Von der Thermodynamik zu Selbstorganisation, Evolution und Information

Peter Schuster
Institut für Theoretische Chemie und Molekulare Strukturbiologie der Universität Wien

Kolloquium des Physikalischen Instituts

Heidelberg, 17.01.2003
Equilibrium thermodynamics is based on two major statements:

1. The energy of the universe is a constant (first law).
2. The entropy of the universe never decreases (second law).

Carnot, Mayer, Joule, Helmholtz, Clausius, ……

Entropy changes in different thermodynamic systems

**Isolated system**

\[ U = \text{const.}, \ V = \text{const.}, \]

\[ \text{d}S = 0 \]

**Closed system**

\[ T = \text{const.}, \ p = \text{const.}, \]

\[ \text{d}S = \text{d}S_{\text{env}} + \text{d}S_{\text{Stat}} \]

\[ \text{d}G = \text{d}U - p\text{d}V - T\text{d}S \parallel 0 \]

**Open system**

\[ \text{d}S = \text{d}S_{\text{env}} + \text{d}S_{\text{Stat}} \]

\[ \text{d}S = \text{d}_iS + \text{d}_eS \]

\[ \text{d}_iS = 0 \]
Thermodynamics of closed systems:

**Second law**

Entropy is a non-decreasing function

$S \rightarrow S_{\text{max}}$

Entropy and fluctuations at equilibrium

Approach towards equilibrium

$S_{\text{max}}$

Fluctuations around equilibrium

$\left( \frac{d^2 S}{dU, V, \text{equil}} \right) < 0$
Reactions in the continuously stirred tank reactor (CSTR)

Flow rate \( r = t_R^{-1} \)

Stock Solution \([a] = a_0\) →

Reaction Mixture \([a],[b]\) →
Reversible first order reaction in the flow reactor.
Autocatalytic second order and uncatalyzed reaction in the flow reactor
Autocatalytic third order and uncatalyzed reaction in the flow reactor
Autocatalytic third order reactions

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

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

Pattern formation in autocatalytic third order reactions

Autocatalytic second order reactions are the basis of selection processes. The autocatalytic step is formally equivalent to replication or reproduction.
Replication in the flow reactor

Flow rate \( r = t_R^{-1} \)

Concentration of stock solution \( a_0 \)

\[
A + I_1 \rightleftharpoons 2 I_1 \\
A + I_2 \rightleftharpoons 2 I_2 \\
A + I_3 \rightleftharpoons 2 I_3 \\
A + I_4 \rightleftharpoons 2 I_4 \\
A + I_5 \rightleftharpoons 2 I_5 \\
\]

Selection in the flow reactor: Reversible replication reactions

\( k_1 > k_2 > k_3 > k_4 > k_5 \)
Selection in the flow reactor: Irreversible replication reactions

Flow rate \( r = \frac{1}{t_R} \)

Concentration of stock solution \( a_0 \)

A + I_1 \rightarrow 2 I_1
A + I_2 \rightarrow 2 I_2
A + I_3 \rightarrow 2 I_3
A + I_4 \rightarrow 2 I_4
A + I_5 \rightarrow 2 I_5

k_1 > k_2 > k_3 > k_4 > k_5
Complementary replication as the simplest copying mechanism of RNA. Complementarity is determined by Watson-Crick base pairs: $G \bar{C}$ and $A=U$. 
dx_i / dt = f_i x_i - x_i Φ = x_i (f_i - Φ)

Φ = Σ_j f_j x_j ;  Σ_j x_j = 1 ;  i,j = 1,2,...,n

[I_i] = x_i [A] ;  i = 1,2,...,n ;

[A] = a = constant

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

x_m(t) ≥ 1 for t ≥ .

**Reproduction** of organisms or replication of molecules as the basis of selection
Selection equation: \[ [I_i] = x_i A^0, \quad f_i > 0 \]

\[
\frac{dx_i}{dt} = x_i \left( f_i - \phi \right), \quad i = 1, 2, \ldots, n; \quad \sum_{i=1}^{n} x_i = 1; \quad \phi = \sum_{j=1}^{n} f_j x_j = \bar{f}
\]

Mean fitness or dilution flux, \( \phi (t) \), is a non-decreasing function of time,

\[
\frac{d\phi}{dt} = \sum_{i=1}^{n} f_i \frac{dx_i}{dt} = \bar{f}^2 - \left( \bar{f} \right)^2 = \text{var}\{f\} \geq 0
\]

**Solutions** are obtained by integrating factor transformation

\[
x_i(t) = \frac{x_i(0) \cdot \exp(f_i t)}{\sum_{j=1}^{n} x_j(0) \cdot \exp(f_j t)}; \quad i = 1, 2, \ldots, n
\]
\[ s = \frac{f_2 - f_1}{f_1}; \quad f_2 > f_1; \quad x_1(0) = 1 - 1/N; \quad x_2(0) = 1/N \]

Selection of advantageous mutants in populations of \( N = 10\,000 \) individuals
Evolution of Populations:

**Mean fitness** is a non-decreasing function

Ronald Fisher’s conjecture

\[ \bar{f} = \frac{S_k x_k(t) f_k}{S_k x_k(t)} \text{ Å } f_{\text{max}} \]

Thermodynamics of closed systems:

**Entropy** is a non-decreasing function

Second law

\[ S \rightarrow S_{\text{max}} \]
The origins of changes in genotypes or mutations are either replication errors or damage of nucleic acid molecules.
Mutations in nucleic acids represent the mechanism of variation of genotypes.
Theory of molecular evolution


\[ \frac{dx_i}{dt} = \sum_j f_j Q_{ji} x_j - x_i \Phi \]

\[ \Phi = \sum_j f_j x_j \]; \quad \sum_j x_j = 1 \quad \sum_i Q_{ij} = 1 \]

\[ [I_i] = x_i \mathcal{A}^i \]; \quad i = 1, 2, \ldots, n \]

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

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

\( p \) .......... Error rate per digit

\( \ell \) .......... Chain length of the polynucleotide

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

Chemical kinetics of replication and mutation as parallel reactions
The Hamming distance induces a metric in sequence space

Hamming distance \( d_H(S_1, S_2) = 4 \)

(i) \( d_H(S_1, S_1) = 0 \)
(ii) \( d_H(S_1, S_2) = d_H(S_2, S_1) \)
(iii) \( d_H(S_1, S_3) \leq d_H(S_1, S_2) + d_H(S_2, S_3) \)
Mutation-selection equation: \([I_i] = x_i \mathcal{A}_0, f_i > 0, Q_{ij} \mathcal{A}_0\)

\[
\frac{dx_i}{dt} = \sum_{j=1}^{n} f_j Q_{ji} x_j - x_i \phi, \quad i=1,2,\ldots,n; \quad \sum_{i=1}^{n} x_i = 1; \quad \phi = \sum_{j=1}^{n} f_j x_j = \bar{f}
\]

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

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

\[
W = \{ f_i Q_{ij}; i, j=1,2,\ldots,n \}; \quad L = \{ \ell_{ij}; i, j=1,2,\ldots,n \}; \quad L^{-1} = H = \{ h_{ij}; i, j=1,2,\ldots,n \}
\]

\[
L^{-1} \cdot W \cdot L = \Lambda = \{ \lambda_k; k=0,1,\ldots,n-1 \}
\]
Quasispecies as a function of the replication accuracy $q$
The molecular quasispecies in sequence space

- Master sequence
- Mutant cloud
The quasispecies on the concentration simplex $S_3 = \{ x_i \geq 0, \ i = 1, 2, 3; \sum_{i=1}^{3} x_i = 1 \}$
Decrease in mean fitness due to quasispecies formation

The increase in RNA production rate during a serial transfer experiment
Ronald Fisher’s conjecture of optimization of mean fitness in populations does not always hold for replication-mutation systems: In general evolutionary dynamics the mean fitness of populations may also decrease monotonously or even go through a maximum or minimum.

Optimization of fitness is, nevertheless, fulfilled in most cases, and it can be understood as a useful heuristic.
**Learning** and **origin of information** on the level of populations:

The change in populations as a consequence of adaptation to the environment creates **information** (on the environmental conditions).

The population „**learns“** through mutation and selection.

**Information** is conserved through replication that reproduces mutants and error-free copies in the same way.
Theory of genotype – phenotype mapping


The RNA secondary structure is a listing of GC, AU, and GU base pairs. It is understood in contrast to the full 3D- or tertiary structure at the resolution of atomic coordinates. RNA secondary structures are biologically relevant. They are, for example, conserved in evolution.
The **inverse folding algorithm** searches for sequences that form a given RNA secondary structure under the minimum free energy criterion.
Criterion of Minimum Free Energy

Sequence Space

Shape Space
$S_k = \psi(I.)$

$f_k = f(S_k)$

Mapping from sequence space into phenotype space and into fitness values
$S_k = \psi(I.)$

$f_k = f(S_k)$

Sequence space  Phenotype space  Non-negative numbers
The pre-image of the structure $S_k$ in sequence space is the neutral network $G_k$. 

$S_k = \psi(I.)$

$f_k = f(S_k)$
Neutral networks are sets of sequences forming the same structure. $G_k$ is the pre-image of the structure $S_k$ in sequence space:

$$G_k = y^{-1}(S_k) \cup \{y_j \mid y(I_j) = S_k\}$$

The set is converted into a graph by connecting all sequences of Hamming distance one.

Neutral networks of small RNA molecules can be computed by exhaustive folding of complete sequence spaces, i.e. all RNA sequences of a given chain length. This number, $N=4^n$, becomes very large with increasing length, and is prohibitive for numerical computations.

Neutral networks can be modelled by random graphs in sequence space. In this approach, nodes are inserted randomly into sequence space until the size of the pre-image, i.e. the number of neutral sequences, matches the neutral network to be studied.
Random graph approach to neutral networks

Step 00

Sketch of sequence space
Step 01

Random graph approach to neutral networks

Sketch of sequence space
Step 02

Sketch of sequence space

Random graph approach to neutral networks
Step 03

Random graph approach to neutral networks
Random graph approach to neutral networks
Step 05

Sketch of sequence space

Random graph approach to neutral networks
Random graph approach to neutral networks
Random graph approach to neutral networks

Sketch of sequence space
Random graph approach to neutral networks

Sketch of sequence space
Random graph approach to neutral networks
Random graph approach to neutral networks

Sketch of sequence space

Step 75
Random graph approach to neutral networks
\[ G_k = y^{-1}(S_k) \cup \{ I_j \mid y(I_j) = S_k \} \]

\[ \lambda_j = \frac{12}{27}, \quad \bar{\lambda}_k = \frac{\varnothing \setminus j(k)}{|G_k|} \]

Connectivity threshold:

\[ \lambda_{cr} = 1 - \kappa^{-1/(\kappa-1)} \]

Alphabet size \( k \): \( \text{AUGC} \) \( n \) \( k = 4 \)

<table>
<thead>
<tr>
<th>( k )</th>
<th>( l_{cr} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>0.4226</td>
</tr>
<tr>
<td>4</td>
<td>0.3700</td>
</tr>
</tbody>
</table>

\( \bar{\lambda}_k > \lambda_{cr} \) \ldots network \( G_k \) is connected

\( \bar{\lambda}_k < \lambda_{cr} \) \ldots network \( G_k \) is not connected

Mean degree of neutrality and connectivity of neutral networks
A multi-component neutral network
A connected neutral network
Compatibility of sequences with structures

A sequence is compatible with its minimum free energy structure and all its suboptimal structures.
The compatible set $C_k$ of a structure $S_k$ consists of all sequences which form $S_k$ as its minimum free energy structure (neutral network $G_k$) or one of its suboptimal structures.
A sequence at the **intersection** of two neutral networks is compatible with both structures.
The intersection of two compatible sets is always non empty: \( C_1 \cap C_2 \neq \emptyset \)
Optimization of RNA molecules *in silico*


Randomly chosen initial structure  Phenylalanyl-tRNA as target structure
Fitness function:
\[ f_k = \frac{g}{a + Dd_s^{(k)}} \]
\[ Dd_s^{(k)} = d^s(I_k,I_t) \]

The flowreactor as a device for studies of evolution *in vitro* and *in silico*
The molecular quasispecies in sequence space
Evolutionary dynamics
including molecular phenotypes
**In silico** optimization in the flow reactor: Trajectory (biologists' view)
In silico optimization in the flow reactor: Trajectory (physicists' view)
In silico optimization in the flow reactor: Main transitions
Main or discontinuous transitions: *Structural innovations*, occur rarely on single point mutations.

Closing of Constrained Stacks
In silico optimization in the flow reactor
Elongation of Stacks

Shortening of Stacks

Opening of Constrained Stacks

Minor or continuous transitions: Occur frequently on single point mutations

Opening of Constrained Stacks
## Statistics of evolutionary trajectories

<table>
<thead>
<tr>
<th>Population size N</th>
<th>Number of replications $&lt; n_{\text{rep}} &gt;$</th>
<th>Number of transitions $&lt; n_{\text{tr}} &gt;$</th>
<th>Number of main transitions $&lt; n_{\text{dtr}} &gt;$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 000</td>
<td>$(5.5 \pm [6.9,3.1]) \times 10^7$</td>
<td>$92.7 \pm [80.3,43.0]$</td>
<td>$8.8 \pm [2.4,1.9]$</td>
</tr>
<tr>
<td>2 000</td>
<td>$(6.0 \pm [11.1,3.9]) \times 10^7$</td>
<td>$55.7 \pm [30.7,19.8]$</td>
<td>$8.9 \pm [2.8,2.1]$</td>
</tr>
<tr>
<td>3 000</td>
<td>$(6.6 \pm [21.0,5.0]) \times 10^7$</td>
<td>$44.2 \pm [25.9,16.3]$</td>
<td>$8.1 \pm [2.3,1.8]$</td>
</tr>
<tr>
<td>10 000</td>
<td>$(1.2 \pm [1.3,0.6]) \times 10^8$</td>
<td>$35.9 \pm [10.3,8.0]$</td>
<td>$10.3 \pm [2.6,2.1]$</td>
</tr>
<tr>
<td>20 000</td>
<td>$(1.5 \pm [1.4,0.7]) \times 10^8$</td>
<td>$28.8 \pm [5.8,4.8]$</td>
<td>$9.0 \pm [2.8,2.2]$</td>
</tr>
<tr>
<td>30 000</td>
<td>$(2.2 \pm [3.1,1.3]) \times 10^8$</td>
<td>$29.8 \pm [7.3,5.9]$</td>
<td>$8.7 \pm [2.4,1.9]$</td>
</tr>
<tr>
<td>100 000</td>
<td>$(3 \pm [2,1]) \times 10^8$</td>
<td>$24 \pm [6,5]$</td>
<td>$9 \pm 2$</td>
</tr>
</tbody>
</table>

The number of **main transitions** or evolutionary innovations is constant.
Stable tRNA clover leaf structures built from binary, GC-only, sequences exist. The corresponding sequences are readily found through inverse folding. Optimization by mutation and selection in the flow reactor has so far always been unsuccessful.

The neutral network of the tRNA clover leaf in GC sequence space is not connected, whereas to the corresponding neutral network in AUGC sequence space is very close to the critical connectivity threshold, $l_{cr}$. Here, both inverse folding and optimization in the flow reactor are successful.

The success of optimization depends on the connectivity of neutral networks.
“...Variations neither useful nor injurious would not be affected by natural selection, and would be left either a fluctuating element, as perhaps we see in certain polymorphic species, or would ultimately become fixed, owing to the nature of the organism and the nature of the conditions. ...”

Charles Darwin, Origin of species (1859)
Evolution in genotype space sketched as a non-descending walk in a fitness landscape.
Evolutionary design of RNA molecules


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

Selection cycle used in applied molecular evolution to design molecules with predefined properties.
The SELEX technique for the evolutionary design of *aptamers*
Secondary structures of aptamers binding theophyllin, caffeine, and related compounds.
Dissociation constants and specificity of theophylline, caffeine, and related derivatives of uric acid for binding to a discriminating aptamer TCT8-4

Table 1. Competition binding analysis with TCT8-4 RNA. The chemical structures are shown for a series of derivatives used in competitive binding experiments with TCT8-4 RNA (Fig. 2) (20). The right column represents the affinity of the competitor relative to theophylline, $K_d(c)/K_d(t)$, where $K_d(c)$ is the individual competitor dissociation constant and $K_d(t)$ is the competitive dissociation constant of theophylline. Certain data (denoted by $>$) are minimum values that were limited by the solubility of the competitor. Each experiment was carried out in duplicate. The average error is shown.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>$K_d(c)$ ($\mu$M)</th>
<th>$K_d(c)/K_d(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theophylline</td>
<td><img src="theophylline.png" alt="Structure" /></td>
<td>0.32 ± 0.13</td>
<td>1</td>
</tr>
<tr>
<td>CP-theophylline</td>
<td><img src="CP-theophylline.png" alt="Structure" /></td>
<td>0.93 ± 0.20</td>
<td>2.9</td>
</tr>
<tr>
<td>Xanthine</td>
<td><img src="xanthine.png" alt="Structure" /></td>
<td>8.5 ± 0.40</td>
<td>27</td>
</tr>
<tr>
<td>1-Methylxanthine</td>
<td><img src="1-methylxanthine.png" alt="Structure" /></td>
<td>9.0 ± 0.30</td>
<td>28</td>
</tr>
<tr>
<td>3-Methylxanthine</td>
<td><img src="3-methylxanthine.png" alt="Structure" /></td>
<td>2.0 ± 0.7</td>
<td>6.3</td>
</tr>
<tr>
<td>7-Methylxanthine</td>
<td><img src="7-methylxanthine.png" alt="Structure" /></td>
<td>&gt; 500</td>
<td>&gt;1500</td>
</tr>
<tr>
<td>3,7-Dimethylxanthine</td>
<td><img src="3,7-dimethylxanthine.png" alt="Structure" /></td>
<td>&gt; 500</td>
<td>&gt;1500</td>
</tr>
<tr>
<td>1,3-Dimethyluric acid</td>
<td><img src="1,3-dimethyluric.png" alt="Structure" /></td>
<td>&gt; 1000</td>
<td>&gt;3100</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td><img src="hypoxanthine.png" alt="Structure" /></td>
<td>49 ± 10</td>
<td>153</td>
</tr>
<tr>
<td>Caffeine</td>
<td><img src="caffeine.png" alt="Structure" /></td>
<td>3500 ± 1500</td>
<td>10,900</td>
</tr>
</tbody>
</table>
Schematic drawing of the aptamer binding site for the theophylline molecule.
Aptamer binding to aminoglycosid antibiotics: Structure of ligands

Formation of secondary structure of the tobramycin binding RNA aptamer

The three-dimensional structure of the tobramycin aptamer complex

A ribozyme switch

Two ribozymes of chain lengths \( n = 88 \) nucleotides: An artificial ligase (A) and a natural cleavage ribozyme of hepatitis-\( d \)-virus (B)
The sequence at the **intersection**: 

An RNA molecule which is 88 nucleotides long and can form both structures
RANDOM GRAPH THEORY is used to model and analyze the relationships between sequences and secondary structures of RNA molecules, which are understood as mappings from sequence space into shape space. These maps are non-invertible since there are always many orders of magnitude more sequences than structures. Sequences folding into identical structures form neutral networks. A neutral network is embedded in the set of sequences that are compatible with the given structure. Networks are modeled as graphs and constructed by random choice of vertices from the space of compatible sequences. The theory characterizes neutral networks by the mean fraction of neutral neighbors (λ). The networks are connected and percolate sequence space if the fraction of neutral nearest neighbors exceeds a threshold value (λ > λ*). Below threshold (λ < λ*), the networks are partitioned into a largest “giant” component and several smaller components. Structures are classified as “common” or “rare” according to the sizes of their pre-images, i.e., according to the fractions of sequences folding into them. The neutral networks of any pair of two different common structures almost touch each other, and, as expressed by the conjecture of shape space covering sequences folding into almost all common structures, can be found in a small ball of an arbitrary location in sequence space. The results from random graph theory are compared to data obtained by folding large samples of RNA sequences. Differences are explained in terms of specific features of RNA molecular structures. © 1997 Society for Mathematical Biology

THEOREM 5. INTERSECTION-THEOREM. Let s and s' be arbitrary secondary structures and C[s], C[s'] their corresponding compatible sequences. Then,

C[s] \cap C[s'] \neq \emptyset.

Proof. Suppose that the alphabet admits only the complementary base pair [XY] and we ask for a sequence x compatible to both s and s'. Then x(s, s') \equiv D_n operates on the set of all positions \{x_1, ..., x_n\}. Since we have the operation of a dihedral group, the orbits are either cycles or chains and the cycles have even order. A constraint for the sequence compatible to both structures appears only in the cycles where the choice of bases is not independent. It remains to be shown that there is a valid choice of bases for each cycle, which is obvious since these have even order. Therefore, it suffices to choose an alternating sequence of the pairing partners X and Y. Thus, there are at least two different choices for the first base in the orbit.

Remark. A generalization of the statement of theorem 5 to three different structures is false.
Two neutral walks through sequence space with conservation of structure and catalytic activity
Sequence of mutants from the intersection to both reference ribozymes
From sequences to shapes and back: a case study in RNA secondary structures

PETER SCHUSTER\textsuperscript{1,2,3}, WALTER FONTANA\textsuperscript{1}, PETER F. STADLER\textsuperscript{2,3} and IVO L. HOFACKER\textsuperscript{2}

\textsuperscript{1}Institut für Molekulare Biotechnologie, Beutenbergstrasse 11, PF 1000813, D-07708 Jena, Germany
\textsuperscript{2}Institut für Theoretische Chemie, Universität Wien, Austria
\textsuperscript{3}Santa Fe Institute, Santa Fe, U.S.A.

SUMMARY

RNA folding is viewed here as a map assigning secondary structures to sequences. At fixed chain length the number of sequences far exceeds the number of structures. Frequencies of structures are highly non-uniform and follow a generalized form of Zipf’s law: we find relatively few common and many rare ones. By using an algorithm for inverse folding, we show that sequences sharing the same structure are distributed randomly over sequence space. All common structures can be accessed from an arbitrary sequence by a number of mutations much smaller than the chain length. The sequence space is percolated by extensive neutral networks connecting nearest neighbours folding into identical structures. Implications for evolutionary adaptation and for applied molecular evolution are evident: finding a particular structure by mutation and selection is much simpler than expected and, even if catalytic activity should turn out to be sparse in the space of RNA structures, it can hardly be missed by evolutionary processes.

Figure 4. Neutral paths. A neutral path is defined by a series of nearest neighbour sequences that fold into identical structures. Two classes of nearest neighbours are admitted: neighbours of Hamming distance 1, which are obtained by single base exchanges in unpaired stretches of the structure, and neighbours of Hamming distance 2, resulting from base pair exchanges in stacks. Two probability densities of Hamming distances are shown that were obtained by searching for neutral paths in sequence space: (i) an upper bound for the closest approach of trial and target sequences (open circles) obtained as endpoints of neutral paths approaching the target from a random trial sequence (185 targets and 100 trials for each were used); (ii) a lower bound for the closest approach of trial and target sequences (open diamonds) derived from secondary structure statistics (Fontana \textit{et al.} 1993\textsuperscript{a}; see this paper, §4); and (iii) longest distances between the reference and the endpoints of monotonously diverging neutral paths (filled circles) (500 reference sequences were used).

Reference for postulation and \textit{in silico} verification of \textit{neutral networks}
Coworkers

Walter Fontana, Santa Fe Institute, NM

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

Peter Stadler, Universität Leipzig, GE

Ivo L. Hofacker, Christoph Flamm, Universität Wien, AT

Bärbel Stadler, Andreas Wernitznig, Universität Wien, AT

Michael Kospach, Ulrike Langhammer, Ulrike Mückstein, Stefanie Widder

Jan Cupal, Kurt Grünberger, Andreas Svřek-Seiler, Stefan Wuchty

Ulrike Göbel, Institut für Molekulare Biotechnologie, Jena, GE

Walter Grüner, Stefan Kopp, Jaqueline Weber