Small RNA inside and outside the cell

Memories on early evolution or recent developments?

(Title by courtesy of Eberhard Neumann)

Peter Schuster

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

42. Winterseminar

Klosters, 14.– 27.01.2007

Recent review article:

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

Web-Page for further information:

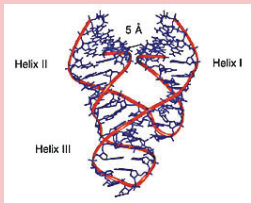
http://www.tbi.univie.ac.at/~pks

1. The exciting RNA story
2. Why is gene regulation so complex?
3. What small RNAs can achieve
4. Structures of small RNAs
5. Riboswitches and kinetic folding

1. **The exciting RNA story**
2. Why is gene regulation so complex?
3. What small RNAs can achieve
4. Structures of small RNAs
5. Riboswitches and kinetic folding

RNA

RNA *as catalyst*



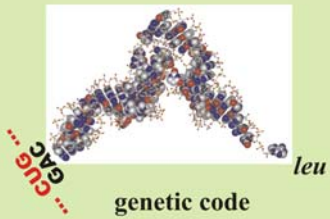
Ribozyme

RNA *as transmitter of genetic information*

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

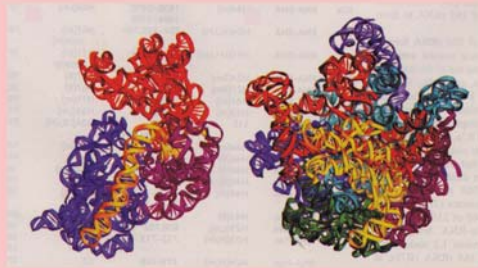
RNA as **working copy** of genetic information

RNA *as adapter molecule*



genetic code

RNA *is the catalytic subunit in supramolecular complexes*



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

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

RNA *is modified by epigenetic control*

RNA editing

Alternative splicing of messenger **RNA**

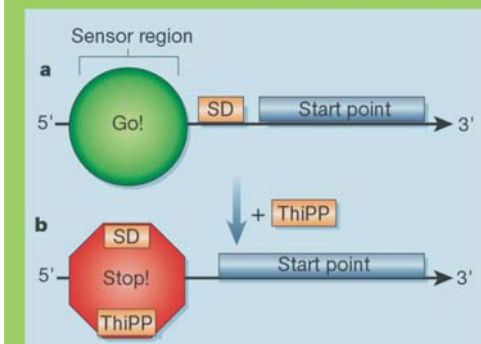
RNA *as carrier of genetic information*

RNA viruses and retroviruses

RNA evolution *in vitro*

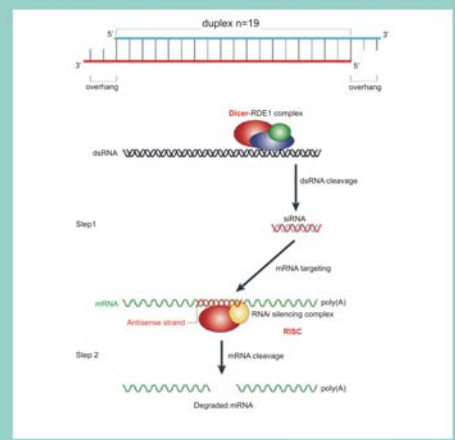
Evolutionary biotechnology
RNA aptamers, artificial ribozymes, allosteric ribozymes

***Allosteric control of transcribed* RNA**



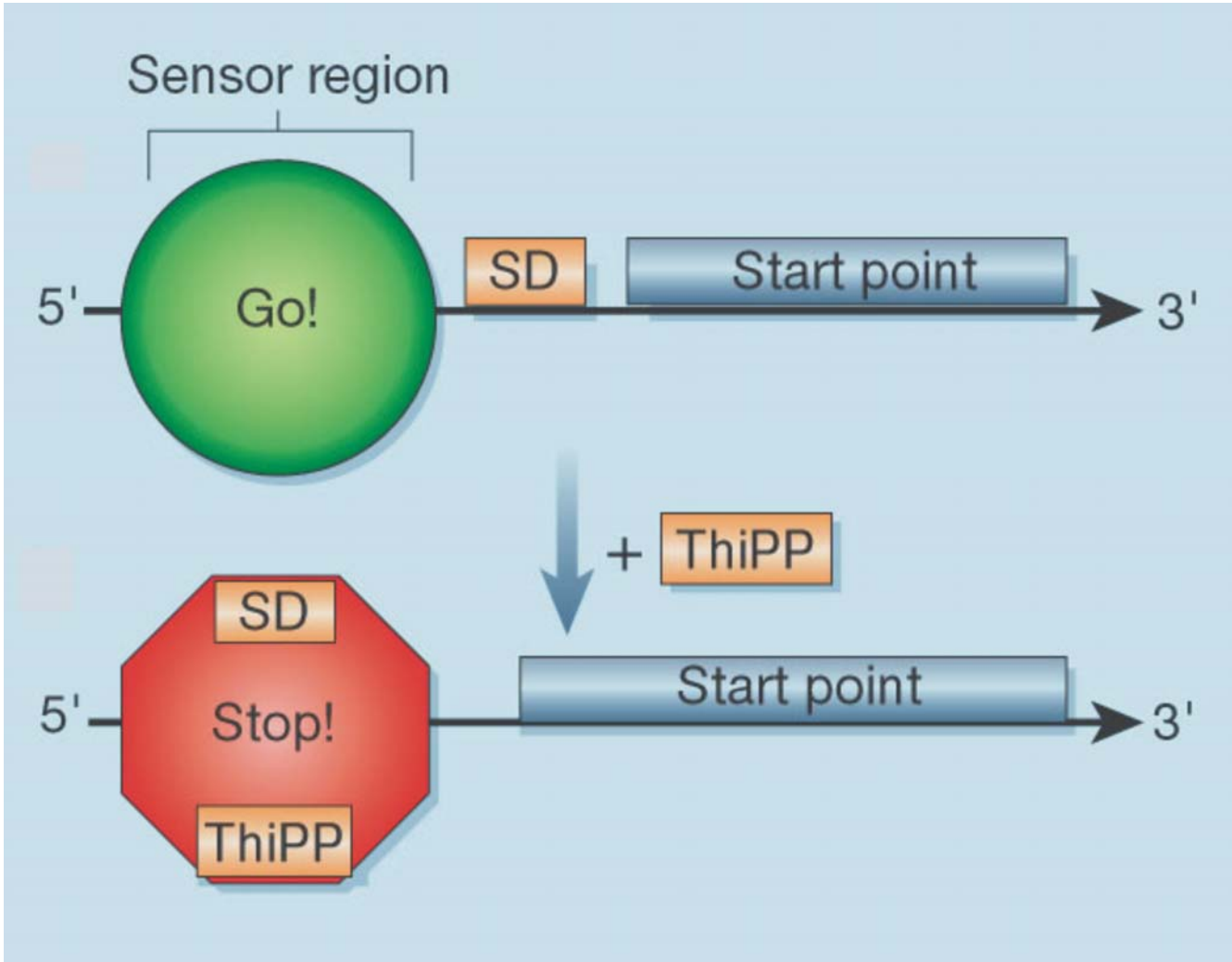
Riboswitches controlled by **metabolites**

RNA *as regulator of gene expression*



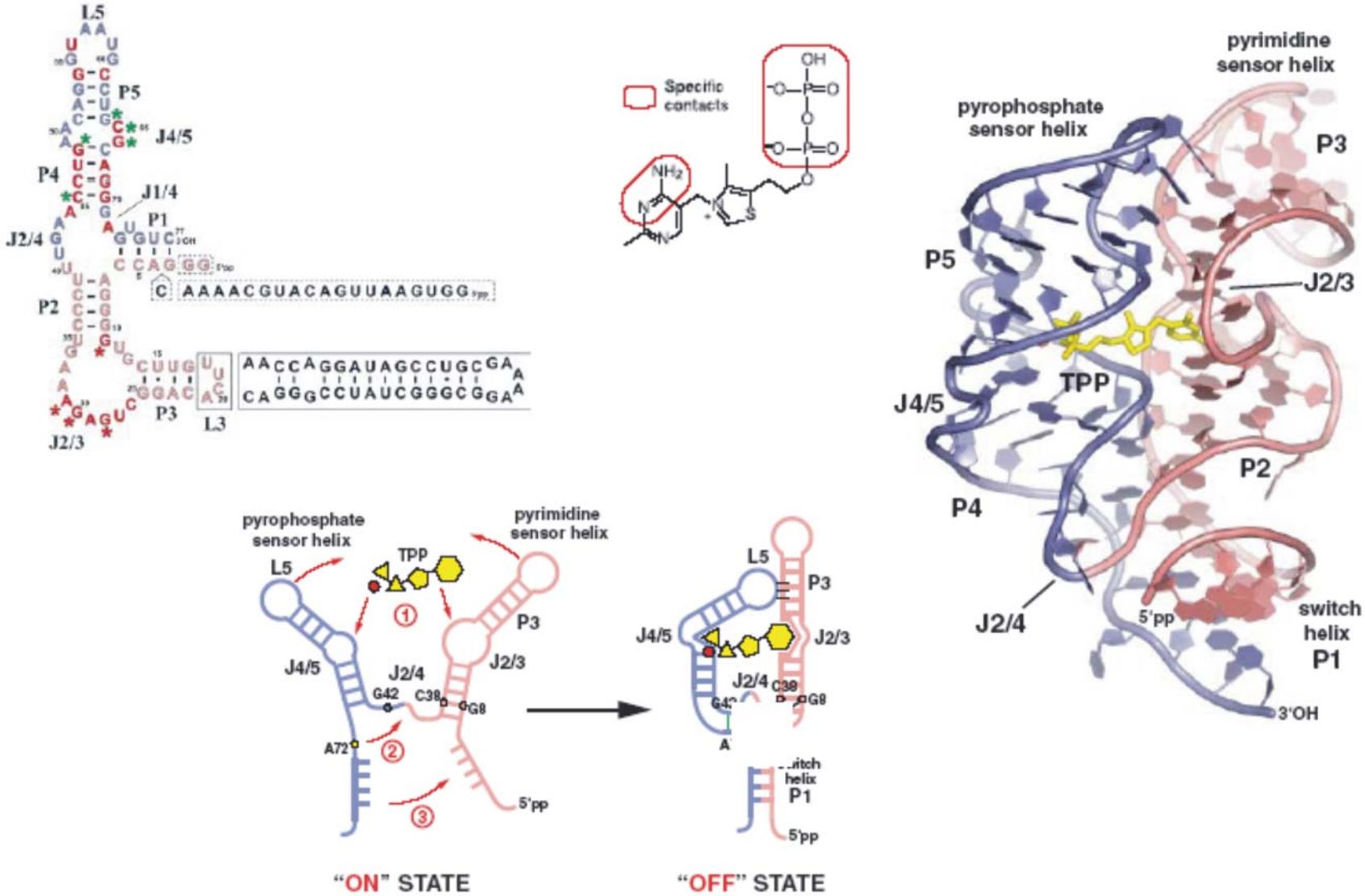
Gene silencing by **siRNA**

Functions of RNA molecules



Jack W. Szostak. RNA gets a grip on translation. *Nature* **419**:890-891 (2002)

Wade Winkler, Ali Nahvi, Ronald R. Breaker. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**:952-956 (2002)



Stéphan Thore, Marc Leibundgut, Nenad Ban. Structure of eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312**:1208-1211 (2006)

Joanna Owens.

Riboswitching off bacterial growth.

Nature Reviews /Drug Discovery **6**:23 (2007)

K.F. Blount et al. Antibacterial lysine analogs that target lysine riboswitches. *Nature Chem. Biol.* **3,** December (2006)

Alexey G. Vitreschak, Dimitry A. Rodinov, Andrey A. Mironov, Mikhail S. Gelfand.

Riboswitches: The oldest mechanism for the regulation of gene expression?

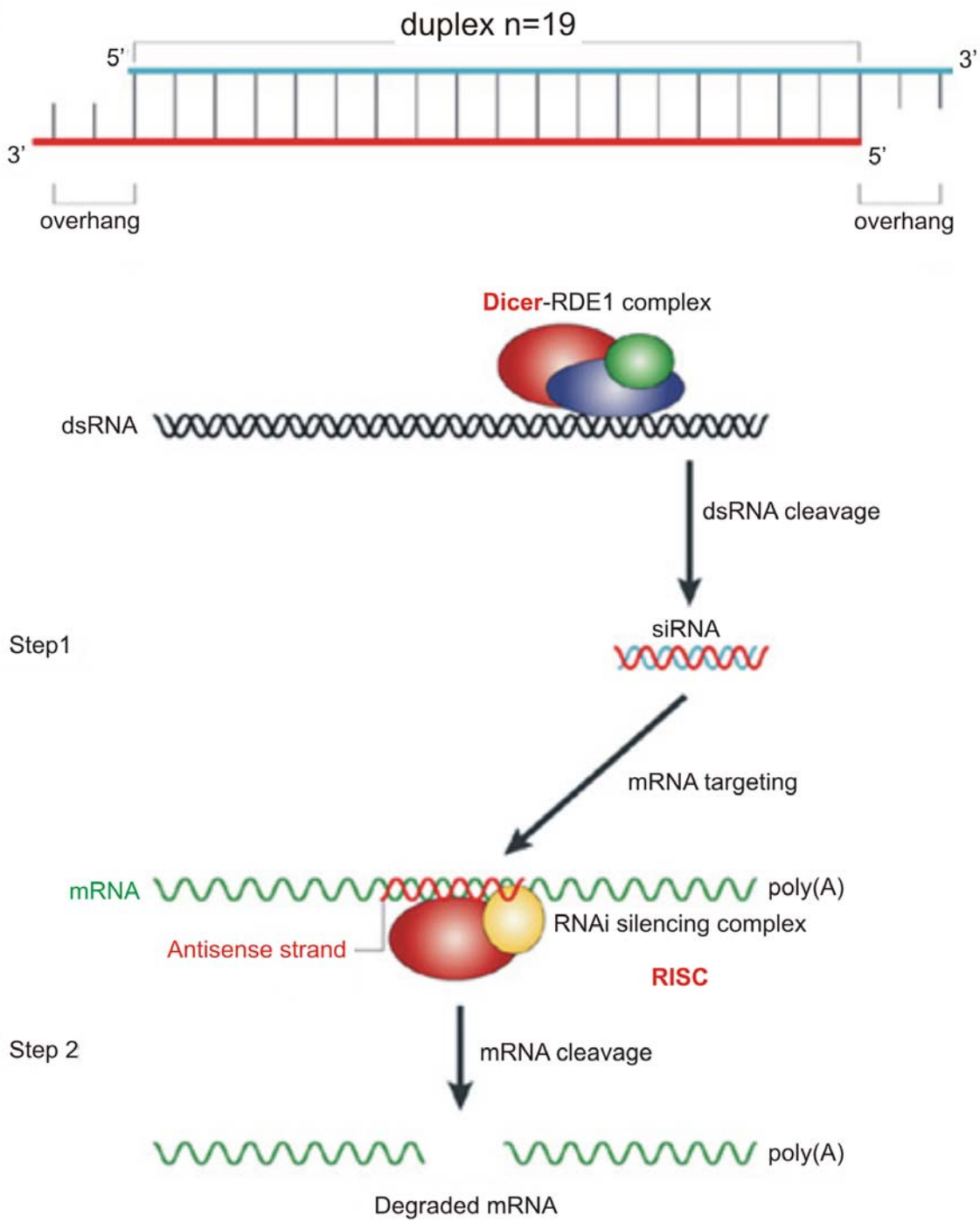
TRENDS in Genetics **20**:44-50 (2004)

Nobel prize for medicine 2006

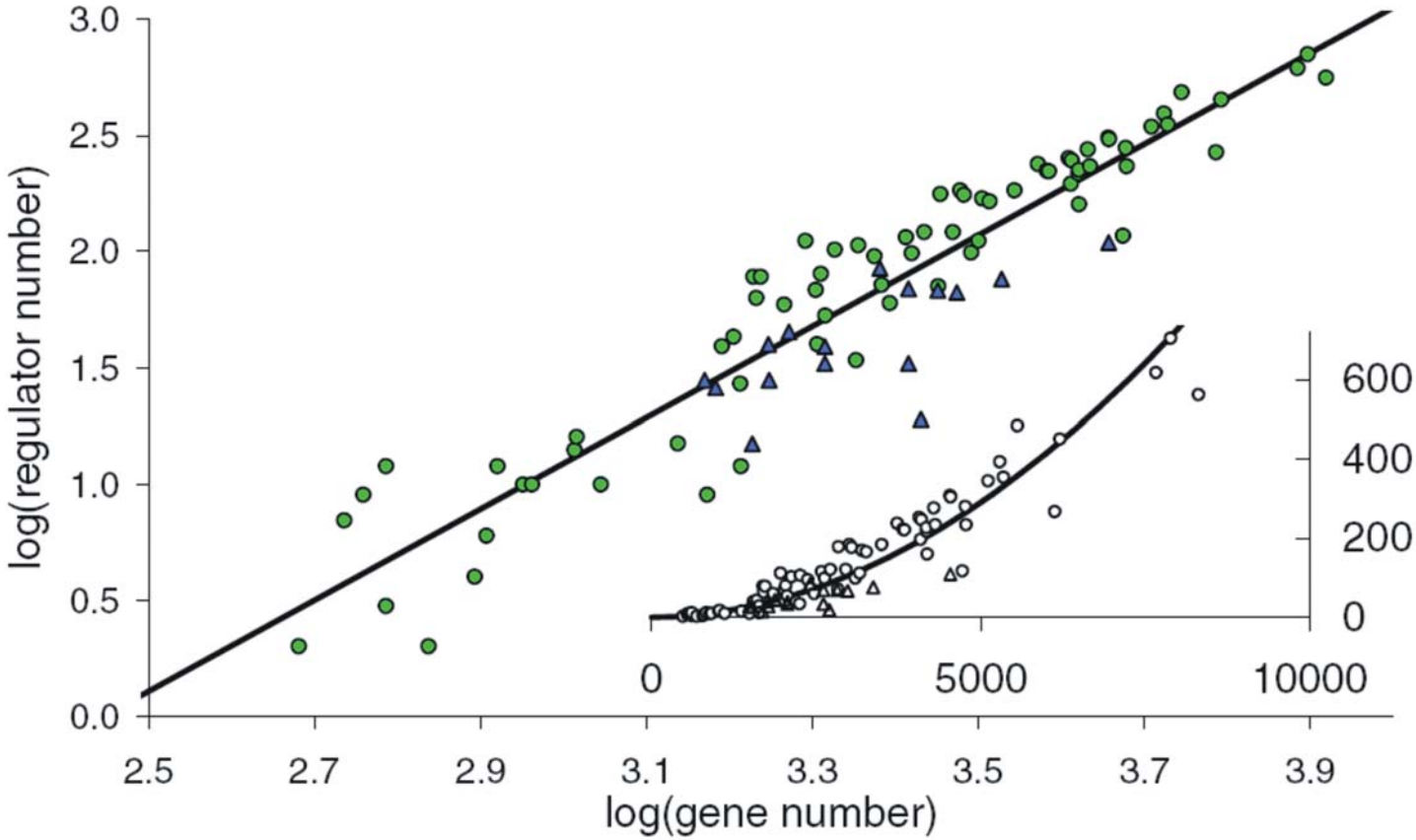
Andrew Z. Fire
Stanford University

Craig C. Mello
University of Massachusetts
Worcester

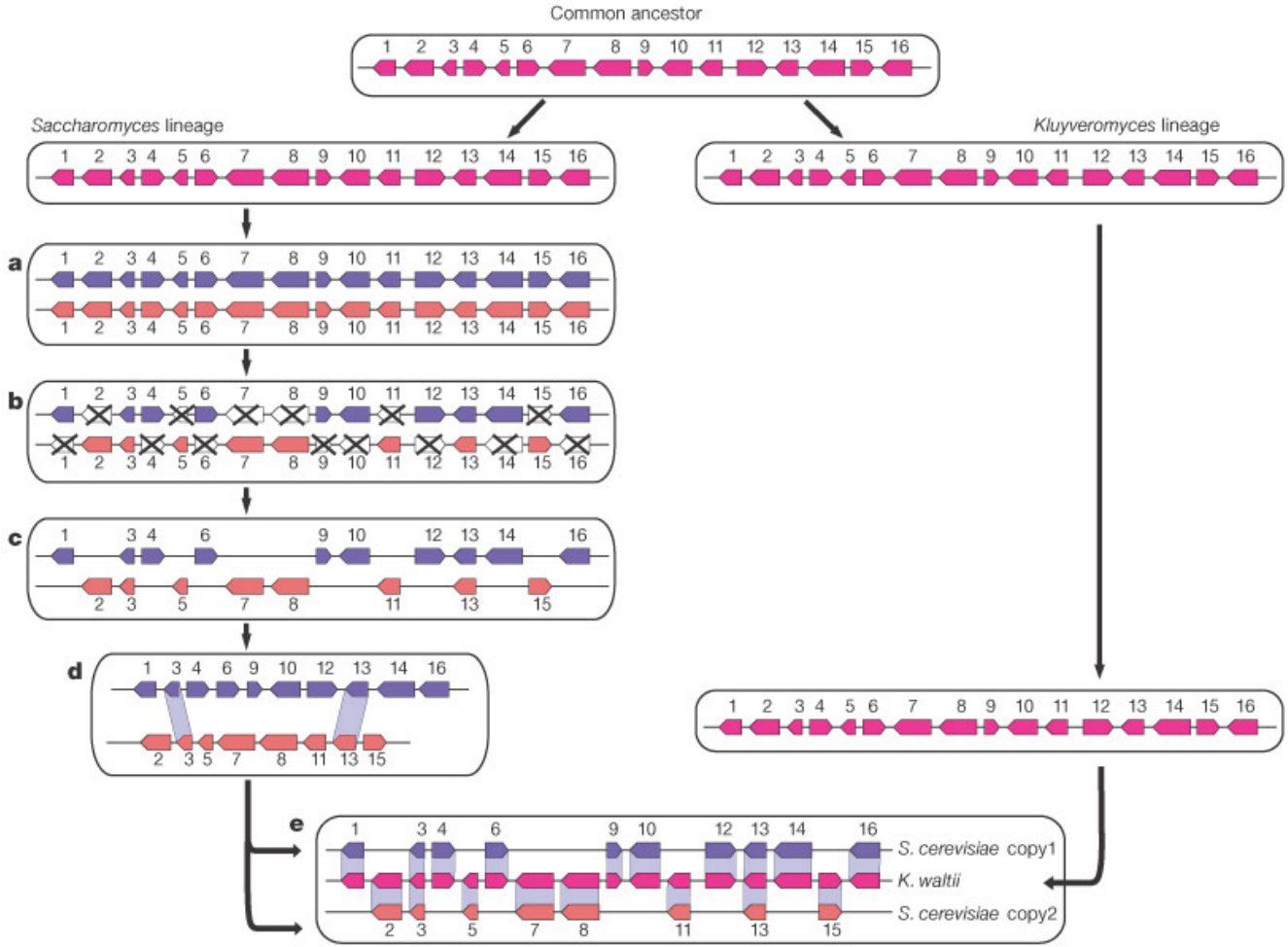
Gene silencing by small interfering RNAs



1. The exciting RNA story
2. **Why is gene regulation so complex?**
3. What small RNAs can achieve
4. Structures of small RNAs
5. Riboswitches and kinetic folding



L.J. Croft, M.J. Lercher, M.J. Gagen, J.S. Mattick. Is prokaryotic complexity limited by accelerated growth in regulatory overhead? *Genome Biology* **5**:P2 (2003)



A model for genome duplication in yeast
≈1×10^8 years ago

2 new genes out of 16 genes,

sequence of genes largely modified

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

1. The exciting RNA story
2. Why is gene regulation so complex?
3. **What small RNAs can achieve**
4. Structures of small RNAs
5. Riboswitches and kinetic folding

Evolution of RNA molecules based on Qβ phage

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

S.Spiegelman, ***An approach to the experimental analysis of precellular evolution***. Quart.Rev.Biophys. **4** (1971), 213-253

C.K.Biebricher, ***Darwinian selection of self-replicating RNA molecules***. Evolutionary Biology **16** (1983), 1-52

G.Bauer, H.Otten, J.S.McCaskill, ***Travelling waves of* in vitro *evolving RNA.*** *Proc.Natl.Acad.Sci.USA* **86** (1989), 7937-7941

C.K.Biebricher, W.C.Gardiner, ***Molecular evolution of RNA* in vitro**. Biophysical Chemistry **66** (1997), 179-192

G.Strunk, T.Ederhof, ***Machines for automated evolution experiments* in vitro *based on the serial transfer concept***. Biophysical Chemistry **66** (1997), 193-202

F.Öhlenschlager, M.Eigen, ***30 years later – A new approach to Sol Spiegelman‘s and Leslie Orgel‘s* in vitro *evolutionary studies***. Orig.Life Evol.Biosph. **27** (1997), 437-457

DIE NATURWISSENSCHAFTEN

58. Jahrgang, 1971 — Heft 10 Oktober

Selforganization of Matter and the Evolution of Biological Macromolecules

MANFRED EIGEN*

Max-Planck-Institut für Biophysikalische Chemie, Karl-Friedrich-Bonhoeffer-Institut, Göttingen-Nikolausberg

I. Introduction

I.1. „Cause and Effect"

The question about the origin of life often appears as a question about "cause and effect". Physical theories of macroscopic processes usually involve answers to such questions, even if a statistical interpretation is given to the relation between "cause" and "effect". It is mainly due to the nature of this question that many scientists believe that our present physics does not offer any obvious explanation for the existence of life, which even in its simplest forms always appears to be associated with complex macroscopic (i.e. multimolecular) systems, such as the living cell.

As a consequence of the exciting discoveries of "molecular biology", a common version of the above question is: *Which came first, the protein or the nucleic acid?* – a modern variant of the old "chicken-and-the-egg" problem. The term "first" is usually meant to define a causal rather than a temporal relationship, and the words "protein" and "nucleic acid" may be substituted by "function" and "information". The question in this form, when applied to the interplay of nucleic acids and proteins as presently encountered in the living cell, leads ad absurdum, because "function"

* Partly presented as the "Robbins Lectures" at Pomona College, California, in spring 1970.

Die Naturwissenschaften 64. Jahrgang Heft 11 November 1977

The Hypercycle

A Principle of Natural Self-Organization

Part A: Emergence of the Hypercycle

Manfred Eigen
Max-Planck-Institut für biophysikalische Chemie, D-3400 Göttingen

Peter Schuster
Institut für theoretische Chemie und Strahlenchemie der Universität, A-1090 Wien

This paper is the first part of a trilogy, which comprises a detailed study of a special type of functional organization and demonstrates its relevance with respect to the origin and evolution of life. Self-replicative macromolecules, such as RNA or DNA in a suitable environment exhibit a behavior, which we may call Darwinian and which can be formally represented by the concept of the quasi-species. A quasi-species is defined as a given distribution of macromolecular species with closely interrelated sequences, dominated by one or several (degenerate) master copies. External constraints enforce the selection of the best adapted distribution, commonly referred to as the wild-type. Most important for Darwinian behavior are the criteria for internal stability of the quasi-species. If these criteria are violated, the information stored in the nucleotide sequence of the master copy will deteriorate irreversibly leading to an error catastrophy. As a consequence, selection and evolution of RNA or DNA molecules is limited with respect to the amount of information that can be stored in a single replicative unit. An analysis of experimental data regarding RNA and DNA replication at various levels of organization reveals, that a sufficient amount of information for the build up of a translation machinery can be gained only via integration of several different replicative units (or reproductive cycles) through functional linkages. A stable functional integration then will raise the system to a new level of organization and thereby enlarge its information capacity considerably. The hypercycle appears to be such a form of organization.

Preview on Part B: The Abstract Hypercycle

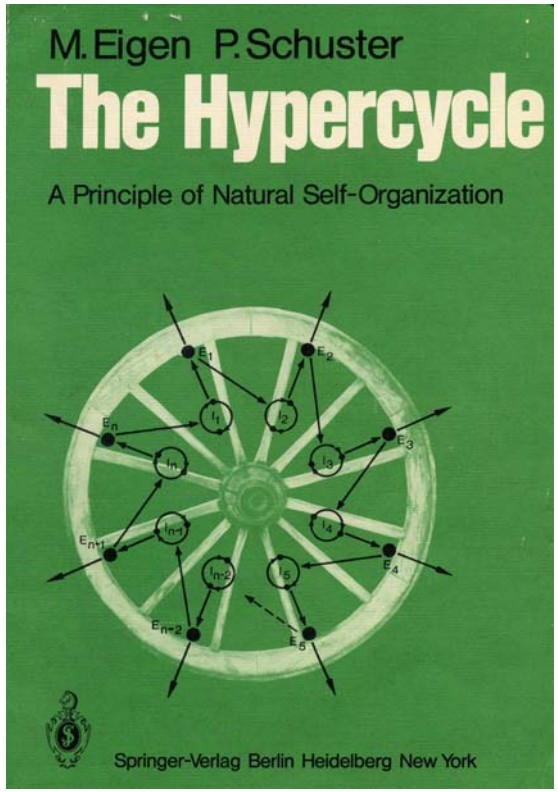
The mathematical analysis of dynamical systems using methods of differential topology, yields the result that there is only one type of mechanisms which fulfills the following requirements: The information stored in each single replicative unit (or reproductive cycle) must be maintained, i.e., the respective master copies must compete favorably with their error distributions. Despite their competitive behavior these units must establish a cooperation which includes all functionally integrated species. On the other hand, the cycle as a whole must continue to compete strongly with any other single entity or linked ensemble which does not contribute to its integrated functions.
These requirements are crucial for a selection of the best adapted functionally linked ensemble and its evolutive optimization. Only hypercyclic organizations are able to fulfil these requirements. Non-cyclic linkages among the autonomous reproduction cycles, such as chains or branched, tree-like networks are devoid of such properties.
The mathematical methods used for proving these assertions are fixed-point, Lyapunov- and trajectorial analysis in higher-dimensional phase spaces, spanned by the concentration coordinates of the cooperating partners. The self-organizing properties of hypercycles are elucidated, using analytical as well as numerical techniques.

Preview on Part C: The Realistic Hypercycle

A realistic model of a hypercycle relevant with respect to the origin of the genetic code and the translation machinery is presented. It includes the following features referring to natural systems:
1) The hypercycle has a sufficiently simple structure to admit an origination with finite probability under prebiotic conditions.
2) It permits a continuous emergence from closely interrelated (t-RNA-like) precursors, originally being members of a stable RNA quasi-species and having been amplified to a level of higher abundance.
3) The organizational structure and the properties of single functional units of this hypercycle are still reflected in the present genetic code in the translation apparatus of the prokaryotic cell, as well as in certain bacterial viruses.

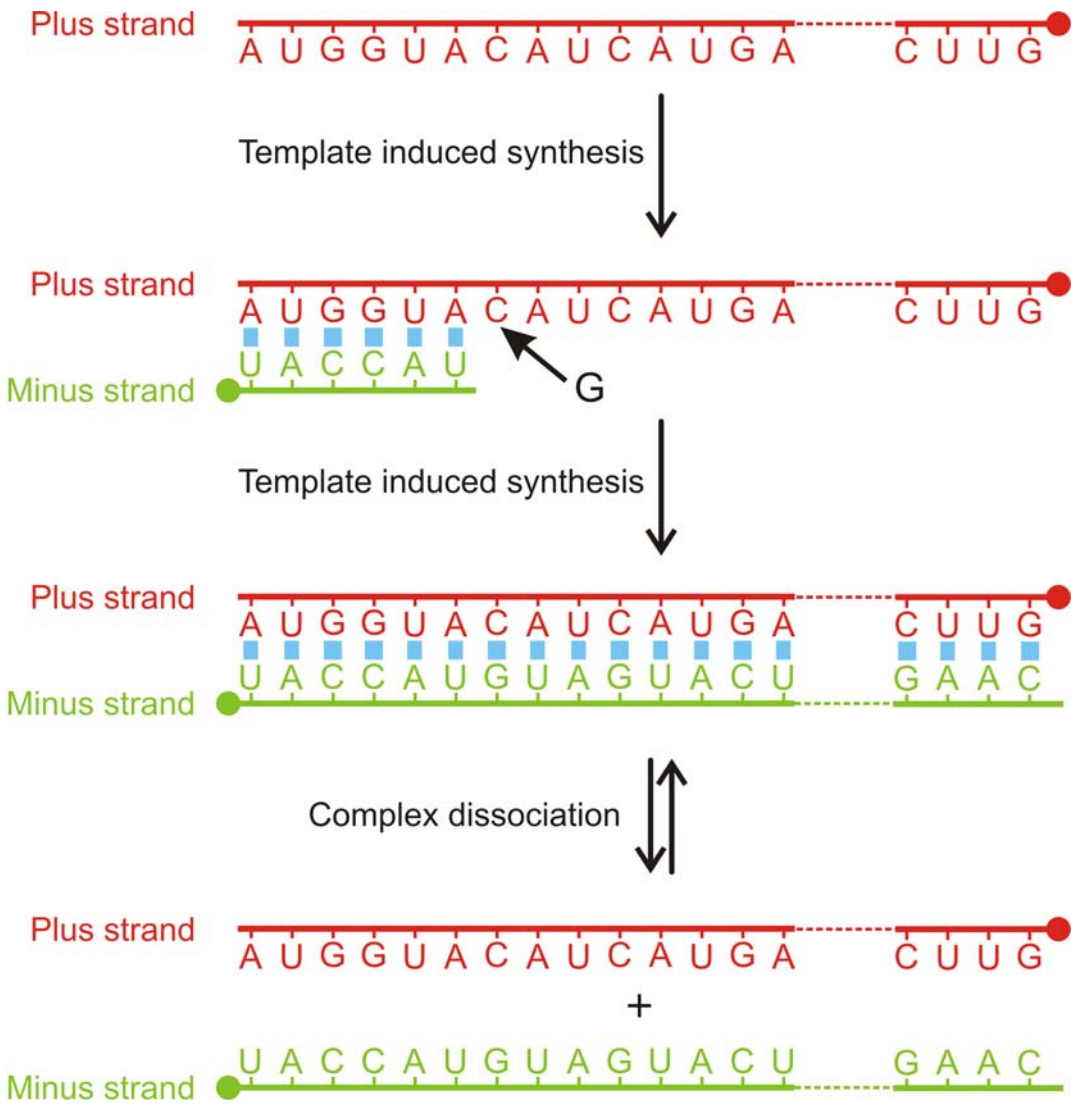
I. The Paradigm of Unity and Diversity in Evolution

Why do millions of species, plants and animals, exist, while there is only one basic molecular machinery of the cell: one universal genetic code and unique chiralities of the macromolecules?
The geneticists of our day would not hesitate to give an immediate answere to the first part of this question. Diversity of species is the outcome of the tremendous branching process of evolution with its myriads of single steps of reproduction and mutation. It in-



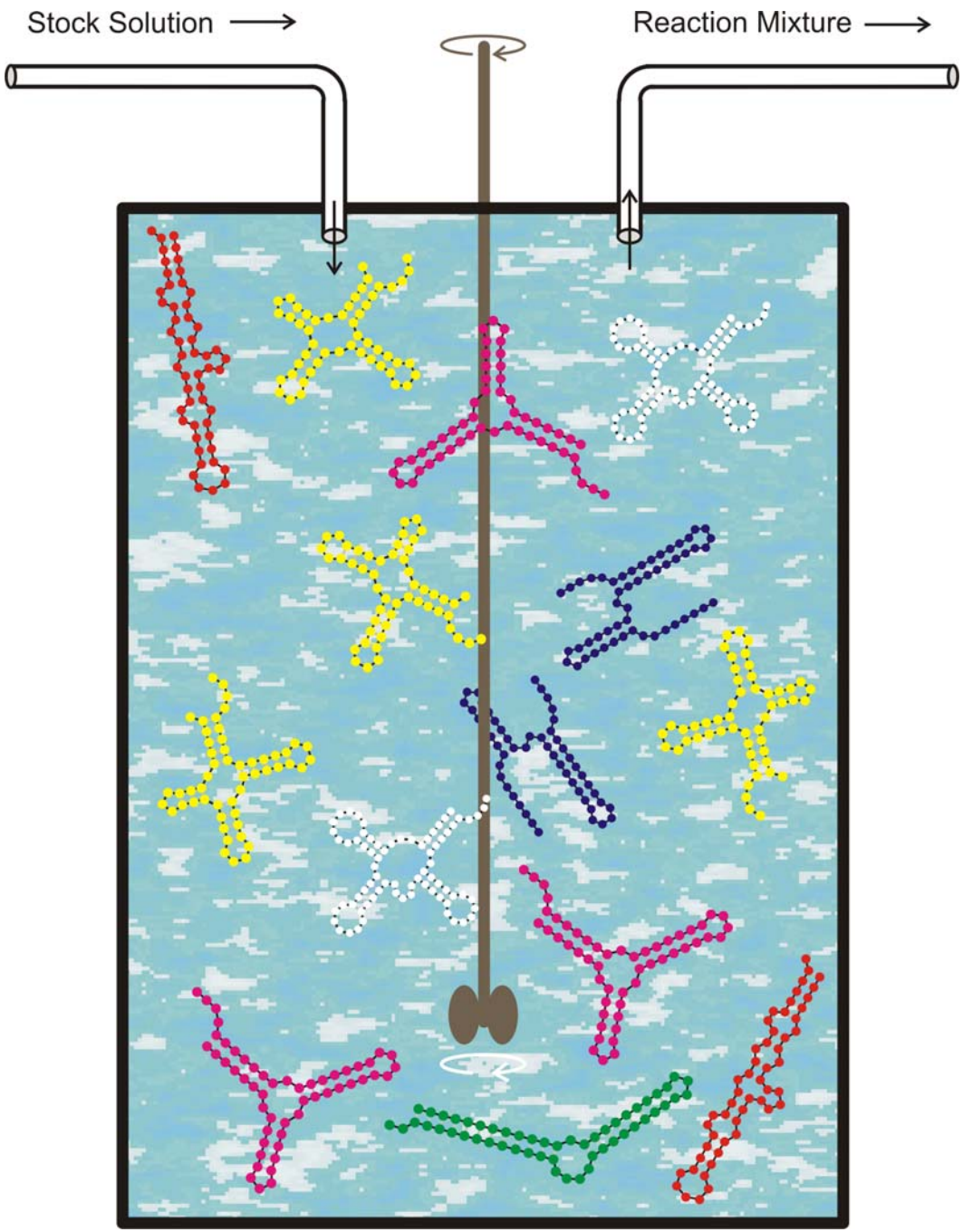
Chemical kinetics of molecular evolution

M. Eigen, P. Schuster, `The Hypercycle´, Springer-Verlag, Berlin 1979



Complementary replication is the simplest copying mechanism of RNA. Complementarity is determined by Watson-Crick base pairs:

G≡C and **A=U**



Replication rate constant:

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

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

Selection constraint:

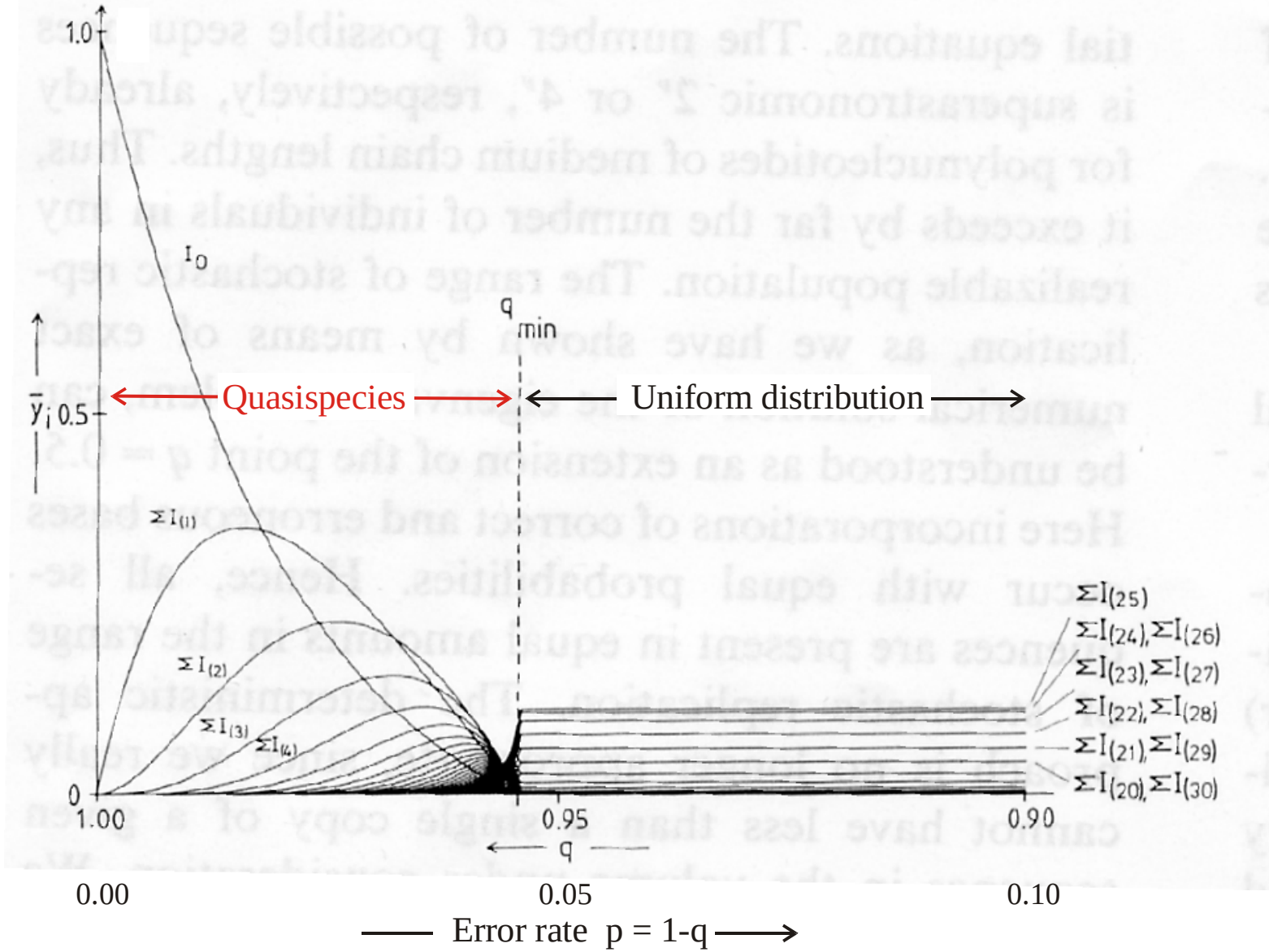
Population size, N = # RNA molecules, is controlled by the flow

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

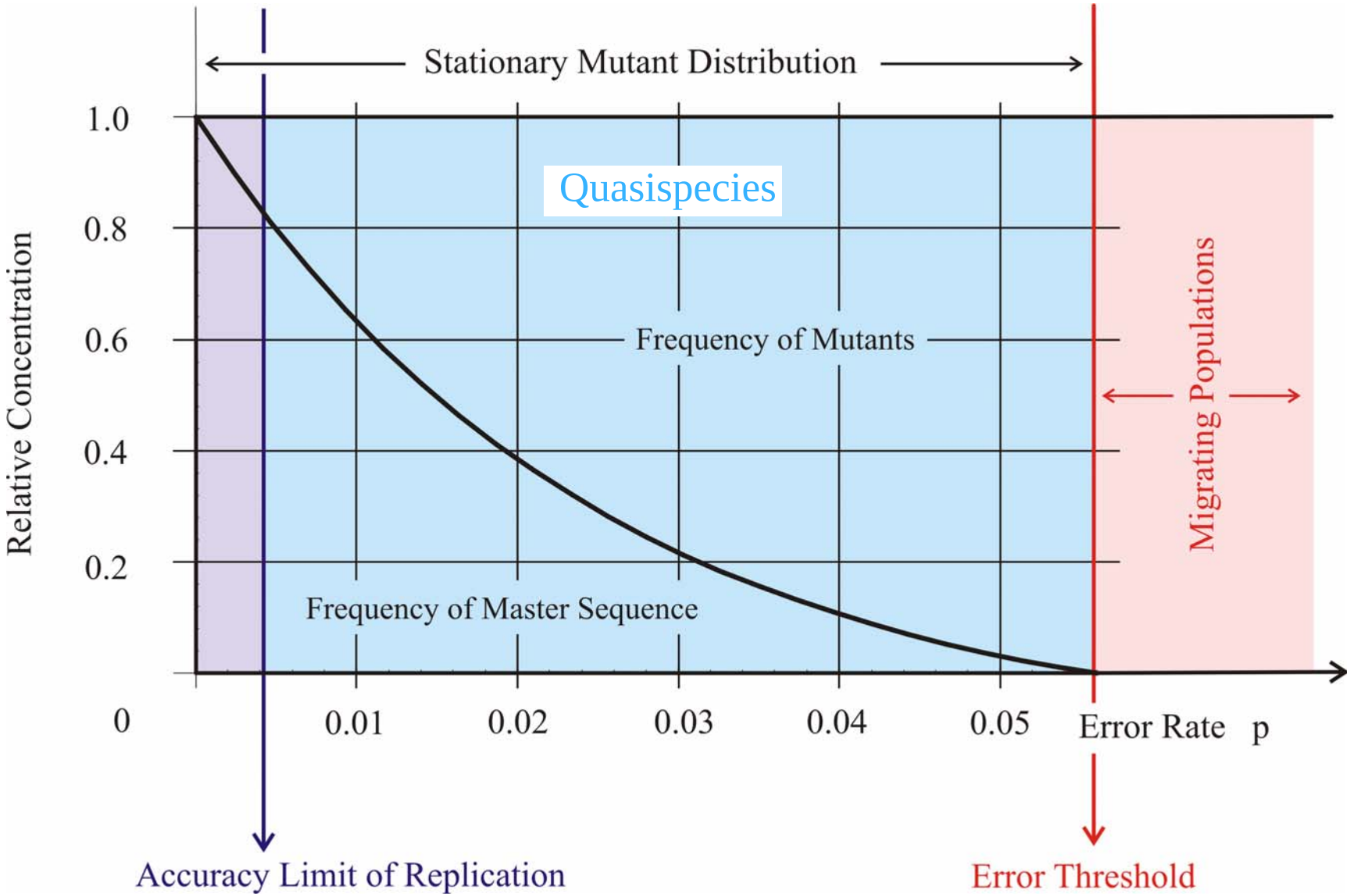
Mutation rate:

p = 0.001 / site × replication

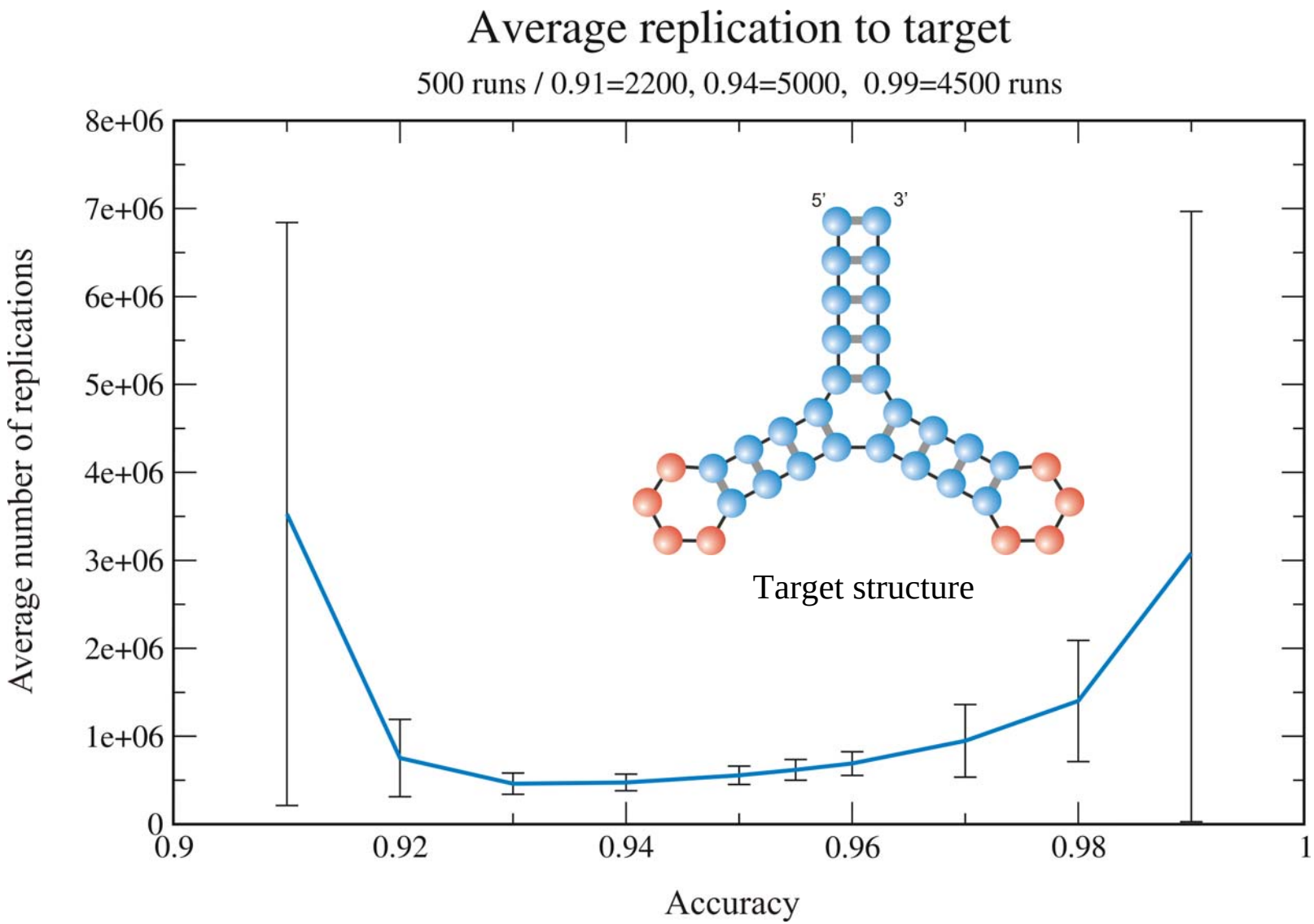
The flowreactor as a device for **studies** of evolution ***in vitro*** and ***in silico***



Quasispecies as a function of the replication accuracy q



The error threshold in replication



Simulation of the approach to a target structure with a population size of N=3000 RNAs

Evolutionary design of RNA molecules

D.B.Bartel, J.W.Szostak, **In vitro *selection of RNA molecules that bind specific ligands***. Nature **346** (1990), 818-822

C.Tuerk, L.Gold, **SELEX - *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage* T4 *DNA polymerase***. Science **249** (1990), 505-510

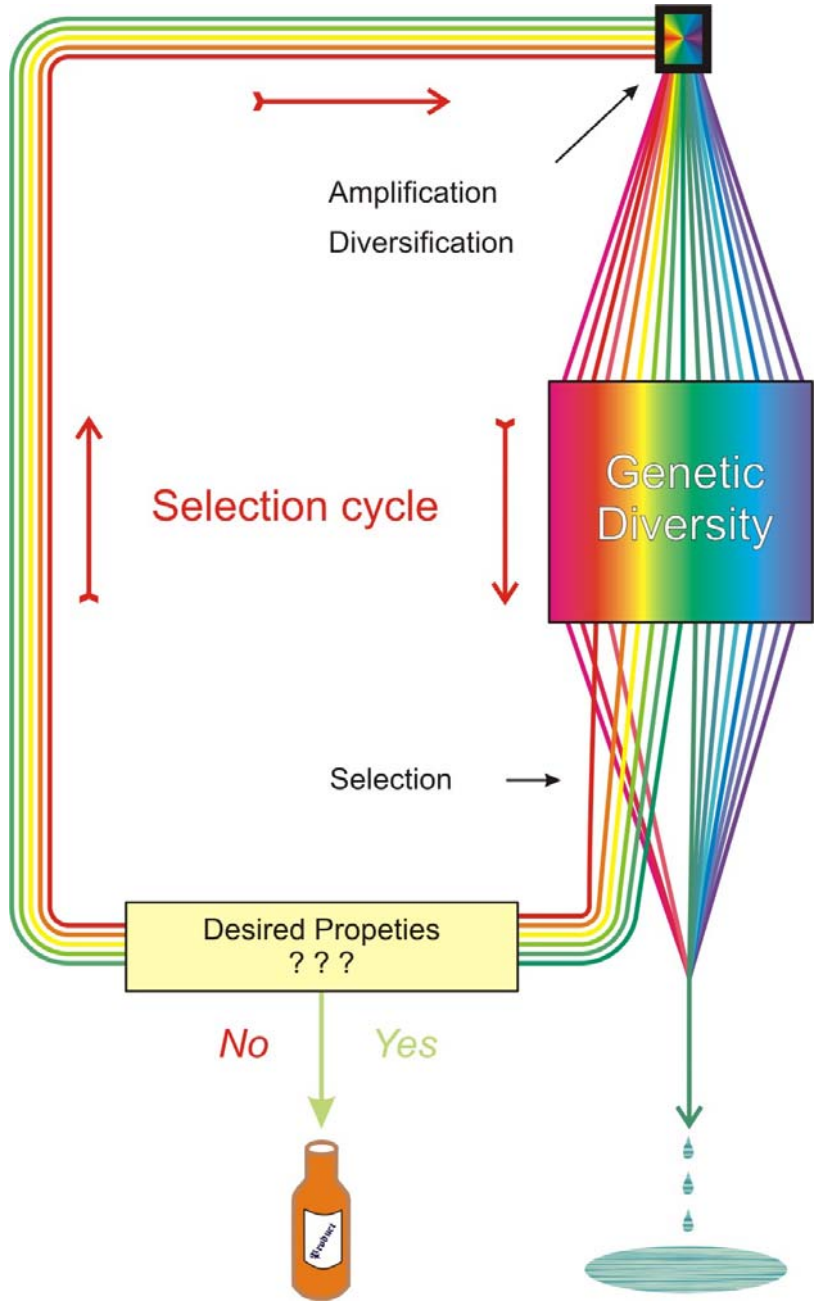
D.P.Bartel, J.W.Szostak, ***Isolation of new ribozymes from a large pool of random sequences***. Science **261** (1993), 1411-1418

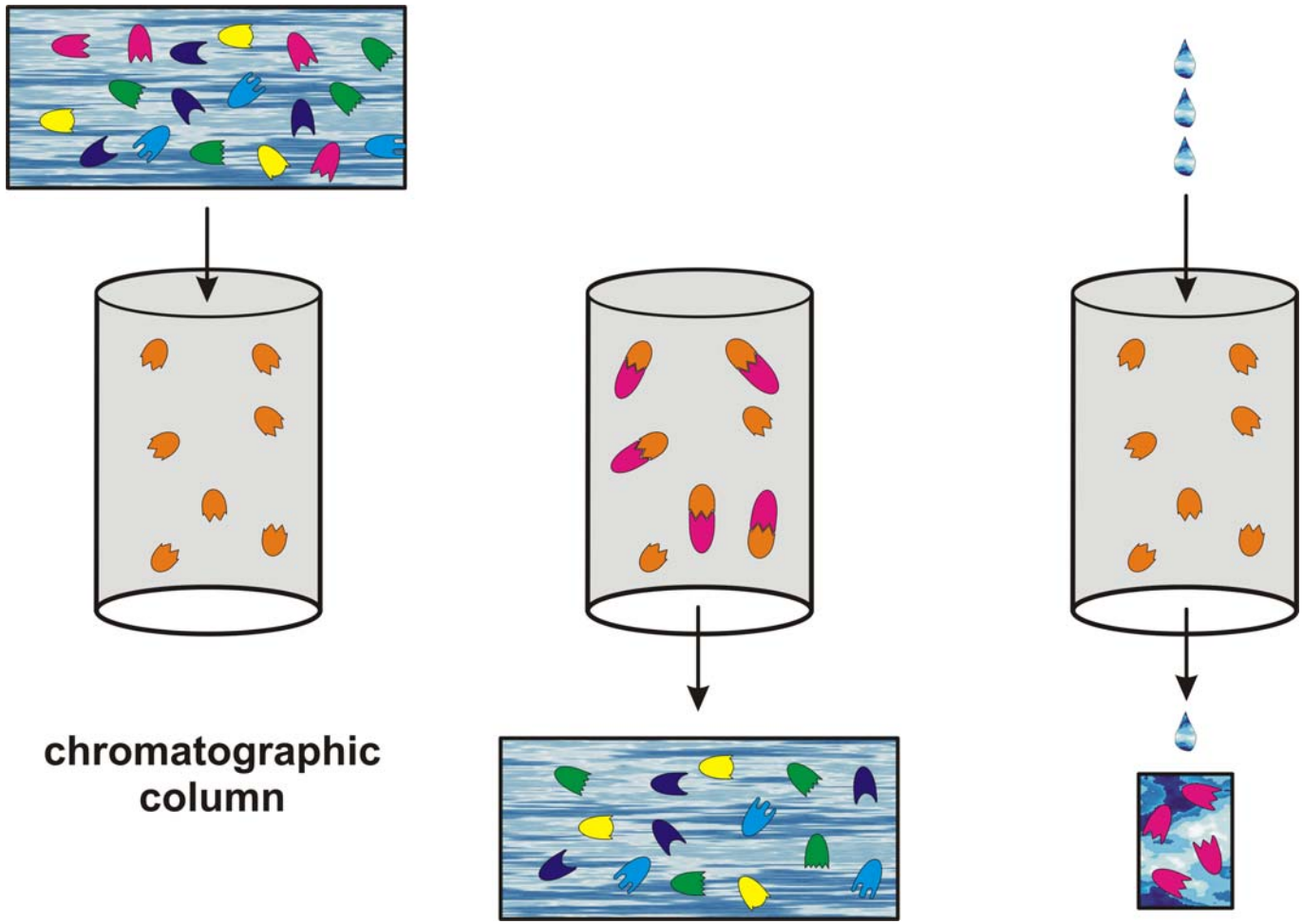
R.D.Jenison, S.C.Gill, A.Pardi, B.Poliski, ***High-resolution molecular discrimination by RNA***. Science **263** (1994), 1425-1429

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

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

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





The SELEX technique for the evolutionary preparation of aptamers

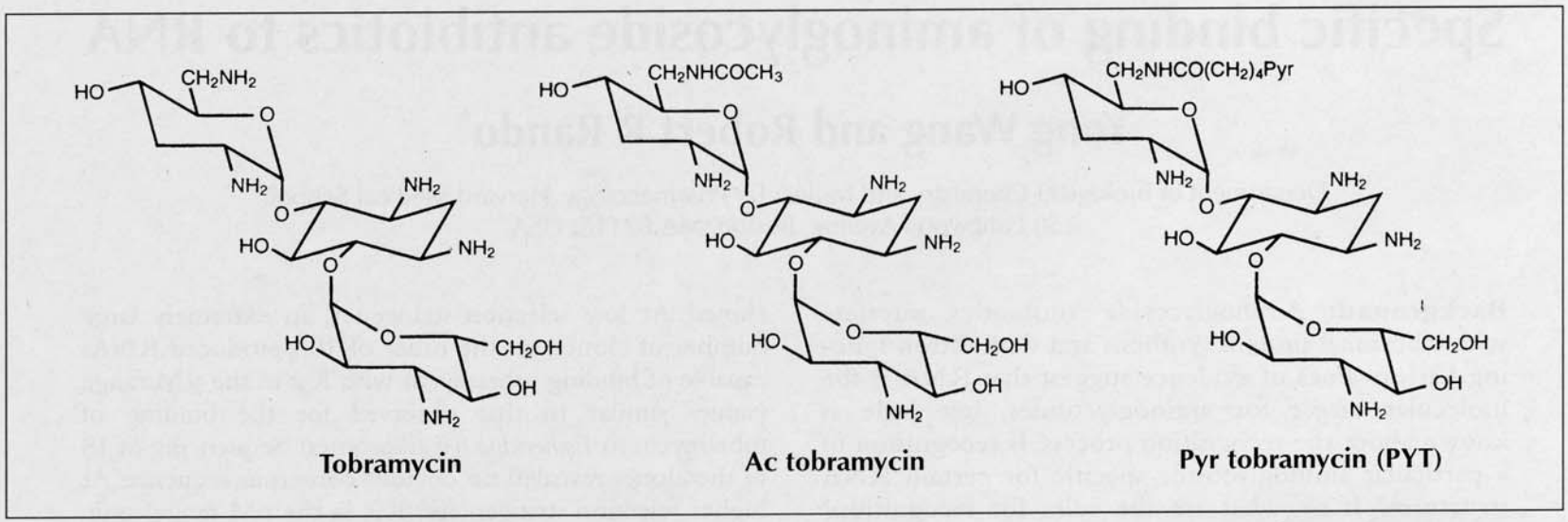
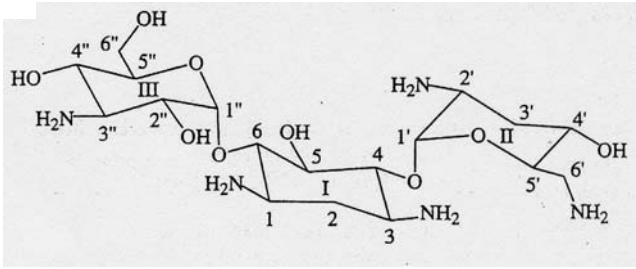


Fig. 1. Structures of tobramycin and analogs used in these studies.

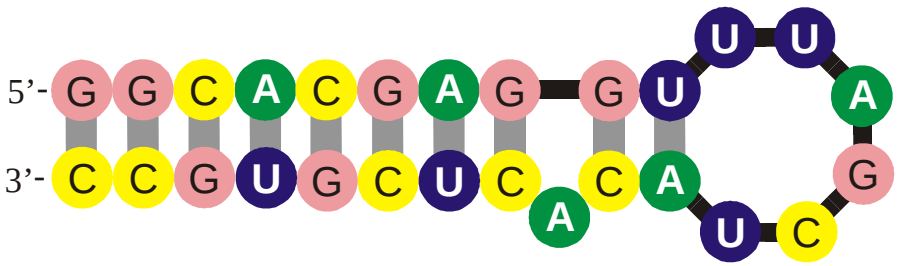
Aptamer binding to aminoglycosid antibiotics: Structure of ligands

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



tobramycin

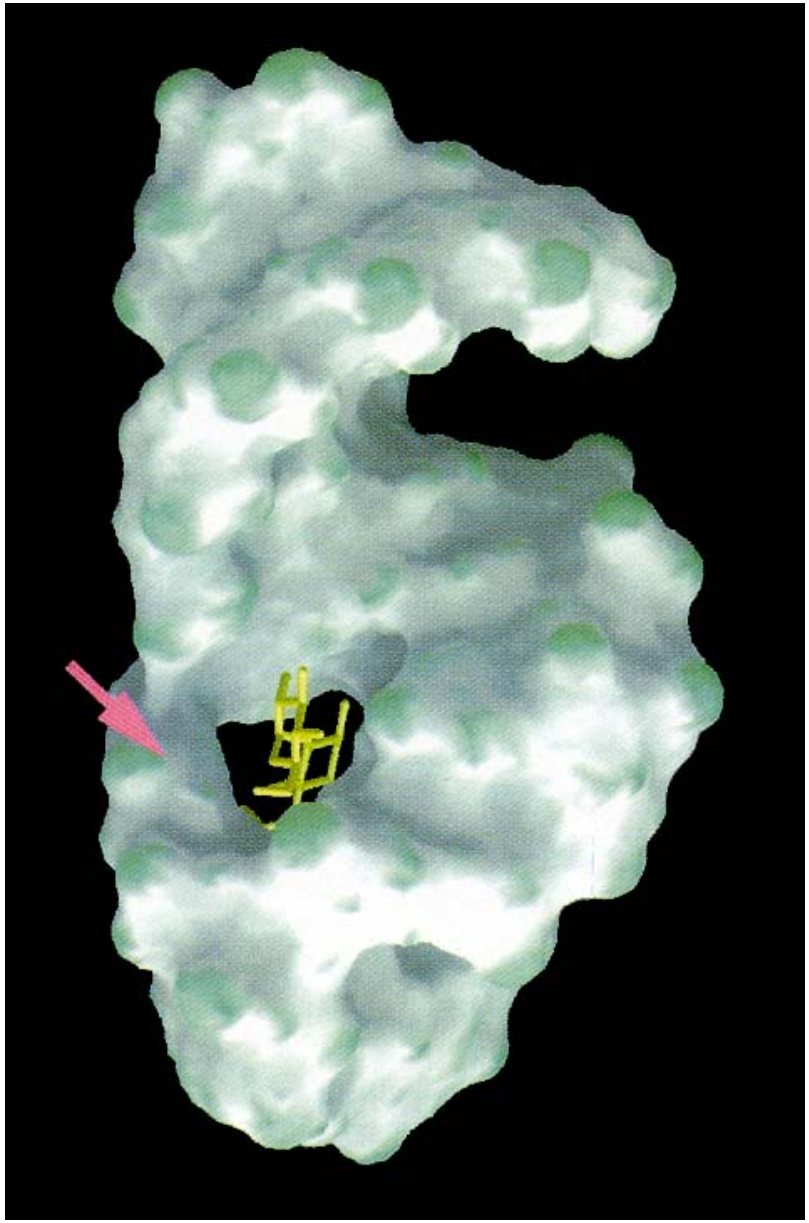
5’- G G C A C G A G G U U U A G C U A C A C U C G U G C C -3’



RNA aptamer

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

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



The three-dimensional structure of the tobramycin aptamer complex

L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, Chemistry & Biology **4**:35-50 (1997)

Hammerhead ribozyme – The smallest RNA based catalyst

H.W.Pley, K.M.Flaherty, D.B.McKay, ***Three dimensional structure of a hammerhead ribozyme***. Nature **372** (1994), 68-74

W.G.Scott, J.T.Finch, A.Klug, ***The crystal structures of an all-RNA hammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage***. Cell **81** (1995), 991-1002

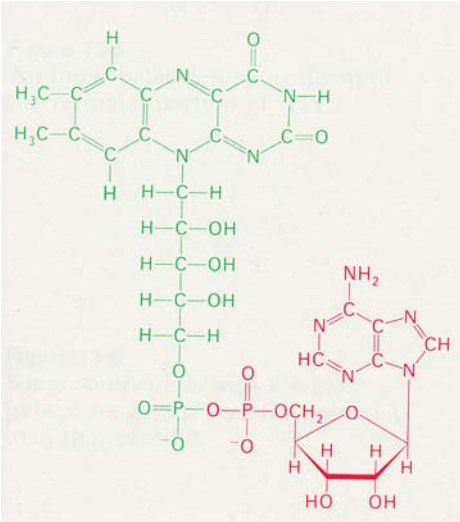
J.E.Wedekind, D.B.McKay, ***Crystallographic structures of the hammerhead ribozyme: Relationship to ribozyme folding and catalysis***. Annu.Rev.Biophys.Biomol.Struct. 27 (1998), 475-502

G.E.Soukup, R.R.Breaker, ***Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization***. Structure **7** (1999), 783-791

Allosteric effectors:

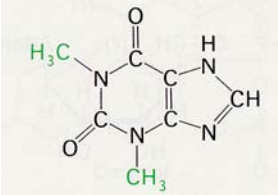
FMN = flavine mononucleotide

H10 – H12



theophylline

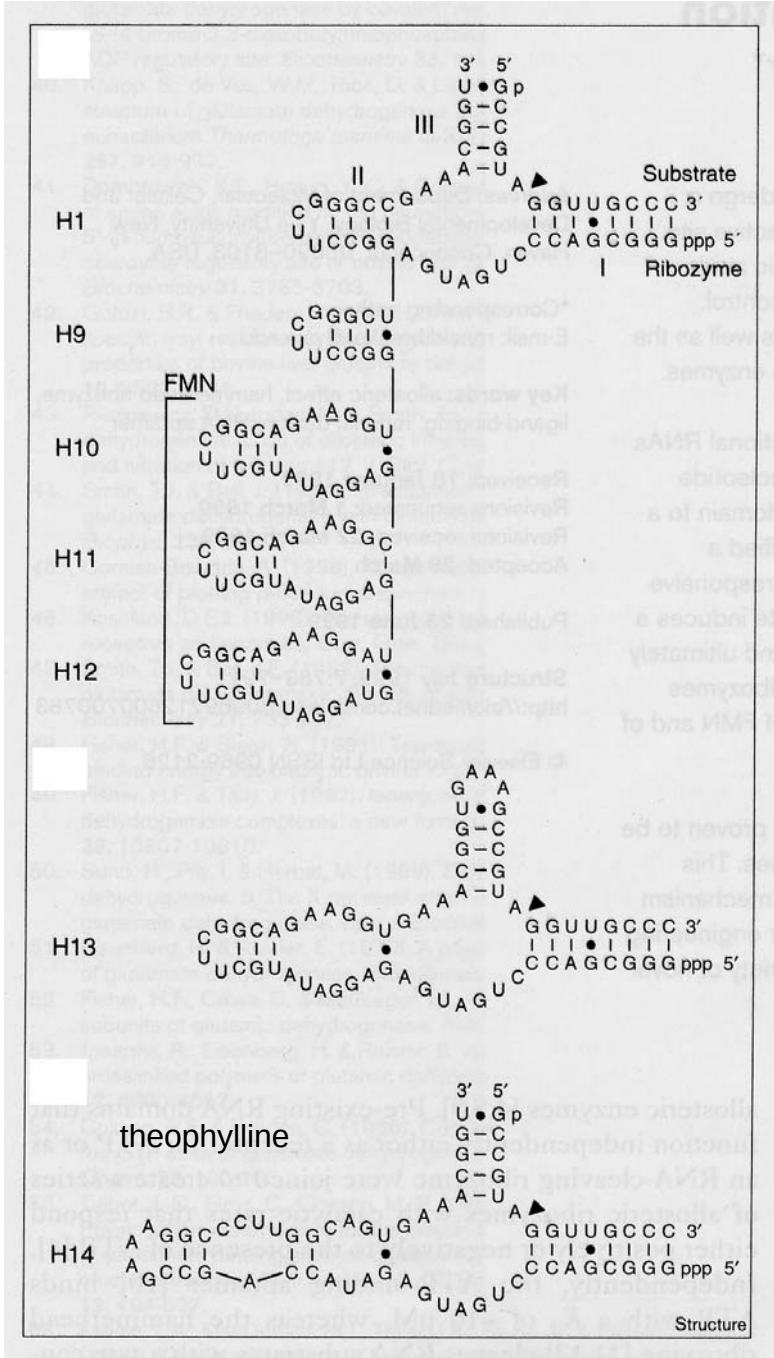
H14



Self-splicing allosteric ribozyme

H13

Hammerhead ribozymes with allosteric effectors



A ribozyme switch

E.A.Schultes, D.B.Bartel, Science **289** (2000), 448-452

minus the background levels observed in the HSP in the control (Sar1-GDP–containing) incubation that prevents COPII vesicle formation. In the microsome control, the level of p115-SNARE associations was less than 0.1%.
46. C. M. Carr, E. Grote, M. Munson, F. M. Hughson, P. J. Novick, *J. Cell Biol.* **146**, 333 (1999).
47. C. Ungermann, B. J. Nichols, H. R. Pelham, W. Wickner, *J. Cell Biol.* **140**, 61 (1998).
48. E. Grote and P. J. Novick, *Mol. Biol. Cell* **10**, 4149 (1999).
49. P. Uetz *et al.*, *Nature* **403**, 623 (2000).
50. GST-SNARE proteins were expressed in bacteria and purified on glutathione-Sepharose beads using standard methods. Immobilized GST-SNARE protein (0.5 μM) was incubated with rat liver cytosol (20 mg) or purified recombinant p115 (0.5 μM) in 1 ml of NS buffer containing 1% BSA for 2 hours at 4°C with rotation. Beads were briefly spun (3000 rpm for 10 s) and sequentially washed three times with NS buffer and three times with NS buffer supplemented with 150 mM NaCl. Bound proteins were eluted three times in 50 μl of 50 mM tris-HCl (pH 8.5), 50 mM reduced glutathione, 150 mM NaCl, and 0.1% Triton X-100 for 15 min at 4°C with intermittent mixing, and elutes were pooled. Proteins were precipitated by MeOH/CH_3Cl and separated by SDS–polyacrylamide gel electrophoresis (PAGE) followed by immunoblotting using p115 mAb 13F12.
51. V. Rybin *et al.*, *Nature* **383**, 266 (1996).
52. K. G. Hardwick and H. R. Pelham, *J. Cell Biol.* **119**, 513 (1992).
53. A. P. Newman, M. E. Groesch, S. Ferro-Novick, *EMBO J.* **11**, 3609 (1992).
54. A. Spang and R. Schekman, *J. Cell Biol.* **143**, 589 (1998).
55. M. F. Rexach, M. Latterich, R. W. Schekman, *J. Cell Biol.* **126**, 1133 (1994).
56. A. Mayer and W. Wickner, *J. Cell Biol.* **136**, 307 (1997).
57. M. D. Turner, H. Plutner, W. E. Balch, *J. Biol. Chem.* **272**, 13479 (1997).
58. A. Price, D. Seals, W. Wickner, C. Ungermann, *J. Cell Biol.* **148**, 1231 (2000).
59. X. Cao and C. Barlowe, *J. Cell Biol.* **149**, 55 (2000).
60. G. G. Tall, H. Hama, D. B. DeWald, B. F. Horazdovsky, *Mol. Biol. Cell* **10**, 1873 (1999).
61. C. G. Burd, M. Peterson, C. R. Cowles, S. D. Emr, *Mol. Biol. Cell* **8**, 1089 (1997).
62. M. R. Peterson, C. G. Burd, S. D. Emr, *Curr. Biol.* **9**, 159 (1999).
63. M. G. Waters, D. O. Clary, J. E. Rothman, *J. Cell Biol.* **118**, 1015 (1992).
64. D. M. Walter, K. S. Paul, M. G. Waters, *J. Biol. Chem.* **273**, 29565 (1998).
65. N. Hui *et al.*, *Mol. Biol. Cell* **8**, 1777 (1997).
66. T. E. Kreis, *EMBO J.* **5**, 931 (1986).
67. H. Plutner, H. W. Davidson, J. Saraste, W. E. Balch, *J. Cell Biol.* **119**, 1097 (1992).
68. D. S. Nelson *et al.*, *J. Cell Biol.* **143**, 319 (1998).
69. We thank G. Waters for p115 cDNA and p115 mAbs; G. Warren for p97 and p47 antibodies; R. Scheller for rbet1, membrin, and sec22 cDNAs; H. Plutner for excellent technical assistance; and P. Tan for help during the initial phase of this work. Supported by NIH grants GM 33301 and GM42336 and National Cancer Institute grant CA58689 (W.E.B.), a NIH National Research Service Award (B.D.M.), and a Wellcome Trust International Traveling Fellowship (B.B.A.).

20 March 2000; accepted 22 May 2000

One Sequence, Two Ribozymes: Implications for the Emergence of New Ribozyme Folds

Erik A. Schultes and David P. Bartel*

We describe a single RNA sequence that can assume either of two ribozyme folds and catalyze the two respective reactions. The two ribozyme folds share no evolutionary history and are completely different, with no base pairs (and probably no hydrogen bonds) in common. Minor variants of this sequence are highly active for one or the other reaction, and can be accessed from prototype ribozymes through a series of neutral mutations. Thus, in the course of evolution, new RNA folds could arise from preexisting folds, without the need to carry inactive intermediate sequences. This raises the possibility that biological RNAs having no structural or functional similarity might share a common ancestry. Furthermore, functional and structural divergence might, in some cases, precede rather than follow gene duplication.

Related protein or RNA sequences with the same folded conformation can often perform very different biochemical functions, indicating that new biochemical functions can arise from preexisting folds. But what evolutionary mechanisms give rise to sequences with new macromolecular folds? When considering the origin of new folds, it is useful to picture, among all sequence possibilities, the distribution of sequences with a particular fold and function. This distribution can range very far in sequence space (*1*). For example, only seven nucleotides are strictly conserved among the group I self-splicing introns, yet secondary (and presumably tertiary) structure within the core of the ribozyme is preserved (*2*). Because these disparate isolates have the same fold and function, it is thought that they descended from a common ancestor through a series of mutational variants that were each functional. Hence, sequence heterogeneity among divergent isolates implies the existence of paths through sequence space that have allowed neutral drift from the ancestral sequence to each isolate. The set of all possible neutral paths composes a "neutral network," connecting in sequence space those widely dispersed sequences sharing a particular fold and activity, such that any sequence on the network can potentially access very distant sequences by neutral mutations (*3–5*).

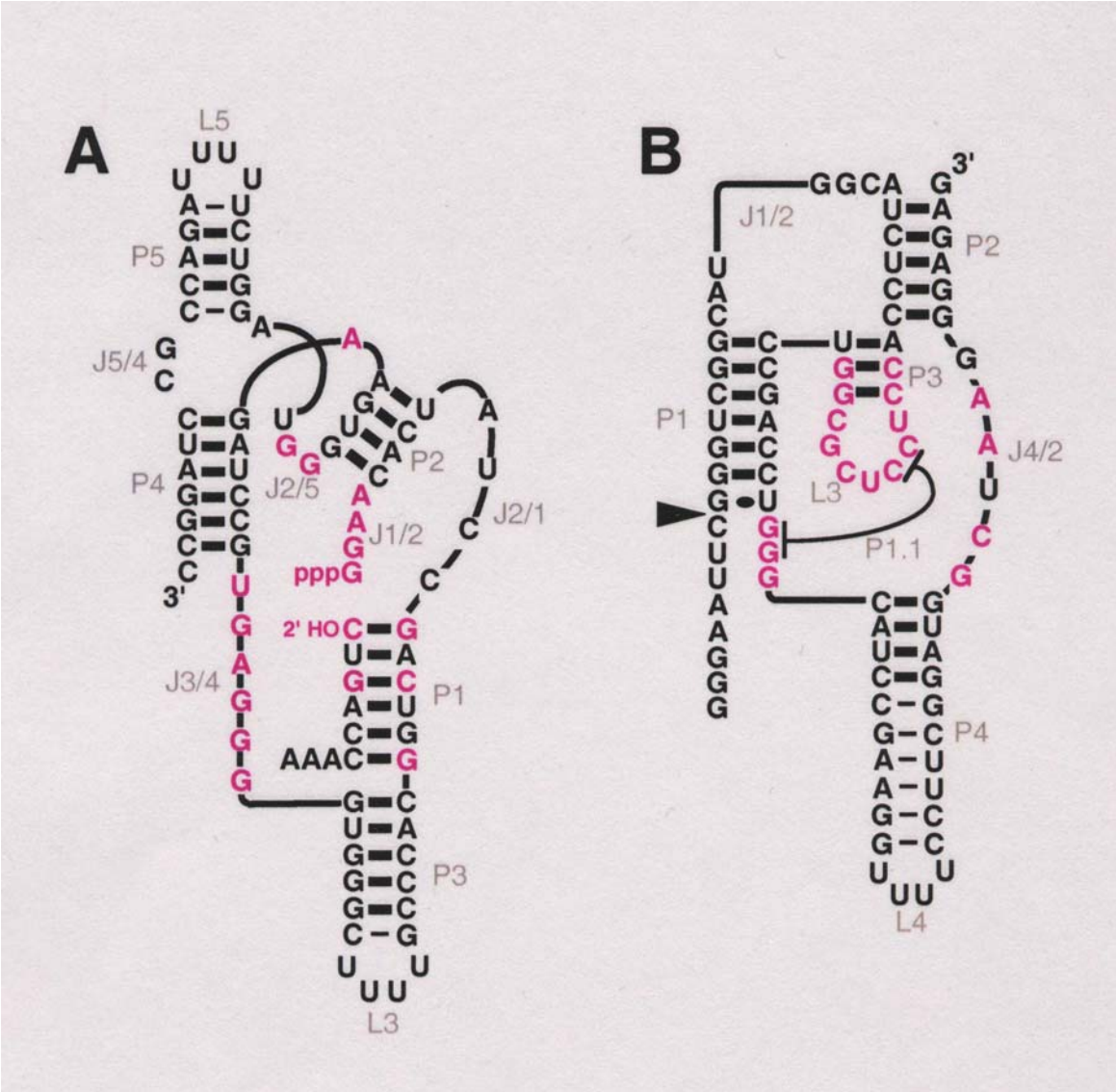
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect. If so, a single sequence would be capable of folding into two different conformations, would have two different catalytic activities, and could access by neutral drift every sequence on both networks. With intersecting networks, RNAs with novel structures and activities could arise from previously existing ribozymes, without the need to carry nonfunctional sequences as evolutionary intermediates. Here, we explore the proximity of neutral networks experimentally, at the level of RNA function. We describe a close apposition of the neutral networks for the hepatitis delta virus (HDV) self-cleaving ribozyme and the class III self-ligating ribozyme.

In choosing the two ribozymes for this investigation, an important criterion was that they share no evolutionary history that might confound the evolutionary interpretations of our results. Choosing at least one artificial ribozyme ensured independent evolutionary histories. The class III ligase is a synthetic ribozyme isolated previously from a pool of random RNA sequences (*9*). It joins an oligonucleotide substrate to its 5′ terminus. The prototype ligase sequence (Fig. 1A) is a shortened version of the most active class III variant isolated after 10 cycles of in vitro selection and evolution. This minimal construct retains the activity of the full-length isolate (*10*). The HDV ribozyme carries out the site-specific self-cleavage reactions needed during the life cycle of HDV, a satellite virus of hepatitis B with a circular, single-stranded RNA genome (*11*). The prototype HDV construct for our study (Fig. 1B) is a shortened version of the antigenomic HDV ribozyme (*12*), which undergoes self-cleavage at a rate similar to that reported for other antigenomic constructs (*13*, *14*).

The prototype class III and HDV ribozymes have no more than the 25% sequence identity expected by chance and no fortuitous structural similarities that might favor an intersection of their two neutral networks. Nevertheless, sequences can be designed that simultaneously satisfy the base-pairing requirements

Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA.

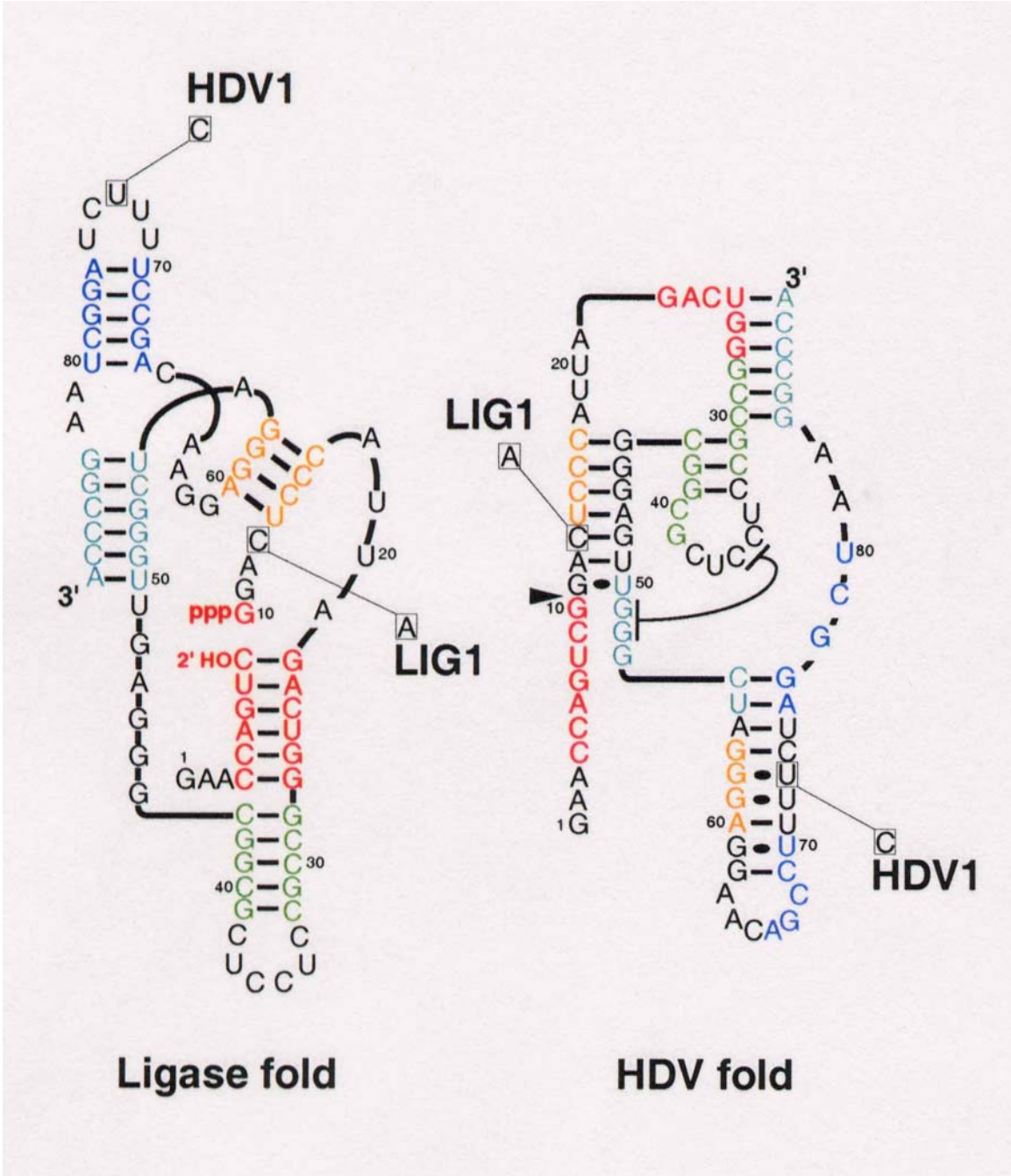
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

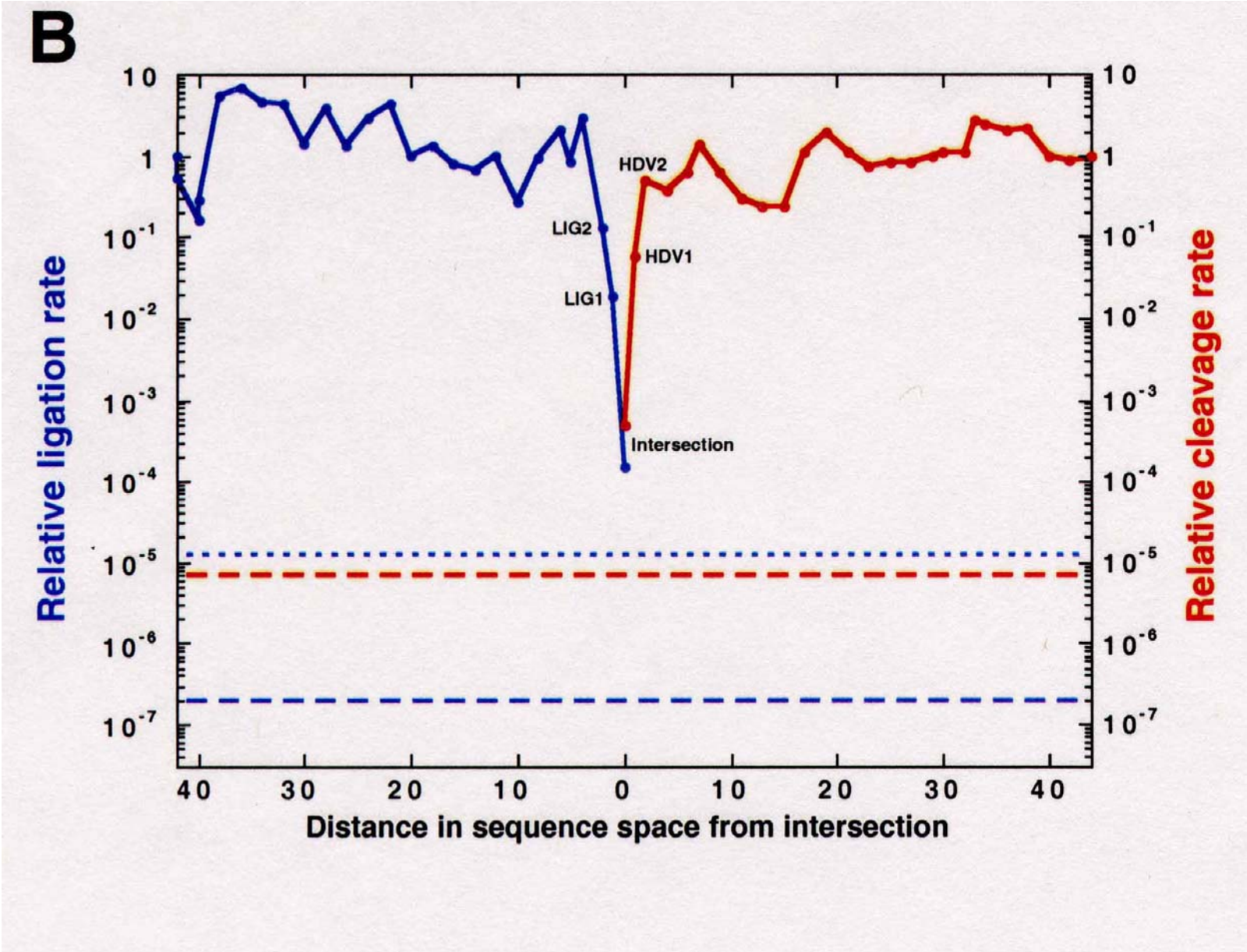


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

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

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





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

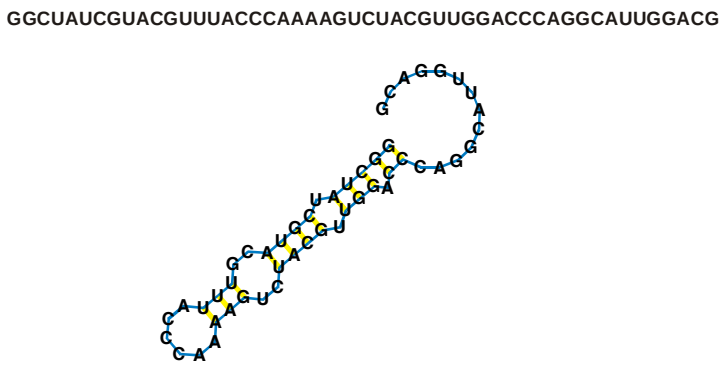
1. The exciting RNA story
2. Why is gene regulation so complex?
3. What small RNAs can achieve
4. **Structures of small RNAs**
5. Riboswitches and kinetic folding

GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG

One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG

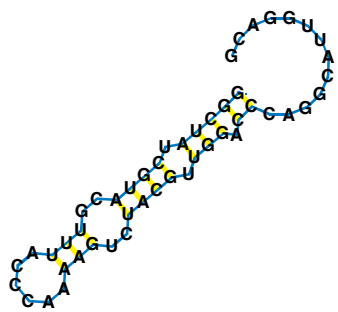
One error neighborhood – Surrounding of an RNA molecule in sequence and shape space



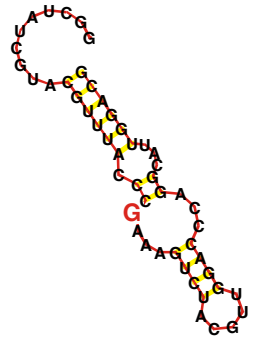
One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

GGCUAUCGUACGUUUACCCGAAAGUCUACGUUGGACCCAGGCAUUGGACG

GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG

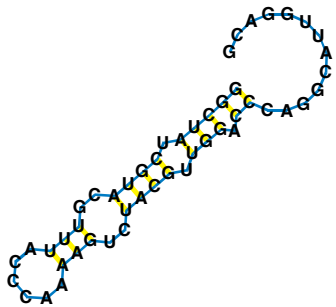


One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

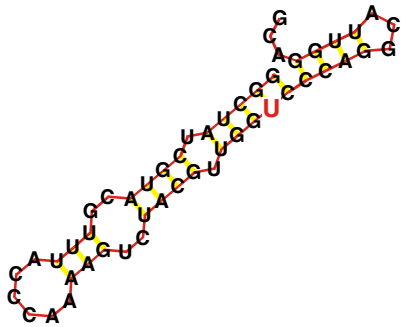


GGCUAUCGUACGUUUACCCGAAAGUCUACGUUGGACCCAGGCAUUGGACG

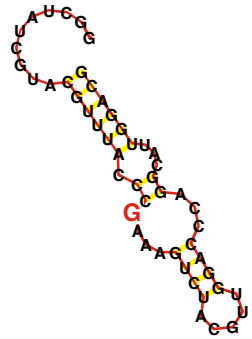
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG



One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

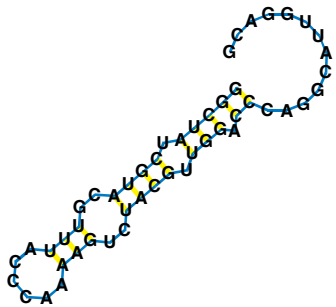


GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGUCCCAGGCAUUGGACG

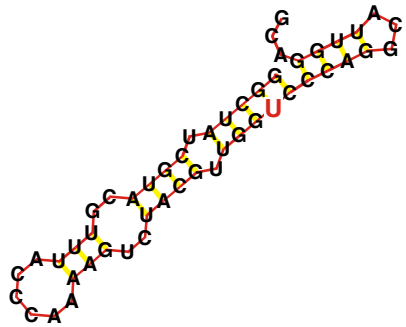


GGCUAUCGUACGUUUACCCGAAAGUCUACGUUGGACCCAGGCAUUGGACG

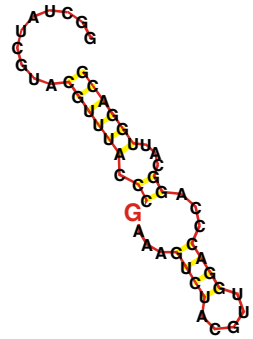
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG



One error neighborhood – Surrounding of an RNA molecule in sequence and shape space



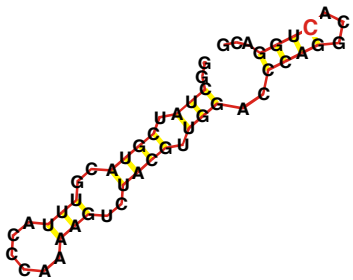
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGUCCCAGGCAUUGGACG

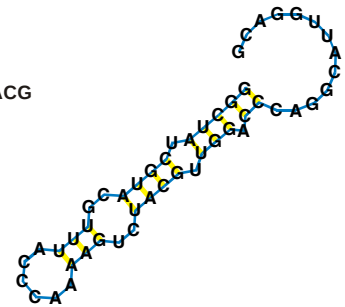


GGCUAUCGUACGUUUACCCGAAAGUCUACGUUGGACCCAGGCAUUGGACG

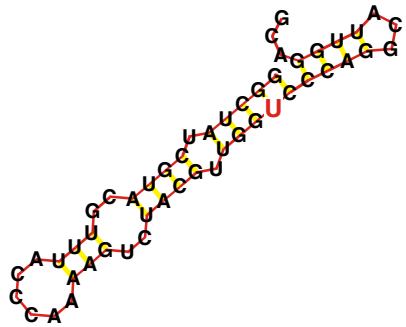
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG

GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCACUGGACG

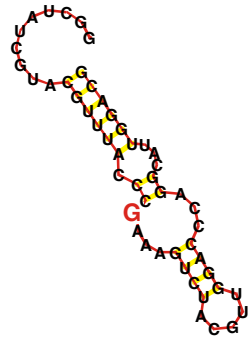




One error neighborhood – Surrounding of an RNA molecule in sequence and shape space



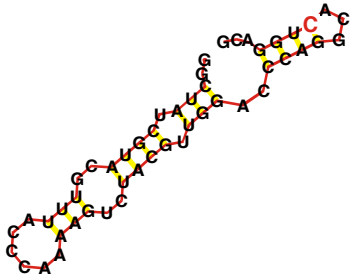
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGUCCCAGGCAUUGGACG

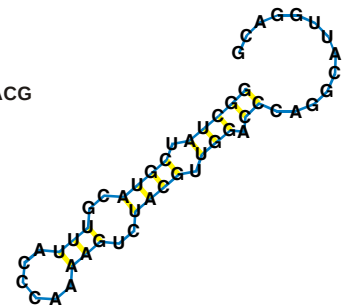


GGCUAUCGUACGUUUACCCGAAAGUCUACGUUGGACCCAGGCAUUGGACG

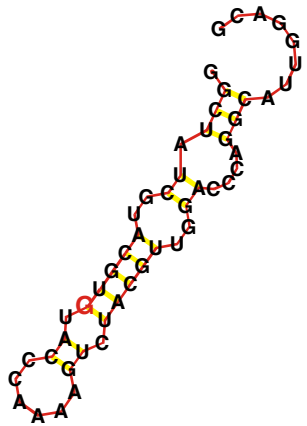
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG

GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCACUGGACG



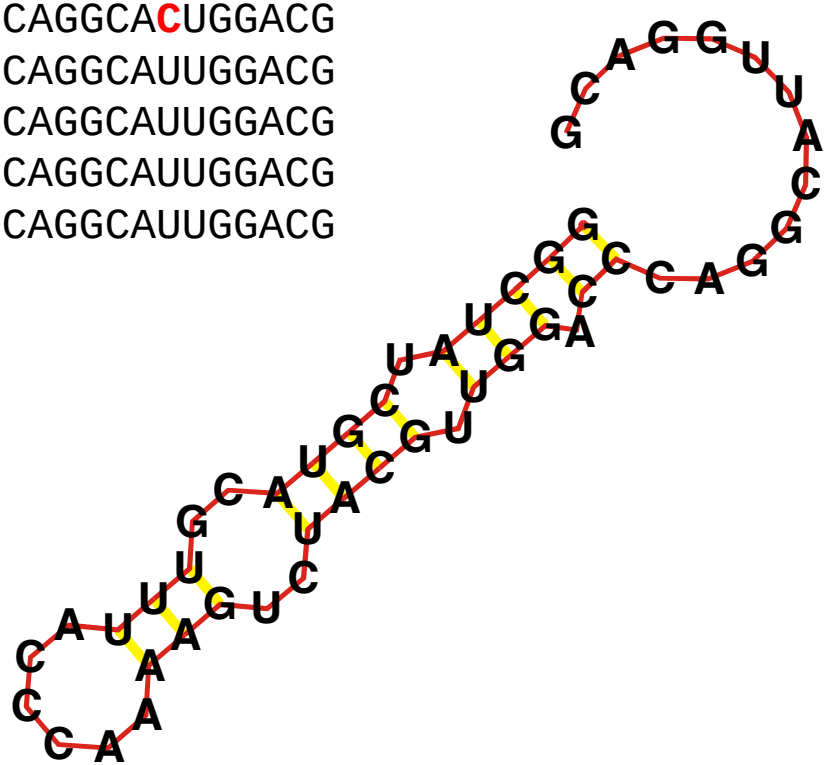


GGCUAUCGUACGUGUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG



One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

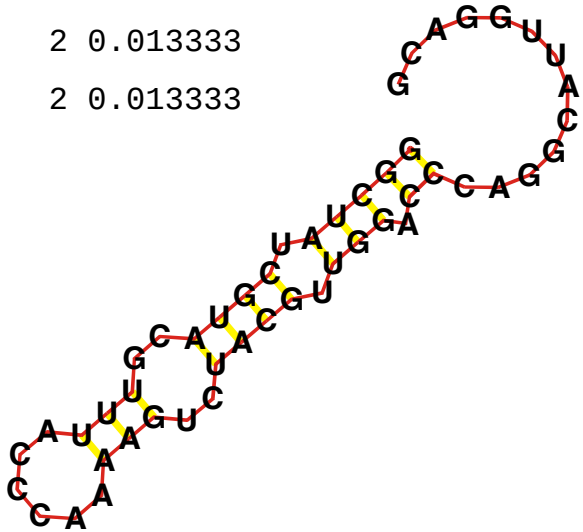
GGCUAUCGUA**U**GUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUU**A**GACG
GGCUAUCGUACGUUUAC**U**CAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACG**C**UUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGC**C**AUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGU**G**UACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUA**A**CGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCC**U**GGCAUUGGACG
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCA**C**UGGACG
GGCUAUCGUACGUUUACCCAAAAGUCUACGUUGG**U**CCCAGGCAUUGGACG
GGCUA**G**CGUACGUUUACCCAAAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGUUUACCC**G**AAAGUCUACGUUGGACCCAGGCAUUGGACG
GGCUAUCGUACGUUUACCCAAAAG**C**CUACGUUGGACCCAGGCAUUGGACG



One error neighborhood – Surrounding of an RNA molecule in sequence and shape space

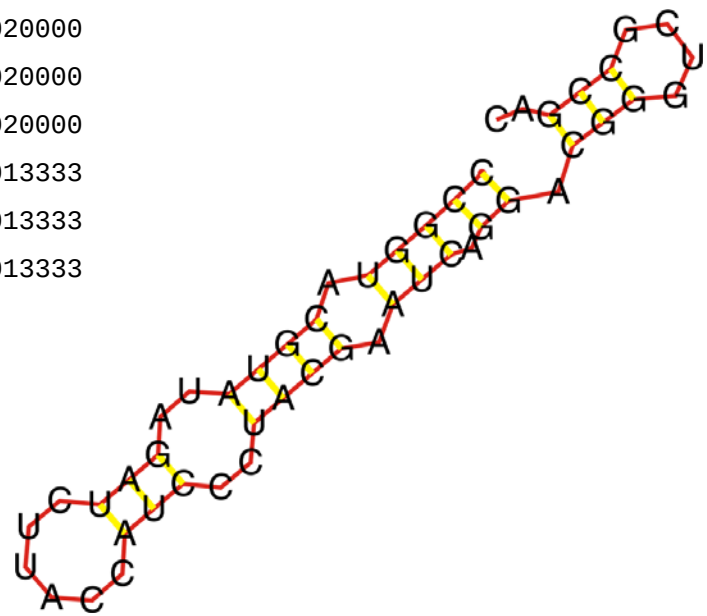
GCAGCUUGCCCAAUGCAACCCCAUGUGGCGCGCUAGCUAACACCAUCCCC

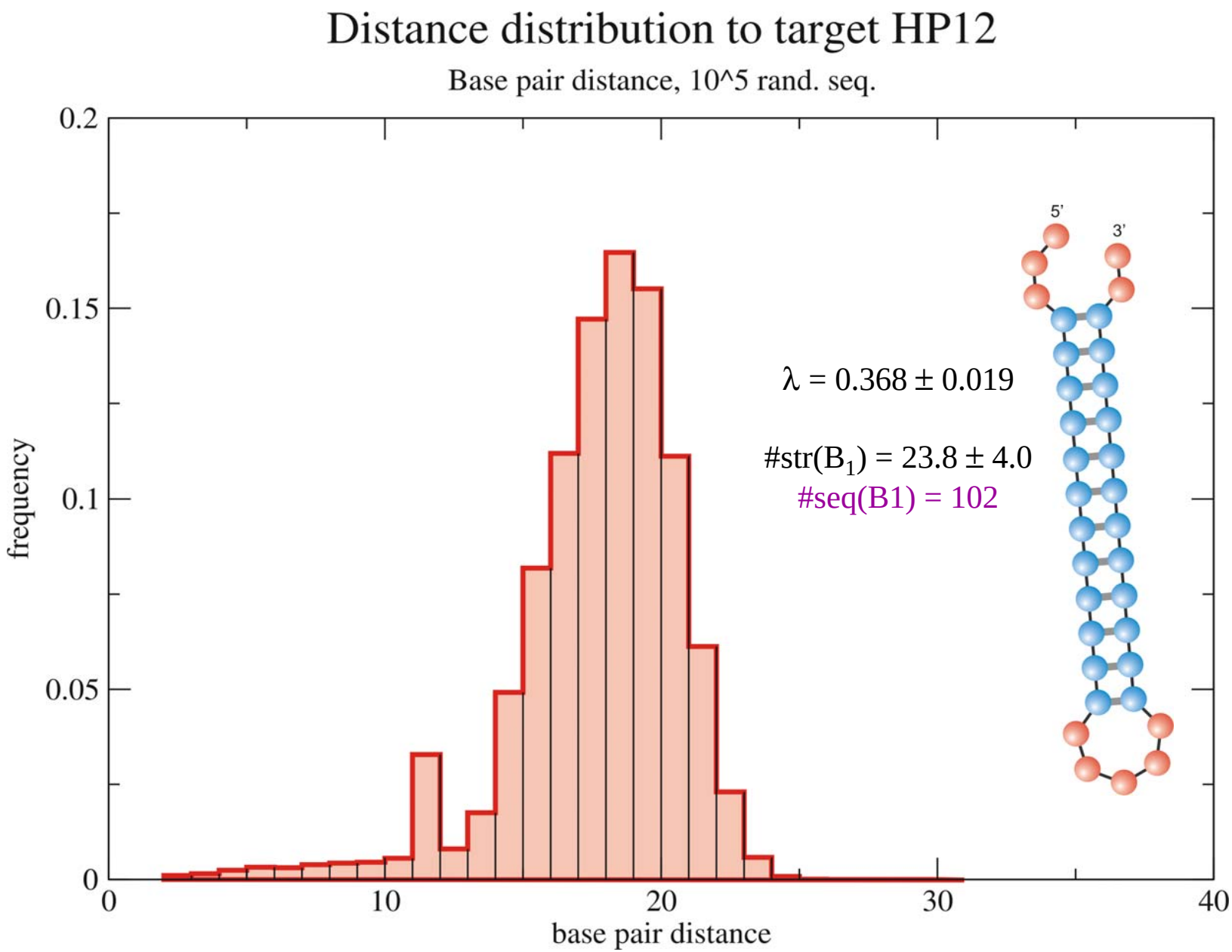
1	**(((((.((((..(((.....))))..)))))..))).))..........**	**65**	**0.433333**
2	..((((((((((((.....))).))).))).))..))))..........**	9	0.060000
3	(((((.((((....(((.....)))))))).))).))..........**	5	0.033333
4	..(((.((((..(((.....))))..)))).))).............**	5	0.033333
5	..((((((((((((.....))).))).)...)))))))..........**	4	0.026667
6	(((((.((((((.((....)).)).)))).))).))..........**	3	0.020000
7	(((((.((((.((((.....)))).)))).))).))..........**	3	0.020000
8	(((((.(((((.(((.....))).)))))).))).))..........**	3	0.020000
9	((((((((((..(((.....)))..)))))))).))..........**	3	0.020000
10	(((((.((((((..........)).)))).))).))..........**	3	0.020000
11	(((((..(((..(((.....)))..)))..))).))..........**	2	0.013333
12	(((((.((((..(((.....)))..)))).)).)))..........**	2	0.013333
13	..((((.((.(..(((.....))))..).)).))))..........**	2	0.013333
14	(((((.((.((((((.....))).)))))).))).))..........**	2	0.013333
15	.((((((((((((.....))).))).))))..)))))..........**	2	0.013333

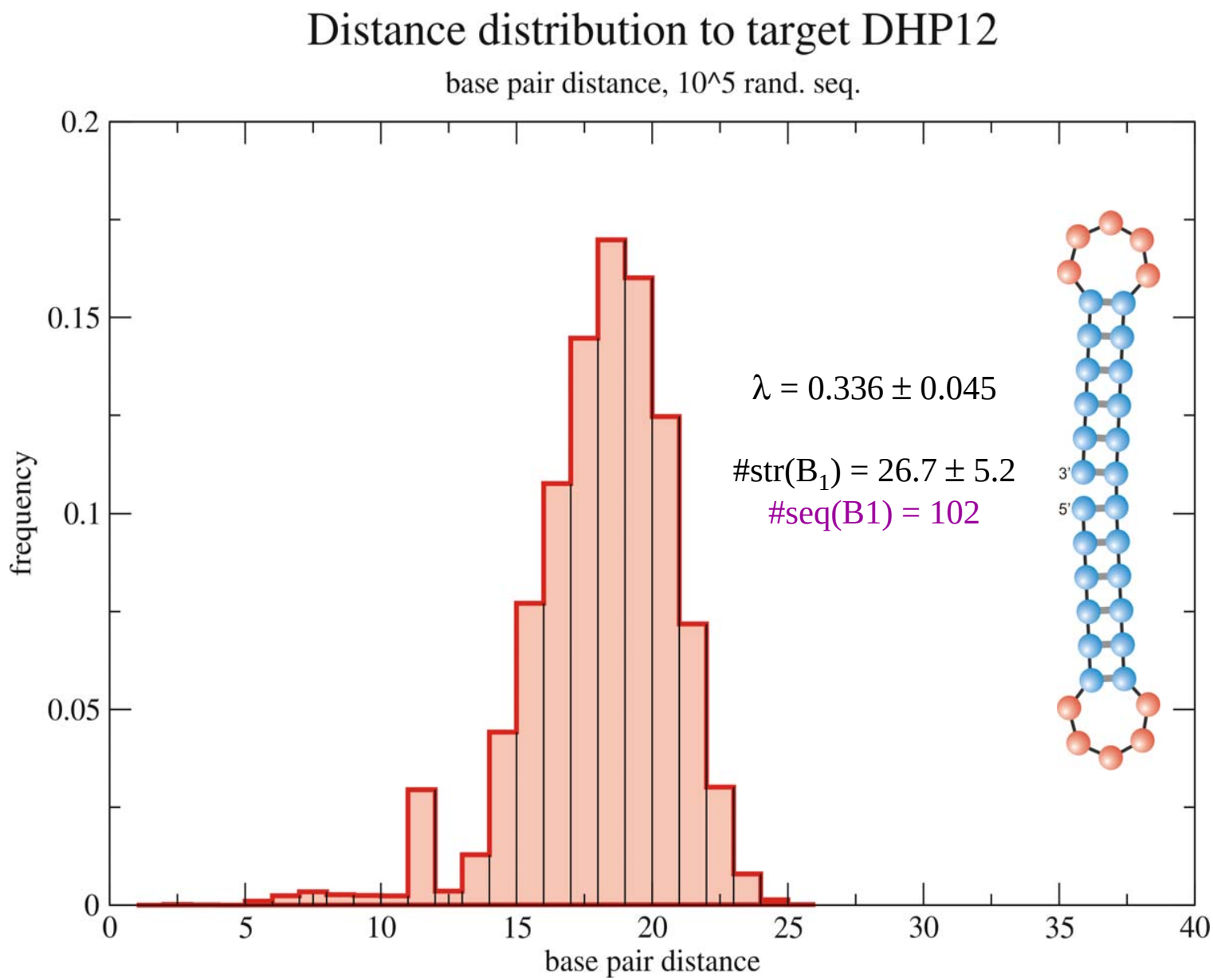


GGAGCUUGCCGAAUGCAACCCCAUGAGGCGCGCUGCCUGGCACCAGCCCC

1	**(((((.((((..(((.....))))..))))).))).(((....))))..**	**49**	**0.326667**
2	**(((((.((((..(((.....))))..))))).))).............**	**7**	**0.046667**
3	..(((.((((..(((.....))))..))))).))).....(((....))))..	6	0.040000
4	(((((.((((..((.......))..))))).))).)).(((....)))..	5	0.033333
5	((.(((((((((...(((.((((....)).).).))).)))))..)))))).	5	0.033333
6	(((((.((((...((.....))...))))).))).)).(((....)))..	5	0.033333
7	(((((.((((..(((.....))))..))))).))).))..((....))...	4	0.026667
8	(((((.((((..(((.....))))..))))).)))))..(((....)))..	4	0.026667
9	(((((.(((...(((.....))))...))).))).)).(((....)))..	3	0.020000
10	((((((((((..(((.....))))..)))))))).)).(((....)))..	3	0.020000
11	((.(((.((((..(((..(.....)..))))..))))..))).))......	3	0.020000
12	(((((...((..(((.....))))..))...))).)).(((....)))..	3	0.020000
13	(.(((.((((..(((.....))))..))))).))).)..(((....)))..	3	0.020000
14	((..(.((((..(((.....))))..))))).)...)).(((....)))..	3	0.020000
15	(((((.(((((.(((.....))).)))))).))).)).(((....)))..	3	0.020000
16	(((((.((((.(((.....))))).))))).))).)).(((....)))..	3	0.020000
17	(((((..(((..(((.....))))..))).)..))).)).(((....)))..	3	0.020000
18	((.(((((((((...(((.(.(.......).).))).)))))..)))))).	2	0.013333
19	(((((.((((..(((.....))))..))))).)).))).(((....)))..	2	0.013333
20	((.(((((((((...(((((((....)).).)))))).)))))..)))))).	2	0.013333

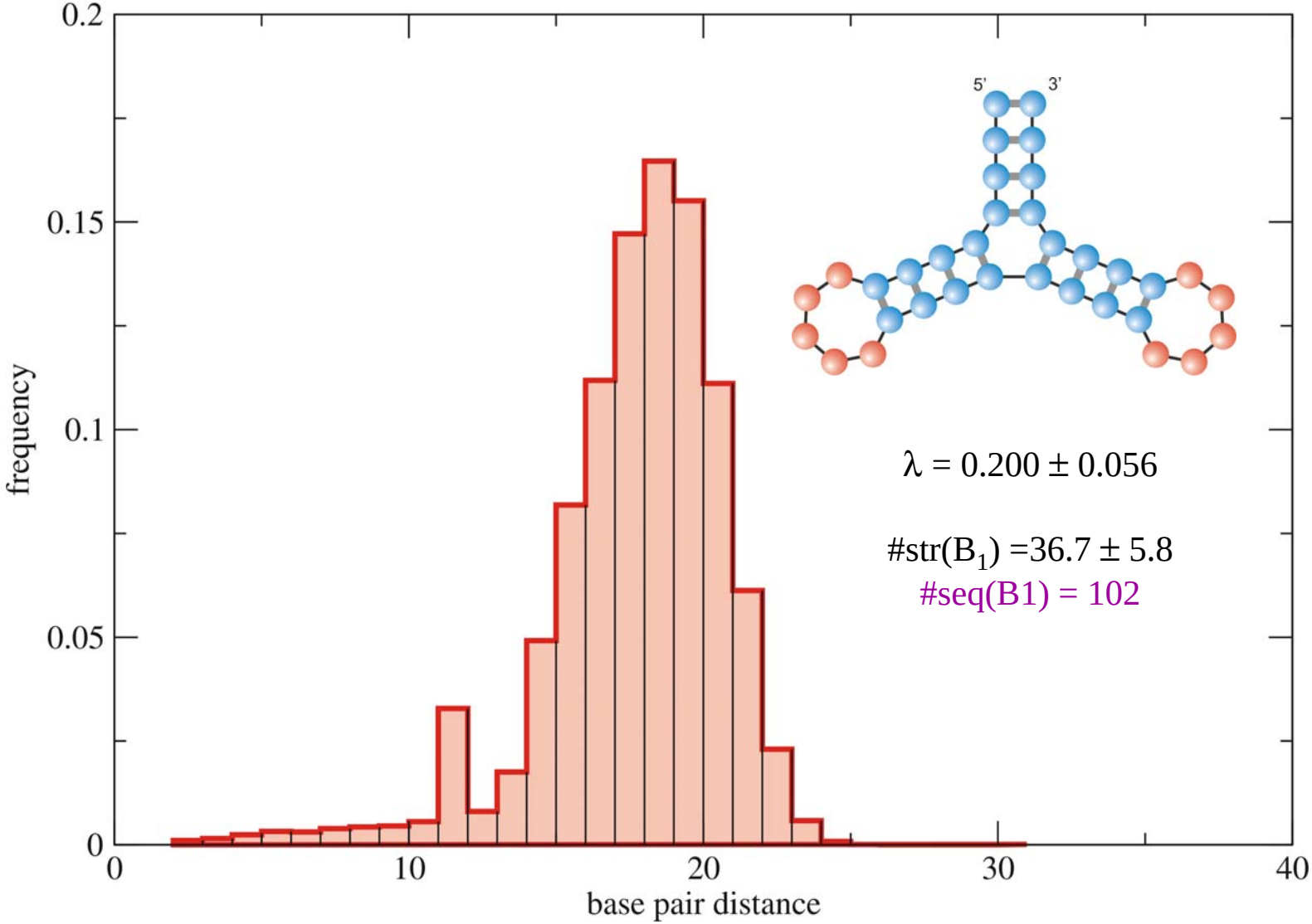



Distance distribution to target HP12
Base pair distance, 10^5 rand. seq.
frequency
base pair distance
0.2
0.15
0.1
0.05
0
0
10
20
30
40
λ = 0.368 ± 0.019
#str(B1) = 23.8 ± 4.0
#seq(B1) = 102
5'
3'


Distance distribution to target DHP12
base pair distance, 10^5 rand. seq.
frequency
base pair distance
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
35
40
λ = 0.336 ± 0.045
#str(B1) = 26.7 ± 5.2
#seq(B1) = 102
3'
5'

Distance distribution to target HP12

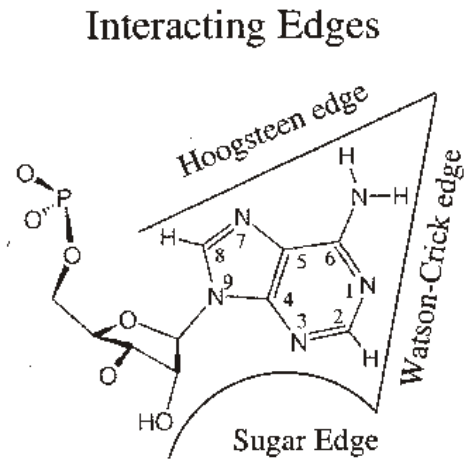
Base pair distance, 10^5 rand. seq.

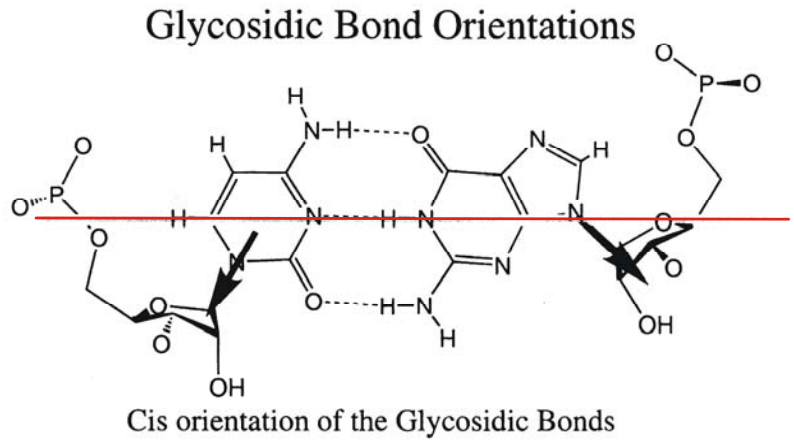


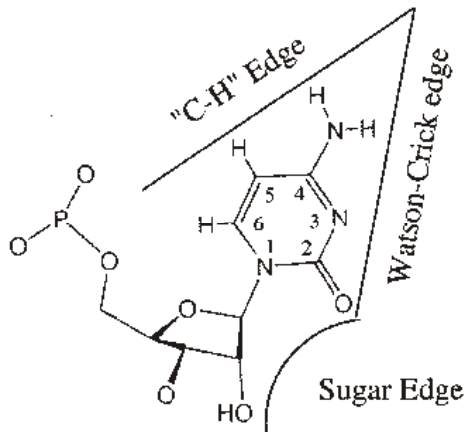
$\lambda = 0.200 \pm 0.056$

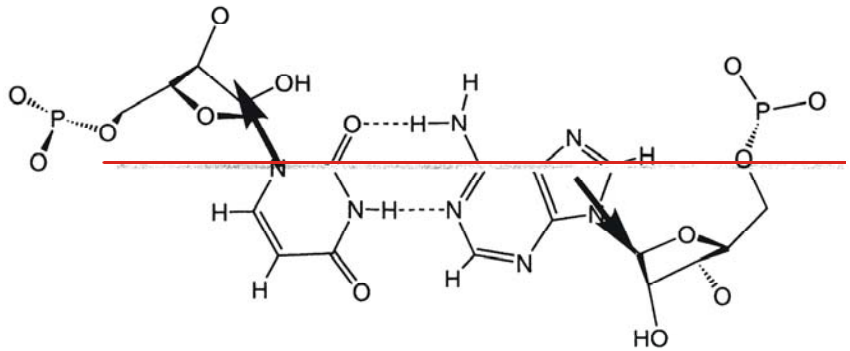
$\#str(B_1) = 36.7 \pm 5.8$

#seq(B1) = 102







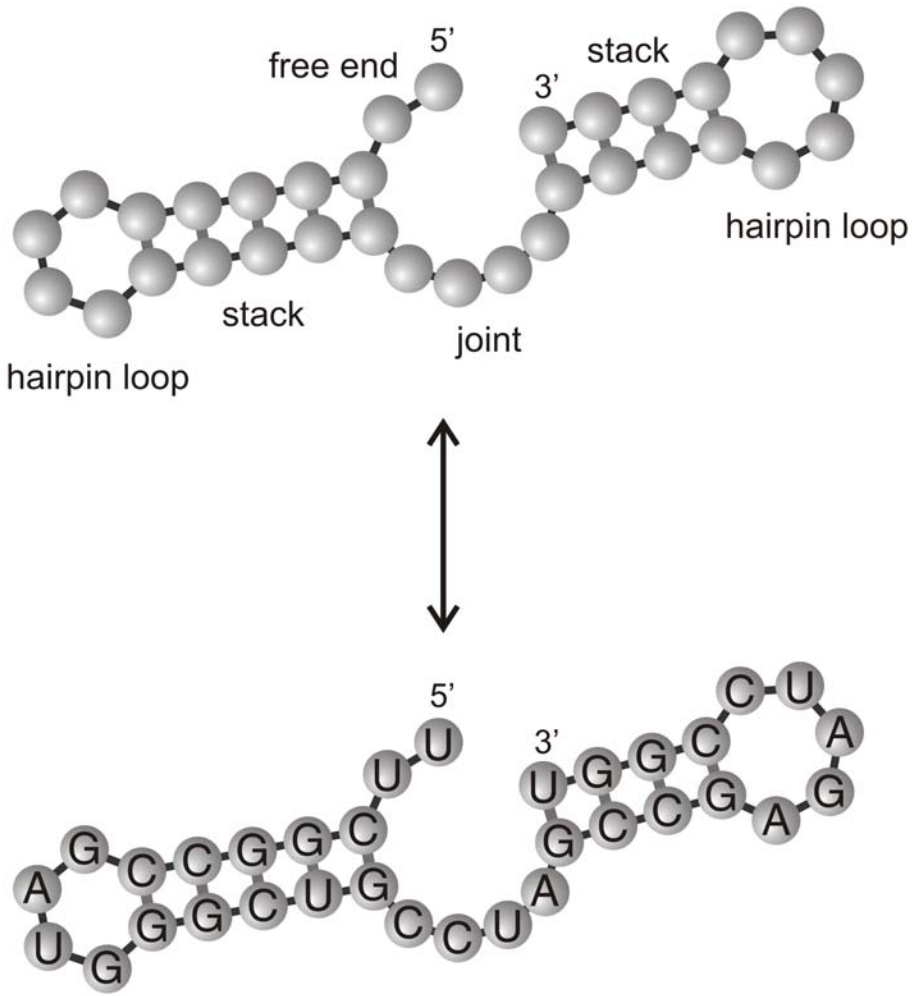


General classification of base pairs

N.B.Leontis and E. Westhof, *RNA* **7**:499-512 (2001)

1. The exciting RNA story
2. Why is gene regulation so complex?
3. What small RNAs can achieve
4. Structures of small RNAs
5. **Riboswitches and kinetic folding**

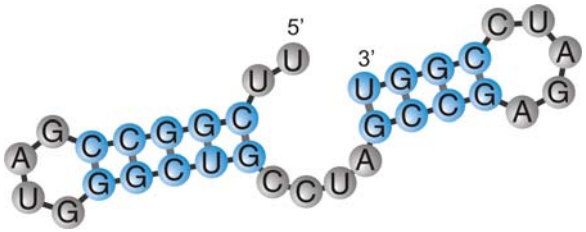
5’ **UUCGGCCGAUGGGCUGCCUAGCCGAGAUCCGGU** 3’

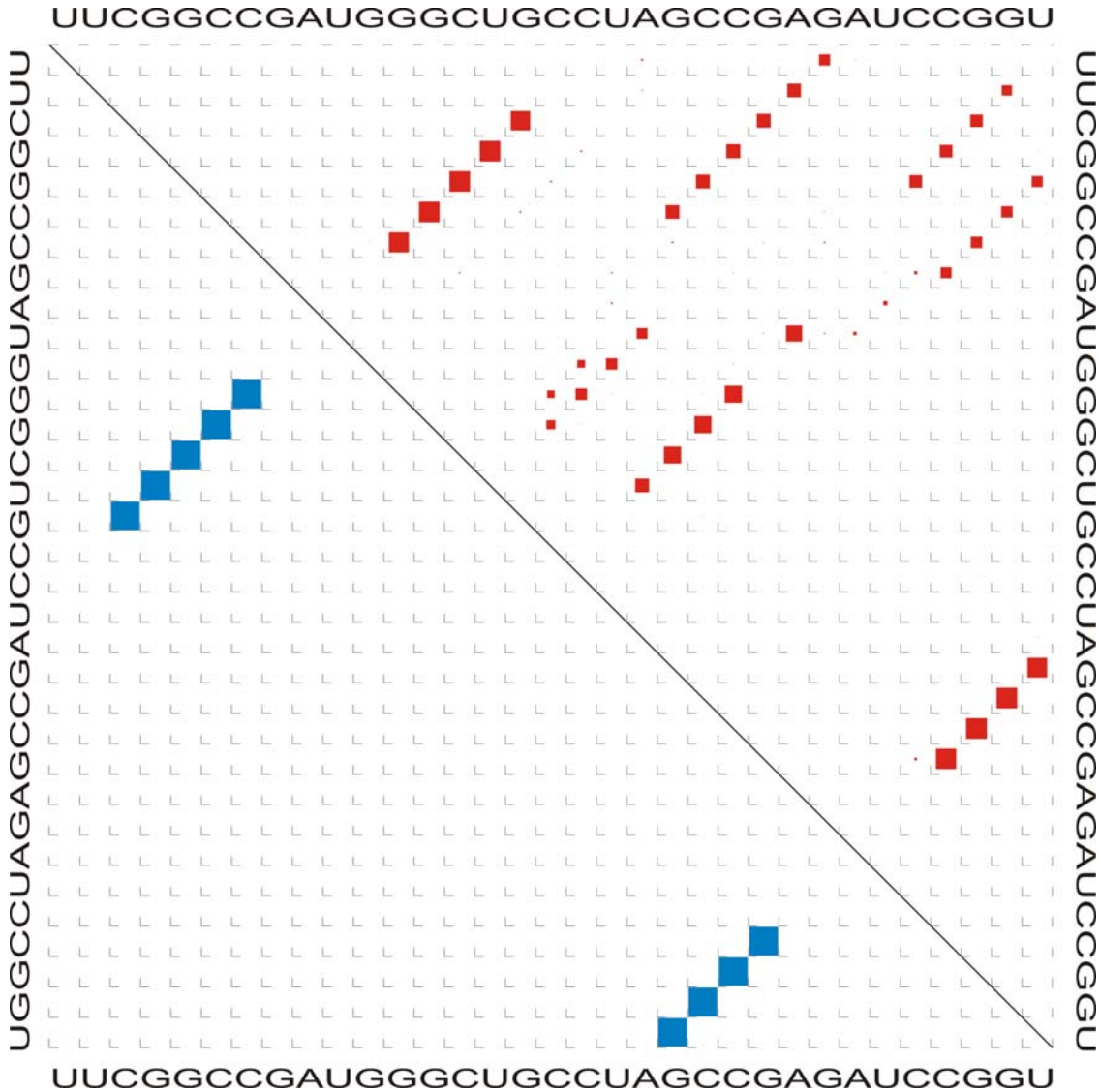


An undesigned RNA switch:

double hairpin 33

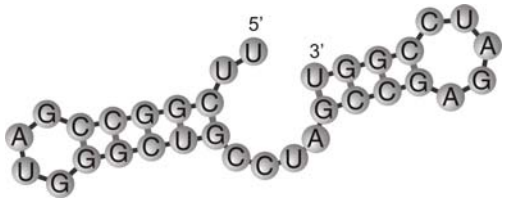
..((((....))))....(((....)))

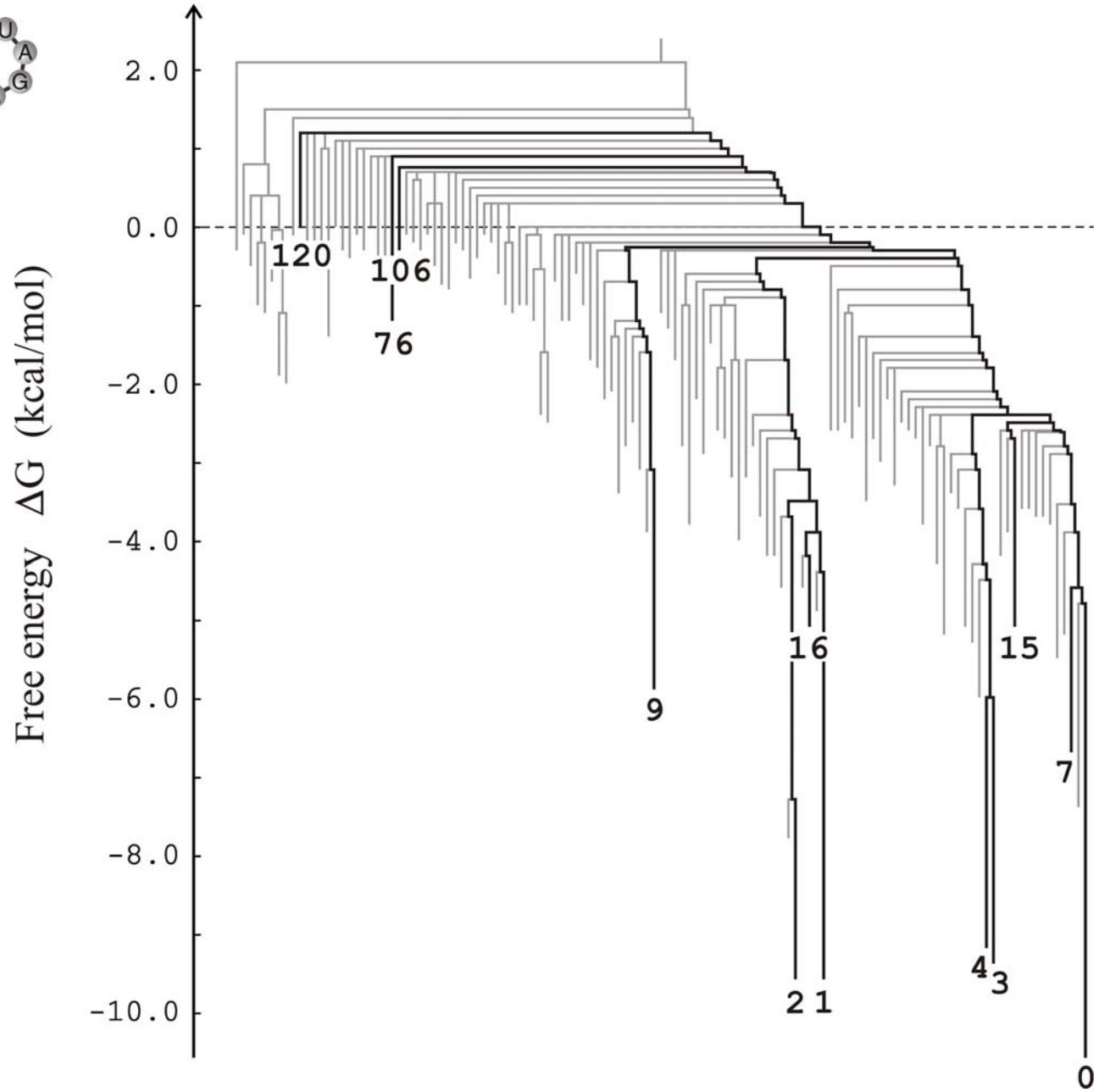




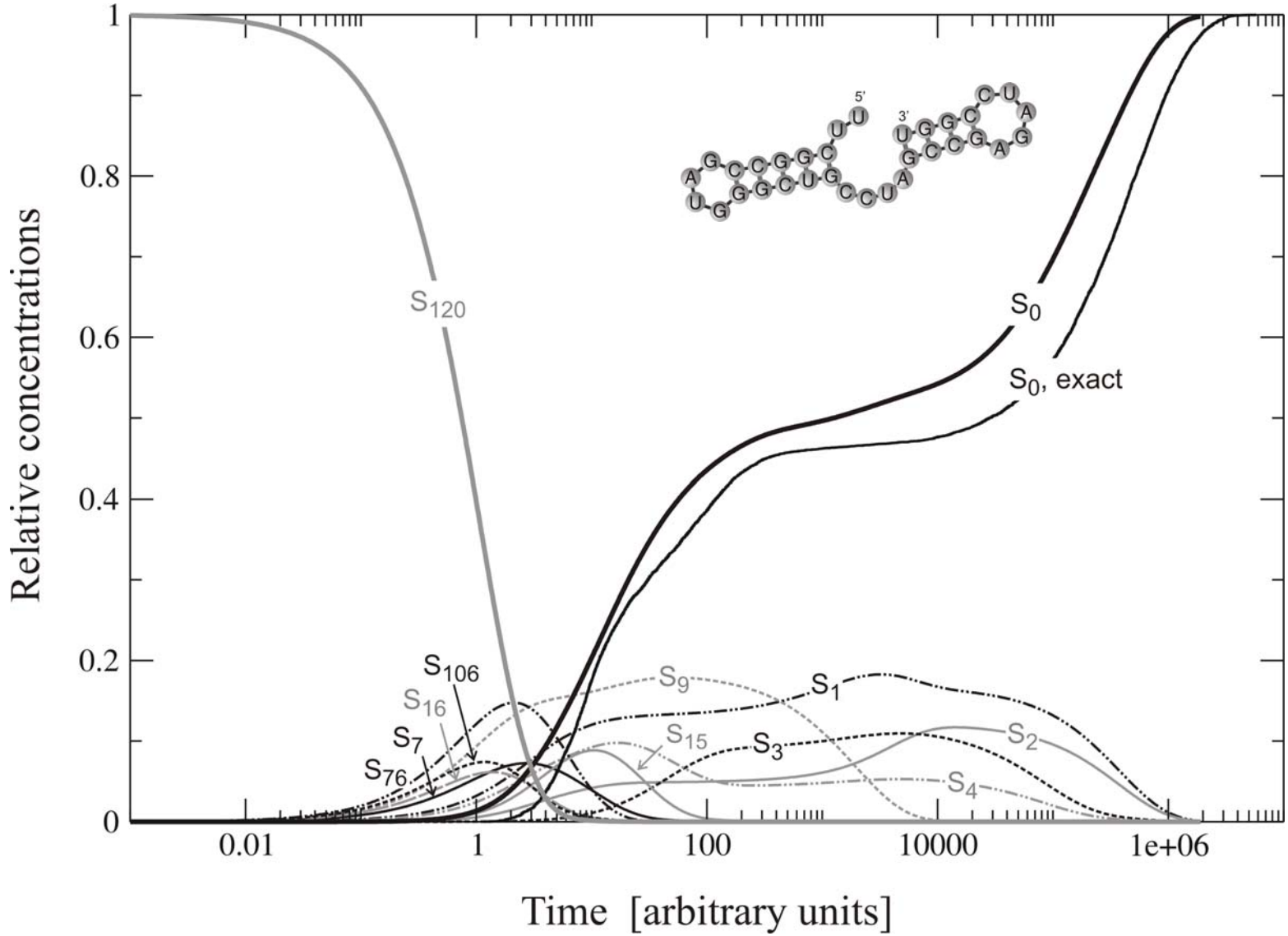
Suboptimal states of double hairpin 33:

dot-plot: ground state and partition function

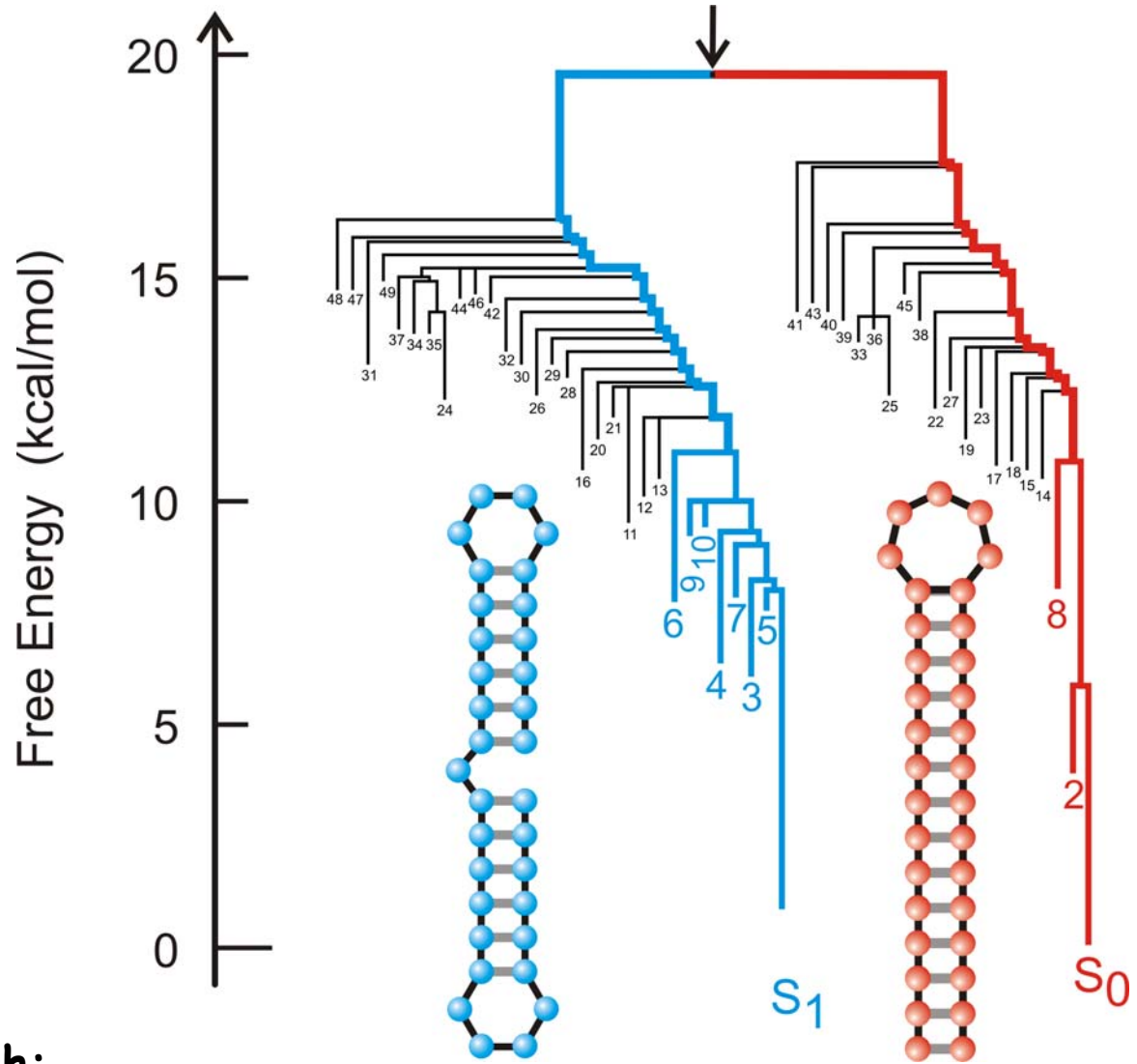




Barrier tree of double hairpin 33

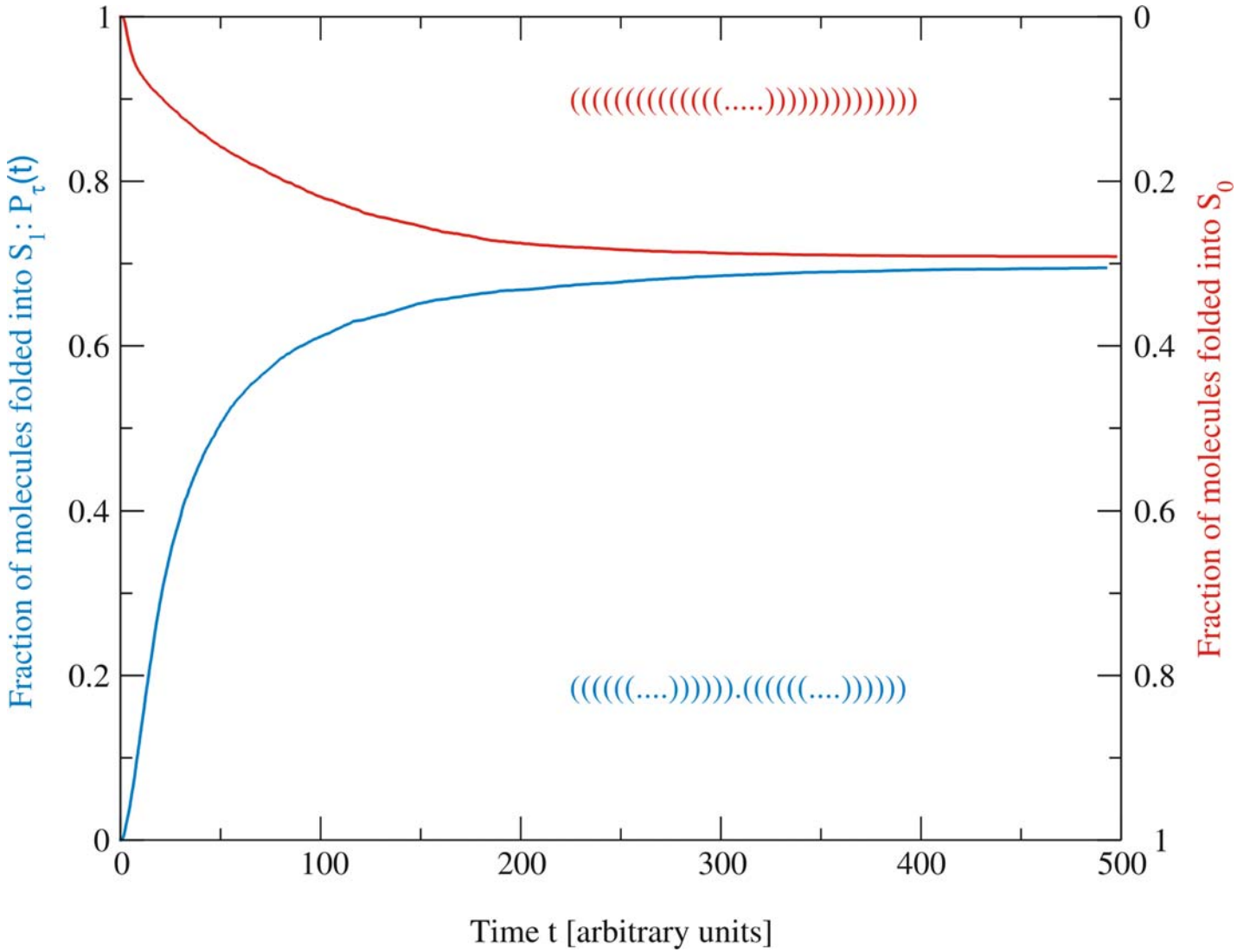


Folding kinetics of double hairpin 33



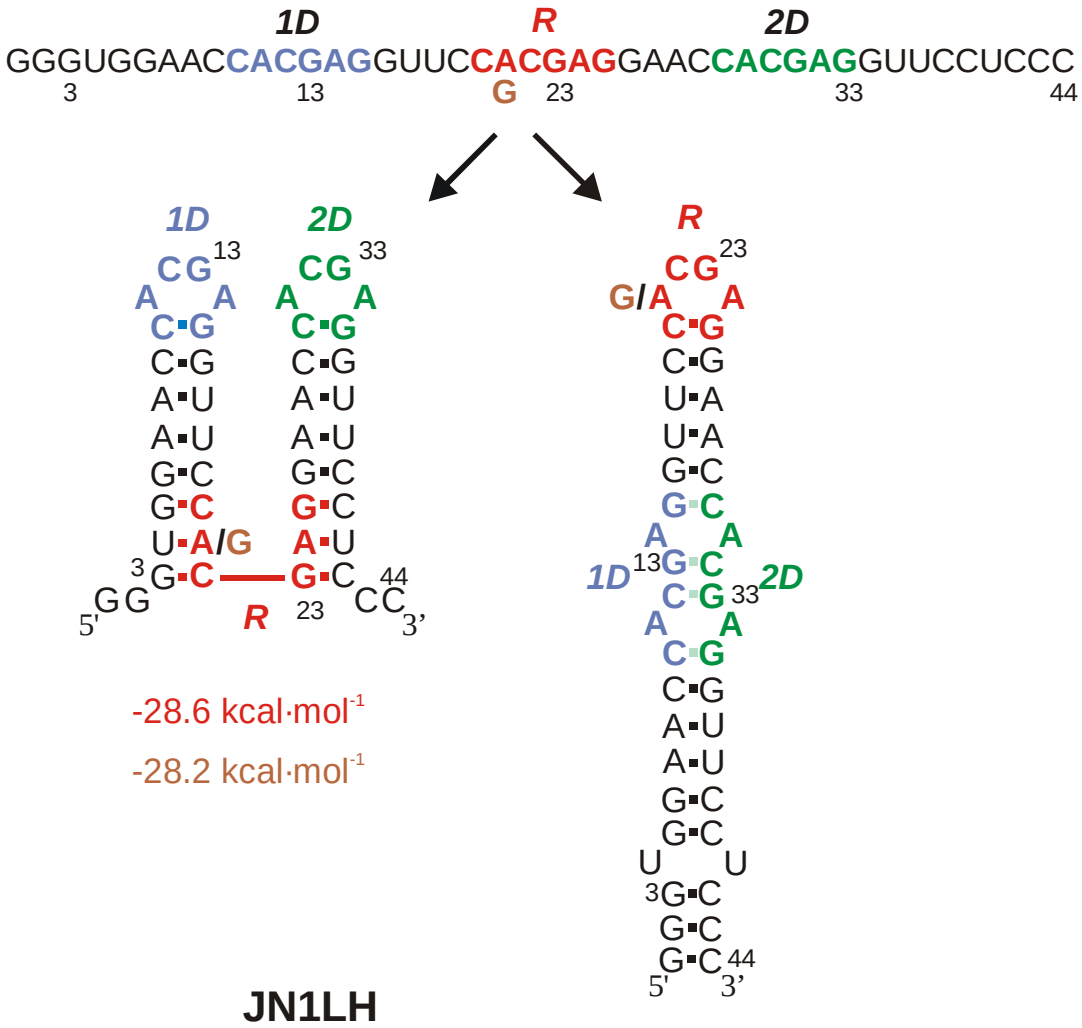
An designed RNA switch:

double hairpin 33



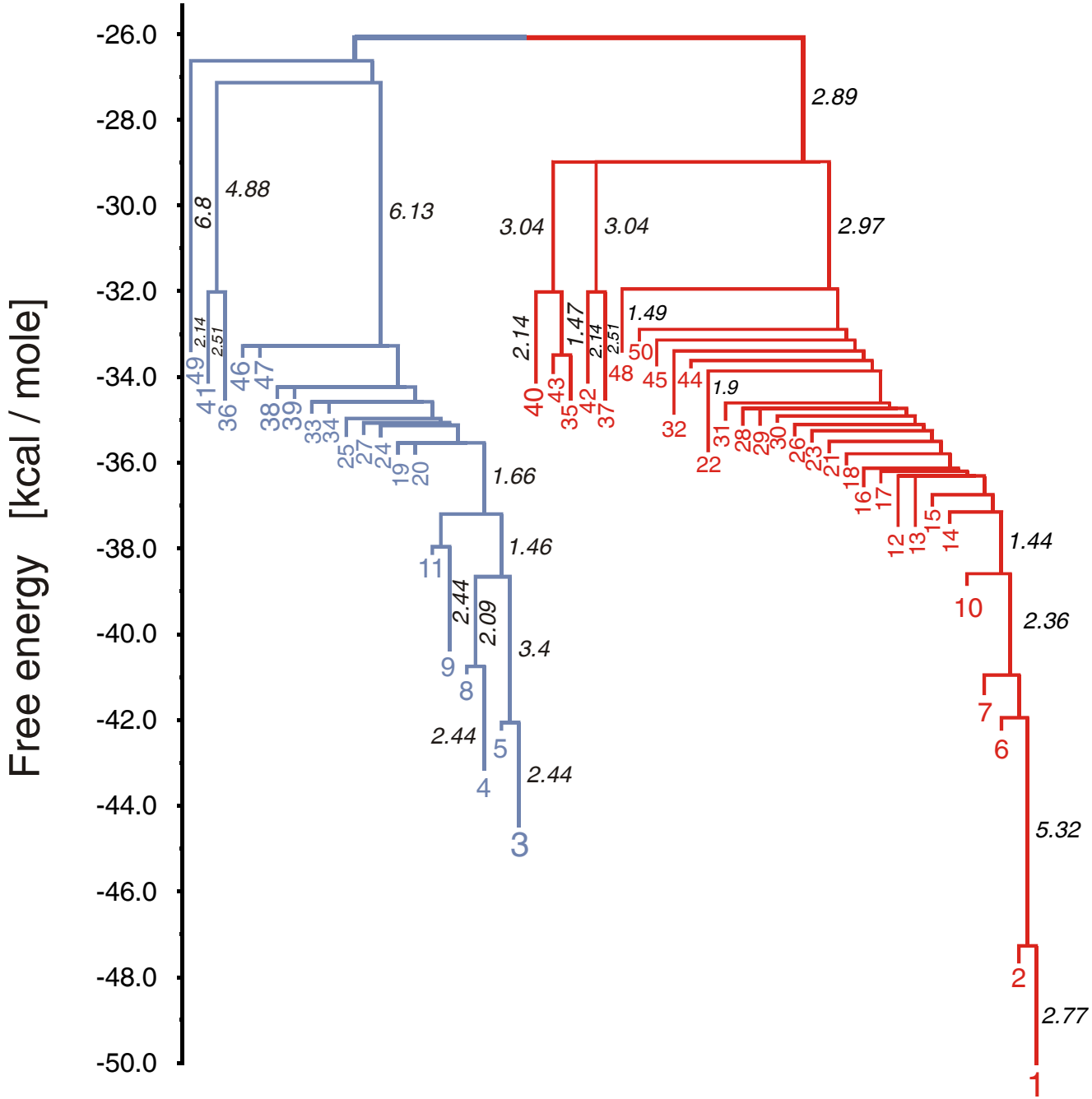
Folding kinetics of the designed double hairpin 33

An experimental RNA switch



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

Structural parameters affecting the kinetic competition of RNA hairpin formation. *Nucleic Acids Res.* **34**:3568-3576 (2006)



J1LH barrier tree

Coworkers

Universität Wien

Peter Stadler, **Bärbel M. Stadler**, Universität Leipzig, GE
Camille Stephan-Otto Atttolini, **Athanasius Bompfüneverer**

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

Walter Fontana, Harvard Medical School, MA

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

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

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

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

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

Acknowledgement of support

Universität Wien

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

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

Jubiläumsfonds der Österreichischen Nationalbank
Project No. Nat-7813

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

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

Österreichische Akademie der Wissenschaften

Siemens AG, Austria

Universität Wien and the Santa Fe Institute

Web-Page for further information:

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