – a series of principles that could be used as reliably as chemical engineering is for chemistry. Thanks to major advances in synthetic biology, those hopes may soon be realized.

Kevin Munnelly. Engineering for the 21st Century: Synthetic Biology.
ACS Synthetic Biology, Viewpoint, April 09, 2013

Danke für die Aufmerksamkeit!

Web-Page für weitere Informationen:

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