Life – A Result of Evolution or Design?

Peter Schuster

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

Meeting of the Honda Foundation

"... Evolution in the sense of common ancestry might be true, but evolution in the Neo-Darwinian sense - an unguided, unplanned process of random variation and natural selection - is not. Any system of thought that denies or seeks to explain away the overwhelming evidence for design in biology is ideology, not science.

... Scientific theories that try to explain away the appearance of design as the result of 'chance and necessity' are not scientific at all, but ... an abdication of human intelligence."


*English translation:*

*Creation and Evolution.*

1. Biology and probabilities
2. Evolution – organismic and molecular
3. Multiplication, mutation, and selection
4. Rational design of molecules
5. Evolution and optimization of molecules
6. Origin of biological complexity
1. Biology and probabilities

2. Evolution - organismic and molecular

3. Multiplication, mutation, and selection

4. Rational design of molecules

5. Evolution and optimization of molecules

6. Origin of biological complexity
Polymer chain of 153 amino acid residues with the sequence:

```
GLSDGEWQLVLNVWGKVEADIPGHGQEVLLIRLFLKGPLPETLEKFDFKHLK
SEDEMKASEDLKKHGATVLTALGGILKKGKHGHEAEIKPLAOSHATKHKP
VKyleFISECIIQVLQSKHPGDAQGAMNKALELFKDMASNYKELG
FOG
```

The myoglobin molecule
Eugene Wigner’s or Fred Hoyle’s argument applied to myoglobin:

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

GLSDGEWQLVLNVWG.....FQG

Alphabet size: 20

Chain length: 153 amino acids

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

Probability to find the myoglobin sequence:

$20^{-153} = 9 \times 10^{-200} = 0.000\ldots009$
Eugene Wigner’s and Fred Hoyle’s arguments revisited:

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

Alphabet size: 20
Chain length: 153 amino acids
Length of longest path to myoglobin sequence: $19 \times 153 = 2907$
Probability to find the myoglobin sequence: $0.00034$
The folding problem of the myoglobin molecule:

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

\[ 15^{153} = 0.9 \times 10^{180} \text{ conformations}. \]

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

The Levinthal paradox of protein folding
Solution to Levinthal’s paradox

The gulf course landscape

Solution to Levinthal’s paradox

The funnel landscape

Solution to Levinthal’s paradox

The structured funnel landscape

Computed folding routes for guanine nucleotide binding (G) protein

An “all-roads-lead-to-Rome” landscape

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

1. Biology and probabilities
2. Evolution – organismic and molecular
3. Multiplication, mutation, and selection
4. Rational design of molecules
5. Evolution and optimization of molecules
6. Origin of biological complexity
Genotype, Genome

Collection of genes

Developmental program

Highly specific environmental conditions

Unfolding of the genotype

Phenotype

Evolution explains the origin of species and their interactions
Genotype, Genome

GCGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGATCTGGAGGTCCTGTGTTCGATCCACAGAATTCGCACCA

Quantitative biology
‘the new biology is the chemistry of living matter’

Biochemistry
molecular biology
structural biology
molecular evolution
molecular genetics
systems biology
bioinformatics
epigenetics

Highly specific environmental conditions

Phenotype

Unfolding of the genotype

The exciting RNA story
evolution of RNA molecules,
ribozymes and splicing,
the idea of an RNA world,
selection of RNA molecules,
RNA editing,
the ribosome is a ribozyme,
small RNAs and RNA switches.

Molecular evolution
Linus Pauling and Emile Zuckerkandl

Hemoglobin sequence
Gerhard Braunitzer

Manfred Eigen

John Kendrew

James D. Watson und
Francis H.C. Crick

Quantitative biology
‘the new biology is the chemistry of living matter’

Biochemistry
molecular biology
structural biology
molecular evolution
molecular genetics
systems biology
bioinformatics
epigenetics

Highly specific environmental conditions

Phenotype

Unfolding of the genotype

The exciting RNA story
evolution of RNA molecules,
ribozymes and splicing,
the idea of an RNA world,
selection of RNA molecules,
RNA editing,
the ribosome is a ribozyme,
small RNAs and RNA switches.

Molecular evolution
Linus Pauling and Emile Zuckerkandl

Hemoglobin sequence
Gerhard Braunitzer

Manfred Eigen

John Kendrew

James D. Watson und
Francis H.C. Crick
Three necessary conditions for Darwinian evolution are:

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

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

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

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

Insertion
Point mutation

Insertion

Deletion
Reconstruction of phylogenies through comparison of molecular sequence data
Results from molecular evolution:

• The molecular machineries of all present day cells are very similar and provide a strong hint that all life on Earth descended from one common ancestor (called „last universal common ancestor“, LUCA).

• Comparison of DNA sequences from present day organisms allows for a reconstruction of phylogenetic trees, which are (almost) identical with those derived from morphological comparison of species and the paleontologic record of fossils.
1. Biology and probabilities
2. Evolution – organismic and molecular
3. Multiplication, mutation, and selection
4. Rational design of molecules
5. Evolution and optimization of molecules
6. Origin of biological complexity
Complementary replication is the simplest copying mechanism of RNA. Complementarity is determined by Watson-Crick base pairs:

\[ G \equiv C \quad \text{and} \quad A = U \]
Mutation as an error in replication
Chemical kinetics of replication and mutation as parallel reactions
Formation of a quasispecies in sequence space
Formation of a quasispecies in sequence space
Formation of a quasispecies in sequence space

- **Master sequence**
- **Mutant cloud**
Formation of a quasispecies in sequence space

- Master sequence
- Mutant cloud
Uniform distribution in sequence space
Quasispecies

Driving virus populations through threshold

The error threshold in replication
Results from the kinetic theory of molecular evolution:

• Replicating ensembles of molecules form stationary populations called quasispecies, which represent the genetic reservoir of asexually reproducing species.

• For stable inheritance of genetic information mutation rates must not exceed a precisely defined and computable error-threshold.

• The error-threshold can be exploited for the development of novel antiviral strategies.
1. Biology and probabilities

2. Evolution - organismic and molecular

3. Multiplication, mutation, and selection

4. Rational design of molecules

5. Evolution and optimization of molecules

6. Origin of biological complexity
Sequence and structure of phenylalanyl-transfer-RNA

\[ \Delta G = -20.20 \text{ kcal/mol} \]
ΔG = -22.90 (-21.90) kcal/mol
1. Trial

\[ \Delta G = -43.10 \text{ (-36.40) kcal/mol} \]
2. Trial

\[ \Delta G = -45.10 (-39.40) \text{ kcal/mol} \]
3. Trial

\[ \Delta G = -41.80 \text{ (-39.90) kcal/mol} \]
4. Trial

Target structure

\[ \Delta G = -40.70 \text{ kcal/mol} \]
1. Biology and probabilities
2. Evolution – organismic and molecular
3. Multiplication, mutation, and selection
4. Rational design of molecules
5. Evolution and optimization of molecules
6. Origin of biological complexity
Application of the serial transfer technique to RNA evolution in the test tube

Stock solution: Qβ RNA-replicase, ATP, CTP, GTP and UTP, buffer
An example of ‘artificial selection’ with RNA molecules or ‘breeding’ of biomolecules.
Formation of secondary structure of the tobramycin binding RNA aptamer with $K_D = 9 \text{ nM}$

The three-dimensional structure of the tobramycin aptamer complex

Results from laboratory experiments in molecular evolution:

• Evolutionary optimization does not require cells and occurs in molecular systems too.

• *In vitro* evolution allows for production of molecules for predefined purposes and gave rise to a branch of biotechnology.

• Direct evidence that neutrality is a major factor for the success of evolution.
1. Biology and probabilities
2. Evolution - organismic and molecular
3. Multiplication, mutation, and selection
4. Rational design of molecules
5. Evolution and optimization of molecules
6. Origin of biological complexity
Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

Three-dimensional structure of the complex between the regulatory protein \textit{cro-repressor} and the binding site on $\lambda$-phage B-DNA.
Sketch of a genetic and metabolic network
The reaction network of cellular metabolism published by Boehringer-Mannheim.
The citric acid or Krebs cycle (enlarged from previous slide).
**E. coli:**  
Genome length $4 \times 10^6$ nucleotides  
Number of cell types 1  
Number of genes 4460

*Four books, 300 pages each*

**Man:**  
Genome length $3 \times 10^9$ nucleotides  
Number of cell types 200  
Number of genes $\approx 30000$

*A library of 3000 volumes, 300 pages each*

Complexity in biology
WHAT IS A GENE?

The idea of genes as beads on a DNA string is fast fading. Protein-coding sequences have no clear beginning or end and a RNA key is a part of the information package, reports Helen Pearson.

"We've come to the realization that the genome is full of overlapping transcripts." — Phillip Karapetian

"The degree of complexity we've seen was not anticipated." — Helmut Seelenfreund

The difficulty to define the notion of "gene".

Helen Pearson,
Nature 441: 399-401, 2006
ENCODE stands for **EN**cyclopedia **O**f **D**NA **E**lements.

**ENCODE** Project Consortium.
Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.
Biology and complexity:

• Evolution does not design with the eyes of an engineer but uses available objects for new purposes.

• The tinkering or bricolage principle gives rise to new objects of increasing complexity.

• The increase of complexity in biological evolution is an empirical fact.
Web-Page for further information:

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