From Schrödinger's „What is Life?" to „All Life is Chemistry"

Peter Schuster, Institute for Theoretical Chemistry, University of Vienna, Austria and
The Santa Fe Institute, Santa Fe, New Mexico, USA

This afternoon we celebrate the 75 anniversary of the appearance of Erwin Schrödinger's book "What is Life? The Physical Aspect of the Living Cell". The booklet presents the material of a course of public lectures that where delivered by Schrödinger under the auspices of the Dublin Institute for Advanced Studies at Trinity College, Dublin, in 1943.

Schrödinger's "What is Life?"

"What is Life?" has been enormously successful, inspiring and influential.[1,2] Max Perutz writes 1987: "Up to 1948 it drew 65 reviews and has sold up to now about 100000 copies." An appreciable number of molecular biologist – among them the famous proposers of the double helical structure of DNA, James Watson and Francis Crick, Max Delbrück, Gunter Stent, Maurice Wilkins and Seymour Benzer – admitted that they were truly inspired and encouraged in their work through reading Schrödinger's book. Historians see mainly three reasons for the enormously positive appraisal of the book by the public, which is exceptional for a scientific publication even for a popular science writing: (i) the booklet is written in an elegant, lively and easy to follow, almost poetic style,[3] (ii) time was ripe for a rethinking of the scientific roots above which biology was built, and (iii) questions concerning the origin of life or likewise the origin of the universe have been and are of great public interest since the provide answers to the burning question: "Where are we – the mankind – coming from?".

Schrödinger's "What is Life?" came just in the right moment before the onset of the revolution in modern biology that introduced thinking in terms of molecular structures.[4] In Horace Judson historical and scientific treatise "The Eighth Day of Creation" Schrödinger's booklet is referenced eight times. The enthusiasm about "What is Life?" and the strong impact it had on young scientists, especially on physicists, strikingly contrasts with the evaluation of its scientific content by experts. Linus Pauling, Max Perutz and Francis Crick – all three Nobel laureates themselves – were very critical. Linus Pauling[5] was upset by Schrödinger's metaphor of organisms feeding on "*negentropy*". The main argument for his arousal is the fact that the energetic and entropic balance of the living cell was already understood when Schrödinger gave his lectures. Free energy rather than entropy is proper thermodynamic function in isothermal systems. Adenosine triphosphate (ATP), the "*energetic currency of life*", was known since 1929 through the discovery by Karl Lohmann. The free energy provided by ATP has a much larger energetic than entropic component and hence the negentropy metaphor is questionable already from pure thermodynamics. ATP synthesis and ATP hydrolysis have been studied with great care and here the ratio of the energetic and the entropic contribution is approximately nine.[6,7] The condensation of amino acids into a polypeptide change is accompanied by a certain decrease in free energy no

matter whether the sequence is ordered or random. The ordering of the sequence, Pauling argues, comes about through processes inside the cell like enzyme catalysis and template action and not through import of negentopy.

Max Perutz and Francis Crick criticize in particular the use of the term "aperiodic crystal". Macromolecules and polymers were known already since the early twenties through the works of Hermann Staudinger,[8] Hermann Mark[9] and others, and they are not crystals in the sense that they are flexible and have no solid state structure. Horace Judson says that the details of Schrödinger's science seemed to Francis Crick "almost embarrassingly gauche. 'I was not conscious of any influence of what he called the aperiodic crystal – I don't suppose the man had ever heard of a polymer!'." There is one important issue, which Schrödinger pointed out correctly: Whenever you have a sequence of some moderate length built from several classes of monomers, the number of possible combinations can easily fill the universe as a result of "*combinatorial complexity*". Two symbols are sufficient as we know from the Morse alphabet or computer codes.

Important and influential was Schrödinger's view that the chromosome carries the information for descendant cells of the future in encoded form together with the machinery to make the cell. Although Schrödinger put forward here the concept of a genetic code for the first time so clearly, Sydney Brenner, a molecular biologist of the first hour, was apparently unhappy with this formulation: [10]

> "I have come to call this „Schrödinger's fundamental error": '… The chromosome structures are at the same time instrumental in bringing about the development they foreshadow. They are code law and executive power, or to use another simile, they are the architect and the builder's craft in one. ' (Schrödinger, What is Life?, p.20). … And that is wrong! The chromosomes contain the information to specify the future organism and a description of the means to implement this, but not the means themselves."

John von Neumann presented a paper at the *Hinxon Symposium 1948* in Pasadena, California, where he compared the function of genes to self-reproducing automata.[11] Sydney Brenner was highly impressed by this presentation and wrote (Brenner, pp.33-36):

> "…Von Neumann shows that you have to have a mechanism not only of copying the *machine*, but of copying the *information* that specifies the machine. So he divided the machine--the *automaton* as he called it--into three components. (i) the functional part of the automaton, (ii) a decoding section which actually takes a tape, reads the instructions and builds the automaton, and (iii) a device that takes a copy of this tape and inserts it into the new automaton. …"

He then raises the claim that von Neumann's concept perfectly describes the principle of the genetic machinery of the cell and considers the fact that von Neumann and Watson and Crick presumably did not know anything about the other as a kind of historical irony.

A clear distinction between the code and the machinery is not splitting hairs. It is illustrative to visualize the difference between Schrödinger's and von Neumann's mechanism of inheritance: In the Schrödinger case multiplication would occur if chromosomes are injected into a solution with nutrients whereas von Neumann's self-reproducing automaton would need an intact cell. Later we shall make use of John von Neumann's theory to work out the difference between reproductions in the DNA-protein world or in a hypothetical RNA world, where both functions, coding and executive function are housed in the same molecule.

Structures of biological macromolecules

Molecular structures in chemistry before quantum mechanics and in particular, before the application of the Schrödinger-equation to problems in quantum chemistry were essentially all built by "hook and eye" models with the binding properties of atoms derived from empirical observations and from the periodic table. Schrödinger's wave equations turned out to be very useful for analysis and description of chemical bonds in small and medium size molecules. Linus Pauling[12] and Charles Coulson[13] among others made quantum chemistry popular and the success in the applications led to the famous statements,

> "The fundamental laws necessary for the mathematical treatment of **a large part of physics and the whole of chemistry** are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved." by Dirac,[14] and

> "There is no doubt that the Schrödinger equation provides the **theoretical basis of chemistry**." by Pauling.[5]

The nature of chemical bonds was correctly understood as a quantum mechanical property. Chemists correlate structure with reactivity and function and therefore, many and strong efforts were and are undertaken to determine precise molecular structures. The molecules in the core of the biology of the cell, in particular proteins and later also nucleic acids were recognized as linear polymers with a periodic molecular backbone and side chains provided by several classes of monomers, twenty amino acids in case of proteins. First, the known structures of small units were combined through model building and Pauling's α-helix of polypeptides and proteins was the first triumph of structure prediction by model building.[15] The most influential and most spectacular prediction of a biopolymer structure from X-ray diffraction data of fibers is the proposal of the structure of deoxyribonucleic acid (DNA) by Watson and Crick.[16] The DNA structure suggests two possible mechanisms for biological key processes: (i) the duplication of the genetic material as expressed by the dicta,

> "It has not escaped our notice that the specific pairing we have postulated immediate suggests a possible copying mechanism for the genetic material."[16],

and (ii) the simplest kind of mutations called point mutation, which consists of the replacement of a single nucleotide letter. Enzyme-free copying of poly-deoxyribonucleotides is a highly inefficient and highly inaccurate process. An enzyme, a thermostable DNA polymerase from the bacterium *Thermus aquaticus* (Taq) for example, in needed to render template copying a useful tool for multiplication of DNA molecules (see later).

The elucidation of the DNA structures ignited a true revolution in biology. About the same time the first complete structures of proteins at molecular resolution were published[17,18] and they revealed Pauling's α-helices as a common substructure of globular proteins. The beginnings of molecular biology are intimately connected to the determination of biomolecular structures by means of X-ray analysis of crystals. Fast technical progress in crystal structure determinations, in particular by the upcoming use of computers for the extensive calculation of Fourier synthesis, made soon the growth of sufficiently large single crystals the time limiting step for protein structure determination.

Molecular biologists were tuned to search for a code relating DNA and protein not least inspired by Schrödinger's "What is life?". In relatively short time the code has been recognized as a triplet code[19] – three nucleotides for one amino acid – and accordingly substantial degeneracy in coding the twenty *natural* amino acids had to be expected and was found – there are 4^3=64 different codon triplets for four nucleotides. The assignment of all codons to the twenty individual amino acids[20] has been deciphered in a few years.[21] The next, important finding was that the code is universal in all organisms except a small number of modifications that were introduced later into the standard code.[22] It is important that every message can be interpreted by the genetic machinery: Three "*nonsense triplets*" code for the end of the polypeptide chain. Together with the discovery of bacterial gene regulation through François Jacob and Jacques Monod,[23] the at first still oversimplified but nevertheless complete dynamic image of the primitive cell was finished.

All structures and processes could be and can be fully understood and interpreted in terms of chemistry: Proteins as highly versatile and enormously specific are acting as catalysts for the various reactions in the cell and nucleic acid chemistry illuminated the relation between genetics and protein synthesis. Therefore molecular biology was received with great enthusiasm by the scientific community and also by the public. An illustrative example is an Austrian TV-production in ten parts with the title

"All Life is Chemistry".

The author and historian Hellmut Andics made the production in the year 1978 and the famous and already mentioned polymer chemist Herrmann Mark presented it. Public interest was enormous. A simple sketch of the cellular machinery is shown in the figure.

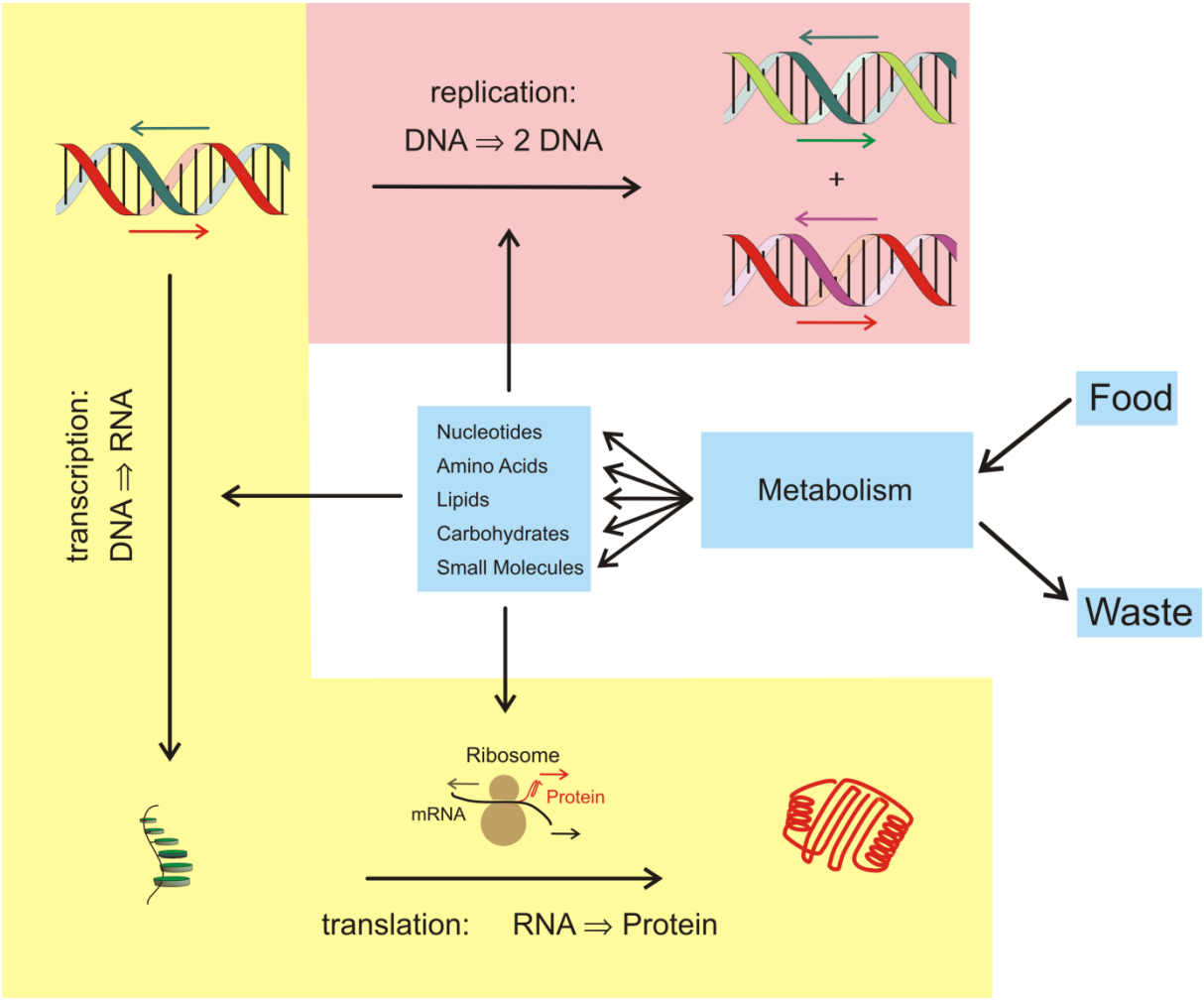

Figure: **A sketch of the core process in the cell**. Three different tasks of the living cell are distinguished: (i) multiplication of the genetic material (pink), (ii) synthesis of protein molecules (yellow), and (iii) metabolism (blue). Practically all cellular reactions are catalyzed by enzyme molecules or protein complexes. Prokaryotic DNA replication, for example, is a highly complex process involving at least nine proteins with different functions. Cellular protein synthesis occurs in two groups of processes: (i) Transcription produces a complementary RNA copy of a stretch of nucleotides on one or the other strand of the DNA and (ii) translation synthesizes a protein from the RNA-template at the ribosome making use of the genetic code. Metabolism provides the various building blocks for the synthesis of biopolymers. The cellular machinery is transmitted in encoded form in the reproduction process and therefore an intact cell is required for multiplication.

What is different in chemistry and biology?

Schrödinger's dream of new physical laws to be detected in biology did not materialize, at least not until now. Is biology then nothing but chemistry with larger molecules built from a subset of atoms of the periodic table? Is the cell a chemical factory much more complex and capable of reproduction in the sense of John von Neumann's self-reproducing automaton or are there basic features in biology that are unknown or at least uncommon in chemistry. No doubt one could list a great number of such features but I shall restrict myself to three of them, which are related to Schrödinger's "What is Life?" in this contribution: (i) biological evolution, (ii) complexity of molecular structures, and (iii) digitalization of chemical information.

The first characteristic property of biological entities – viruses, bacteria, cells, and higher organisms – is evolution as was nicely phrased by Theodosius Dobzhansky:[24,25]

"Nothing in biology makes sense except in the light of evolution".

Inevitably biology has a historical component whereas chemistry does not: Every organism carries information on his phylogeny and the application of the theory of chemical structure analysis in the form of DNA sequence comparisons in molecular evolution turned out extremely useful in the reconstruction of phylogenetic trees. The basis for evolution is the interplay of variation and selection in the multiplication process over many generations as expressed by the term of *natural selection* created by Charles Darwin[26] and Alfred Russel Wallace.[27] Variation in nature is the result of two processes: (i) mutation, a change in the nucleotide sequence, and (ii) recombination, reshuffling of gene variants. The only quantity that is relevant for selection is the number of descendants in forthcoming generations commonly denoted as *fitness*. Natural selection can be easily cast into a mathematical formalism, which is based on the mathematics of two processes: (i) reproduction with limited resources in form of the logistic or Verhulst equation[28] and (ii) chemical kinetics of simple replication in a population (for mathematical details see the Appendix).[29] The Verhulst equation,

$$\frac{dX}{dt} = f\,X\left(1-\frac{X}{C}\right) \quad \Rightarrow \quad X(t) = \frac{C\,X_0}{X_0 + (C - X_0)\,\exp(-f\,t)};\; X_0 = X(0),$$

describes the kinetics of growth in the world where resources are always finite. Chemical kinetics of replication in a heterogeneous population of limited size $\Pi = \{X_1, X_2, \ldots, X_n\}$ leads to selection:

$$\frac{d\xi_j}{dt} = \xi_j\,(f_j - \Phi);\; \Phi = \sum_{i=1}^{n} f_i\,\xi_i \;\Rightarrow\; \xi_i(t) = \frac{\xi_j(0)\exp(f_j t)}{\sum_{i=1}^{n}\xi_i(0)\exp(f_i t)};\; \xi_i(t) = \frac{X_i}{\sum_{i=1}^{n} X_i};\; \sum_{i=1}^{n}\xi_i = 1.$$

In the long time limit we find:

$$\Pi = \{X_m\} \quad \text{or} \quad \lim_{t\to\infty} \xi_m(t) = 1 \text{ and } \lim_{t\to\infty} \xi_{i\neq m}(t) = 0,$$

that implies selection of X_m, which is the genotype of largest fitness. Although all required mathematics has been known long before the publication of the "Origin of species" and the derivation of the selection equation is straightforward apparently no biologist tried to develop a mathematical frame for natural selection before the works of the population geneticists.[30,31]

An often raised question concerns optimality of the entities produced by an evolutionary process. For the survival in generations of the future optimality as such is not needed. Successful competition requires only to be better, i.e. to have more progeny, than the competitors. In contrast to engineers nature cannot redesign from scratch but has to build from the existing materials with or without minor modifications. François Jacob put this fact in an illustrative phrase:[32,33]

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

More recent research has shown that the evolutionary process is more constrained by molecular requirements than a notion like "tinkering" or "bricolage" indicates. [34]

The second characteristic of cellular catalysts, proteins or RNA molecules, is enormous structural complexity, which is otherwise never found in chemistry. Large quantitative differences may become qualitative differences as Phil Anderson points out in his famous essay. "More is different".[35] Two natural "molecular machines" performing biological core processes may serve here as examples: the replication fork of DNA and the ribosome, which performs the translation of messenger RNA into protein. The DNA replication machinery of the cell involves enzymes from – at least –six classes, RNA primers and single strand binding proteins. Both strands, which in the double helix are running in different directions, 5'-end $\rightarrow$ 3'-end and 3'-end $\rightarrow$ 5'-end, respectively, are completed to double helices simultaneously leading to two daughter molecules, each one containing a new and an old strand. Interestingly, nature did not develop two classes of DNA polymerases that can operate in one or the other direction, instead the "lagging strand" is synthesized in small pieces called Okazaki fragments in the direction opposite to the progression of the fork and the fragments are ligated – joined – afterwards. This is at the same time an impressive and highly efficient molecular machinery and a wonderful example of tinkering! The second example of a molecular factory is the ribosome, a large complex of RNA molecules and proteins synthesizing a polypeptide chain through translation of a messenger RNA.[36,37] The link between the coding messenger-RNA and the amino acid residues is made by a class of RNA molecules, the transfer-RNAs. The catalytic activity of the ribosome is exerted by the ribosomal RNA molecules whereas the proteins fulfill the task to hold the RNA in precise spatial positions.[38]

In chemistry the notion of information is often used for interactions in molecular complexes. Binding partners "carry information" on the other molecules in the complex in order to be able to fit into the molecular arrangement. I personally would prefer to reserve information for coding information in the sense of Claude Shannon's theory of information.[39] The third example of a novelty in biology versus chemistry, the digitalization of chemical information, is chosen accordingly. In isolation, the two base pairs, **A**=**T** and **G**≡**C**, are of very different strengths and the binding constants differ by one order of magnitude: **A** and **T** are connected by two hydrogen bonds, **G** and **C** by three, and accordingly the latter pair is much stronger. Nevertheless, the two pairs appear more or less equivalent in the double helix. Since hydrogen bonding leads to different stabilities in the base pairs, Nature apparently does not use hydrogen bond energies for the stabilization of double helices. The role of the hydrogen bonds is to provide specificity that is exerted through the base pair geometries. Careful studies[40] revealed indeed that base pairing contributes very little to the stability of DNA double helices. This fact that hydrogen bonding is not effective in the stabilization of the DNA double helix is easy to explain: What counts is the free energy different between the bound and the unbound state, where hydrogen bonding to the solvent water occurs. The stabilizing factor is base pair stacking but, as said, the hydrogen bonds are important in their own right, because they determine the base pairs geometry and are responsible for fitting the base pairs into the structure of the double helix accordingly.

The difference in binding strengths is even more drastic in case of the codon-anticodon interactions: **G**,**C**-rich triplets are bound much stronger than the corresponding **A**,**T**-rich codon-anticodon systems but there is apparently no such difference on the ribosome since the base composition of the messenger-RNA has very little or no influence on efficiency, speed and accuracy of translation. Evolution has created a molecular environment where the two base pairs are practically equivalent. The gradualism of thermodynamic strength of interactions in chemistry is replaced by **yes-or-no decisions** – two nucleotides form or don't form a base pair. Processing of genetic information has a lot in common with message passing in Shannon's theory.[39] In processing of genetic information chemistry has been truly digitalized or in other words, the cell provides an environment that allows for readout of nucleotide sequences as if it were written in a language with equivalent digits.

Bridging from chemistry to biology

There are many ways to bring chemistry and biology into the same context. Here we shall sketch an origin of life model, which originates from prebiotic chemistry on Earth and leads to a scenario that allows for the onset of Darwinian evolution. Discoveries of Thomas Cech[41,42] and Sydney Altman[43] that have been rewarded with the Nobel Prize 1989 in chemistry were fundamental for understanding early evolution: RNA molecules called ribozymes can catalyze reactions very much like proteins do. An RNA molecule can be doing both, coding and catalyzing simultaneously. RNA catalyzing template induced production of RNA[44] or RNA self-replication in a way is possible. A new combination of structures and properties allowed for an important simplification in von Neumann's theory of self-reproducing automata. The tape itself can form the machine that is reading the tape! In biological language this feature boils down to a class of molecules, which are genotype – legislative – and phenotype – executive – at the same time. In principle the concept of an RNA polymerase, which replicates an RNA template, appears simple but the search for such a molecule turned out to be highly involved.[45] The pendant to self-reproduction in chemistry is autocatalysis and also here simple autocatalytic systems are very rare. Commonly autocatalysis is exhibited by the overall kinetics of complex reaction networks.

The notion of a prebiotic RNA world,[46] which came after a complex mixture of small and medium size organic molecules produced by a network of reactions taking place predominantly in aqueous solution and preceded our present RNA+protein+DNA world, turned out to be very fruitful in many fields from origin of life studies to applications in biotechnology.[47] The process in the core of the RNA world is RNA-template induced polymerization of RNA or in other words RNA self-replication. As said RNA replication by ribozymes is very difficult to achieve.[43] Later in evolution Nature has developed protein enzymes doing this job in RNA virus infection of bacteria: a single enzyme synthesizes the complementary strand of an RNA template whereby two tasks have to be accomplished simultaneously: (i) the accuracy of incorporation of the correct complementary nucleotide has to be sufficiently high to allow for passably correct reproduction and (ii) the separation of the template and the newly synthesized strand requires catalytic help. Otherwise a double stranded RNA molecule would be formed that requires heating for separation of the two strands like it is done in DNA multiplication through polymerase chain reaction (PCR). Charles Weissmann[48] suggested a mechanism in which nucleotide complementarity is used during the incorporation of the incoming nucleotide in order to guarantee a sufficiently low error rate and then, the two strands are separated and each one folds into its specific single strand structure thereby avoiding double helix formation. The entire multistep mechanism has been explored in case of RNA replication by the bacteriophage specific enzyme Qβ-replicase.[49]

Darwinian evolution is based on two requirements: (i) selection and (ii) variation. Selection as we have seen in the case of the Verhulst equation takes place automatically when self-reproducing elements compete in a finite world the only thing that is required is a source of building blocks for the synthesis of new molecules and the replication machinery in order to resupply consumed materials. Variation occurs inevitably since no real process can take place with ultimate accuracy – errors cannot be completely avoided in the real world. Quantitatively it is necessary that the error rate lies in an appropriate range (see later). All features are naturally fulfilled in RNA replication by RNA template specific RNA-polymerases. Sol Spiegelman[50,51] recognized that a system based on the bacteriophage Qβ fulfils all requirements for cell-free Darwinian evolution and performed the first *in vitro* evolution experiments. In the light of the title of this section: "Cell-free evolution builds a bridge between chemistry and biology". Now fifty-two years later *in vitro* evolution has come of age and many exciting experiments were done.[52]

About the same time when Spiegelman started to perform his *in vitro* evolution experiments Manfred Eigen developed a kinetic theory describing the evolution of molecules in the languages of chemical kinetics and molecular biology.[53] Eigen's 1971 publication initiated a rich body of theoretical and experimental work on the evolution of molecules and led to two new concepts: (i) the quasispecies[54,55] and (ii) the hypercycle.[56,57] A quasispecies is the stationary mutant distribution of an asexually multiplying population. It represents the genetic reservoir created by asexual reproduction and consists of a most frequent master sequence and its closely related and sufficiently fit neighbors in sequence space. Commonly but not necessarily, the master sequence is also the fittest sequence in the population (for mathematical details see the Appendix). Hypercycles are catalytic systems combining two classes of catalytic actions: template induction and catalysis of reactions. Hypercycles show dynamical features the are typical for higher order autocatalysis like "once forever" decisions or frozen accidents, oscillations and deterministic chaos, and formation of Turing patterns.

The molecular quasispecies[53] turned out to be a very useful concept not only in modeling test tube evolution of molecules but also in describing virus evolution, bacterial evolution and evolution of transformed cellular clones in cancer.[58] Here two quantitative relations of mutation rates and quasispecies structure will be highlighted: (i) the error threshold of error prone replication and (ii) the relation between mutation rates and genome lengths. An increase in the error rate per nucleotide and generation, p, leads to more mutants in the stationary population. For constant chain lengths ℓ of the polynucleotides a sharp threshold at a certain error rate p_{crit} is observed above which no faithful reproduction is possible:

$$p_{\max} \approx \frac{\ln \sigma_m}{\ell} \quad \text{with} \quad \sigma_m = \frac{(1-\bar{\xi}_m)\, f_m}{\sum_{j \neq m} \bar{\xi}_j\, f_j} \quad \text{and} \quad \sum_{i=1}^{n} \bar{\xi}_i = 1 .$$

The existence of an error threshold has been used in medical applications to drive populations towards extinction through accumulation of incorrect variants or lethal mutagenesis. Special drugs were developed, which interfere with nucleotide pair matching during the replication process and increase the mutation rate thereby.[56]

When the mutation rate is constant because of the accuracy of one and the same replication machinery, sufficiently accurate reproduction sets limits to the chain lengths of the sequences in stable populations. Since the logarithm of the superiority factor of the master sequence is positive and often near one. An over-the-thumb rule assumes that the mutation rate cannot strongly exceed the reciprocal chain length:

$$\ell_{\max} \approx \frac{\ln \sigma_m}{p} \quad \text{with} \quad \sigma_m = \frac{(1-\bar{\xi}_m)\, f_m}{\sum_{j \neq m} \bar{\xi}_j \, f_j} \quad \text{and} \quad \sum_{i=1}^{n} \bar{\xi}_i = 1$$

A paper in *Science* showed[59] that the replication accuracy relation versus genome length relation holds approximately over seven orders of magnitude from viroids to higher eukaryotes including mammals.

Thomas Cech[45] distinguishes three different notions of “RNA worlds”. All three, one way or the other, build bridges from chemistry to biology. The “prebiotic RNA world” closes the gap between prebiotic chemistry and the simplest molecular systems that are capable of evolution. “RNA-world two” is the fascinating multiple task RNA plays in present day cells as a mediator between protein chemistry and DNA genetics. “RNA world three”, finally, is the world of RNA technology and applications. The almost unlimited diversity of RNA structures provides the key to design molecules for predefined purposes. Since RNA can carry digital information at the same time, it can be evolved to exert predefined functions and to fulfill given tasks.

Appendix: The mathematics of natural selection without and with mutation

Natural selection can be easily cast into a mathematical formalism that is based on limited resources in form of the logistic or Verhulst equation[28]

$$\frac{dN}{dt} = f\,N\left(1-\frac{N}{C}\right) \quad \text{and} \quad N(t) = \frac{N(0)\,C}{N(0)+\left(C-N(0)\right)\exp(-f\,t)},$$

where $N(t)$ is the size of a homogeneous population Π = {**X**} with [**X**] = N, C is the carrying capacity of the ecosystem and f the fitness or Malthus parameter, and chemical kinetics of simple reproduction:

$$\mathbf{A} + \mathbf{X} \rightarrow 2\,\mathbf{X} \quad \text{and} \quad \frac{dX}{dt} = k\,A\,X = f\,X \;\; ; A=[\mathbf{A}],\; X=[\mathbf{X}].$$

The symbol **A** stands for the resource, which is necessary to build a molecule **X** that is capable of self-reproduction. The fitness of the variant **X** is denoted by $f = k\,A$. In the original Verhulst equation constant fitness corresponding to a constant concentration [**A**] = A_0 is assumed. Limitation of population growth is controlled by the quadratic term.

In a structured population Π = {X_1,X_2,…,X_n} the same approach leads to an extended Verhulst equation:[31]

$$\frac{dX_j}{dt} = X_j\left(f_j - \frac{N}{C}\phi(t)\right) \;\text{ with }\; \phi(t) = \frac{1}{N}\sum f_i\,X_i\,.$$

The solution of the extended Verhulst equation is performed in two steps:
(i) the time dependence of the population size,

$$N(t) = \frac{N(0)\,C}{N(0)+\left(C-N(0)\right)\exp\left(-\Phi(t)\right)},$$

with Φ(t) being the time integral of the mean fitness

$$\Phi(t) = \int_{\tau=0}^{t} \frac{\sum_{i=1}^{n} f_i\,X_i(t)}{N(t)}\,d\tau = \int_0^t \sum_{i=1}^{n} f_i\,\xi_i(\tau)\,d\tau = \int_0^t \phi(\tau)\,d\tau\,,$$

where ξ_j(t) is the normalized concentration, ξ_j(t) = Xj(t) / N(t) , and
(ii) the time dependence of the internal structure of the population, which can be calculated for normalized concentrations ξ_j(t) = X_j(t) / N(t) :

$$\xi_j(t) = \frac{\xi_j(0)\exp(f_j\,t)}{\sum_{i=1}^{n}\xi_i(0)\exp(f_i\,t)}.$$

The solution (ii) is required for the calculation of (i).

Error free replication and mutation are considered as parallel chemical reactions:

$$\frac{\mathrm{d}X_j}{\mathrm{dt}} = \sum_{i=1}^{n} W_{ji}\, X_i - X_j\, \phi\,; \quad j = 1,2,\ldots,n; \phi = \frac{\sum_{i=1}^{n} f_i X_i}{\sum_{i=1}^{n} X_i} = \sum_{i=1}^{n} f_i\, \xi_i\,.$$

Factorization of the value matrix $\mathrm{W} = \mathrm{Q} \cdot \mathrm{F}$ separates effects of fitness and mutation:

$$\mathrm{W} = \begin{pmatrix} W_{11} & W_{12} & \ldots & W_{1n} \\ W_{21} & W_{22} & \ldots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & W_{n2} & \ldots & W_{nn} \end{pmatrix} = \mathrm{Q} \cdot \mathrm{F} \quad \text{with}$$

$$\mathrm{Q} = \begin{pmatrix} Q_{11} & Q_{12} & \ldots & Q_{1n} \\ Q_{21} & Q_{22} & \ldots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \ldots & Q_{nn} \end{pmatrix} \quad \text{and} \quad \mathrm{F} = \begin{pmatrix} f_1 & 0 & \ldots & 0 \\ 0 & f_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f_n \end{pmatrix}.$$

The mutation-selection equation takes on the form:

$$\frac{d\xi_j}{dt} = \sum Q_{ji}\, f_i\, \xi_i - \xi_j\, \phi;\; j = 1,2,\ldots,n; \phi = \sum f_i\, \xi_i = \bar{f}\,.$$

Solutions are obtained after an integrating factor transformation,

$$\zeta_j(t) = \xi_j(t) \exp\int_0^t \phi(\tau)\, d\tau\,,$$

By means of an eigenvalue problem:

$$\xi_j(t) = \frac{\sum_{k=0}^{n-1} \ell_{jk}\, c_k(0) \exp(\lambda_k t)}{\sum_{i=1}^{n} \sum_{k=0}^{n-1} \ell_{ik}\, c_k(0) \exp(\lambda_k t)} \quad \text{with}$$

$$\mathrm{L} = \left\{ \ell_{ij}\,; i, j = 1,2,\ldots,n \right\} \quad \text{and} \quad \mathrm{L}^{-1} \cdot \mathrm{W} \cdot \mathrm{L} = \Lambda = \left\{ \lambda_k\,; k = 0,1,\ldots,n-1 \right\}.$$

The dominant eigenvalue is λ_0 and the corresponding eigenvector represents the quasispecies.

References:

[1] C.E.Margo, L.E.Harman. The remarkable life of Erwin Schrödinger's What Is Life? The Pharos Alpha Omega Alpha Honor Med Soc. 2015 Autumn; 78(4):19-23, 2015.

[2] K. R. Dronamraju. Erwin Schrödinger and the Origins of Molecular Biology.Genetics 153:1071-1076, 1999

[3] M. F. Perutz. Schrödinger's What is Life? and molecular biology. In: C.W. Kilmister. Schrödinger. Centenary celebration of a polymath. Cambridge University Press, New York 1987, chap.19, pp. 234-251.

[4] H. F. Judson. The Eighth Day of Creation. Makers of the Revolution in Biology. Jonathan Cape, London 1979.

[5] L. Pauling. Schrödinger's contribution to chemistry and biology. In: C.W. Kilmister. Schrödinger. Centenary celebration of a polymath. Cambridge University Press, New York 1987, chap.18, pp. 225-233.

[6] R. Milo, R. Phillips. Cell biology by the numbers. Garland Science Group at Taylor & Francis. New York 2017.

[7] O. Pänke, B. Rumberg. Energy and entropy balance of ATP synthesis. Biochimica et Biophysica Acta 1322:183-194, 1997

[8] H. Staudinger, J. Fritschi. Über Isopren und Kautschuk, 5. Mitt. Über die Hydrierung des Kautschuks und über seine Konstitution. Helvetica Chimica Acta 5(5):785-806, 1922.

[9] H. F. Mark. Aus den frühen Tagen der Makromolekularen Chemie. Naturwissenschaften 67(10):477-483, 1980.

[10] S. Brenner. My Life in Science. BioMed Central Ltd. New York 2001, pp.33-34.

[11] J. von Neumann. Theory of Self-Reproducing Automata. Edited and completed by A. W. Burks. University of Illinois Press, Urbana and London 1966.

[12] L. Pauling. The nature of the chemical bond and the structure of molecules and crystals. An introduction to modern structural chemistry. Cornell University Press. Ithaca, New York 1939.

[13] C.A. Coulson. Valence. Oxford at the Clarendon Press. Oxford, UK 1952.

[14] P.A.M. Dirac. Quantum mechanics of many-electron systems. Proc.Roy.Soc.London A 123:714-733, 1929.

[15] L. Pauling, R.B. Corey, H.R. Branson. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. Proc.Natl.Acad.Sci.USA 37(4):205-211, 1951.

[16] J.D. Watson, F.H.C. Crick. A structure for deoxyribonucleic acid. Nature 171(4356):737-738, 1953.

[17] J.C: Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. Nature 181(4610):662-666, 1958.

[18] M.F. Perutz, M.G. Rossmann, A.F. Cullis, H. Muirhead, G. Will, A.C.T. North. Structure of hemoglobin: A three-dimensional Fourier synthesis at 5.5 Å resolution, obtained by X-ray analysis. Nature 185(4711):416-422, 1960.

[19] M.W. Nirenberg, J.H. Matthaei. The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polynucleotides. Proc.Natl.Acad.Sci.USA 47(10):1588-1602, 1961.

[20] M.W. Nirenberg, W. Jones, P. Leder, B.F.C. Clark, W.S. Sly, S. Pestka. On the coding of genetic information. Cold Spring Harb Symp Quant Biol 28:549-557, 1963.

[21] M.W. Nirenberg. Historical review: Deciphering the genetic code – a personal account. Trends in Biological Sciences 29(1):46-54, 2004.

[22] R.D. Knight, S.J. Freeland, L.F. Landweber. Rewiring the keyboard: Evolvability of the genetic code. Nature Reviews Genetics 2(1):49-58, 2001.

[23] F. Jacob, J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. J.Mol.Biol. 3(3):318-356, 1961

[24] T. Dobzhansky. Nothing in biology makes sense except in the light of evolution. The American Biology Teacher 35(3):125-129, 1973.

[25] T. Dobzhansky. Biology, molecular and organismic. American Zoologist 4:443-452, 1974.

[26] C. Darwin. On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life. John Murray. London 1859.

[27] A.R. Wallace. Contributions to the Theory of Natural Selection. Second Ed., Macmillan & Co. London 1870.

[28] P, Verhulst. Notice sur la loi que la population pursuit dans son accroisement. Corresp.Math.Phys. 10:113-121, 1838.

[29] P. Schuster. Mathematical modeling of evolution. Solved and open problems. Theory Biosci. 130:71-89, 2011.

[30] J.F. Crow. Population genetics history: A personal view. Annu.Rev.Genetics 21:1-22, 1987.

[31] P. Schuster. Mathematical modeling of evolution. Solved and open problems. Theory Biosci. 130:71-89, 2011.

[32] F. Jacob. Evolution and tinkering. Science 196(4295):1161-1166, 1977.

[33] D. Duboule, A.S. Wilkins. The evolution of 'bricolage'. Trends in Genetics 14(2):54-59, 1998

[34] A.S. Wilkins. Between „design" and "bricolage": Genetic networks, level of selection, and adaptive evolution. Proc.Natl.Acad.Sci.USA 104:8590-8596, 2007.

[35] P.W. Anderson. More is different. Broken symmetry and the nature of the hierarchical structure of science. Science 177(4047):393-396, 1972.

[36] N. Ban, P. Nissen, J. Hansen, P.B. Moore, T.A. Steitz. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science 289(5481):905-920, 2000.

[37] P. Nissen, J. Hansen, N. Ban, P.B. Moore, T.A. Steitz. The structural basis of ribosome activity in peptide bond synthesis. Science 289(5481):920-930, 2000.

[38] T.R. Cech. The ribosome is a ribozyme. Science 289(5481):878-879, 2000.

[39] C.E. Shannon, W. Weaver. The mathematical theory of communication. University of Illinois Press. Urbana-Champaign 1949.

[40] P. Yakovchuk, E. Protozanova, M.D. Frank-Kamenetskii. Base stacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Research 34(2):564-574, 2006.

[41] T.R.Cech, A.J. Zaug, P.J. Grabowski. *In vitro* splicing of the ribosomal RNA precursor of Tetrahymena: Involvement of a Guanosine nucleotide in the excision of the intervening sequence. Cell 27:487-496, 1981.

[42] T.R. Cech. Self-splicing of group I introns. Annu.Rev.Biochemistry 59:543-568, 1990.

[43] C. Guerrier-Takada, K. Gardiner, T. March, N. Pace, S. Altman. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35:849-857, 1983.

[44] T.R. Cech. A model for the RNA-catalyzed replication of RNA. Proc.Natl.Acad.Sci.USA 83(12):4360-4363, 1986.

[45] G.F. Joyce, J.W. Szostak. Protocells and RNA self-replication. In: T.R. Cech., J. A. Steitz, J.F. Atkins, Additional perspectives on RNA worlds. Cold Spring Harb.Perspect.Biol. 10:a034801, 2018.

[46] W. Gilbert. Origin of life: The RNA world. Nature 319(6055):618, 1986.

[47] T.R. Cech. The RNA worlds in the context. In: J.F. Atkins, R.F. Gesteland, T.R. Cech. Additional perspectives on RNA worlds. Cold Spring Harb.Perspect.Biol. 4:a006724, 2012.

[48] C. Weissmann. The making of a phage. FEBS Letters 40(S1):S10-S18, 1974.

[49] C.K. Biebricher, M. Eigen, W.C. Gardiner, Jr. Kinetics of RNA replication. Biochemistry 23(10):3186-3194, 1983.

[50] 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(1):217-224, 1967

[51] S. Spiegelman. An approach to the experimental analysis of precellular evolution. Quart.Rev.Biophysics 4:213-253, 1971.

[52] J.F. Joyce. Forty years of *in vitro* evolution. Angew.Chem.Internat.Ed. 46(34):6420-6436, 2007.

[53] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. Die Naturwissenschaften 58(10):465-523, 1971.

[54] M. Eigen, P. Schuster. The hypercycle: A principle of natural self-organization. Part A: Emergence of the hypercycle. Die Naturwissenschaften 64(11):541-565, 1977.

[55] M. Eigen, J. McCaskill, P. Schuster. Molecular quasispecies. J.Phys.Chem. 92(24):6881-6891, 1989.

[56] M. Eigen, P. Schuster. The hypercycle: A principle of natural self-organization. Part B: The abstract hypercycle. Die Naturwissenschaften 65(1):7-41, 1978.

[57] M. Eigen, P. Schuster. The hypercycle: A principle of natural self-organization. Part C: The realistic hypercycle. Die Naturwissenschaften 65(7):341-369, 1978.

[58] E. Domingo, P. Schuster. Quasispecies: From theory to experimental systems. Current Topics in Microbiology and Immunology. Vol.392. Springer International. Cham, CH, 3016.

[59] S.Gago, S.F.Elena, R. Flores, R. Sanjuán. Extremely high mutation rate of a hammerhead viroid. Science 323(5919):1308, 2009.