THE NECKLACE PRINCIPLE OF CODING: PATTERNS OF FORCES IN RELATION TO INFORMATIONAL DETERMINISM

ILSE WALKER

Abstract. It is argued that specific combinations of zero force and of specific forces between the parts of complex structures and processes result in biological organization, characterized by repetition of unpredictable, complex structures and functions. Death, variation and the necessity of copy–reproduction, i.e. of selection, are consequences of these systemic patterns of force vectors.

I am thanking the editors, Dr. M. Arrigoni (mathematics), Dr. F. Pini (physics) and Dr. M. Villa (biochemistry) for stimulating discussions and helpful suggestions and corrections.

1. The problem

Organic, live and inorganic matter consist of the same kind of particles, the atoms, as we know them from the periodic table of the elements. The laws governing the physical chemistry of small molecules are the same in both types of systems, as for instance the laws of stoichiometry and of thermodynamics. What, then is the exact physical difference between live and inorganic systems? Is there a specific level of molecular complexity where the two kinds of systems can be distinguished in terms of the physical parameters of mass–energy in phase–space?

Standard biological texts mention reproduction, mutation and selection as the basic biological criteria. The origin of life is rooted in the faculty of self–replication, i.e. autocatalysis (Eigen, [6]; Miller and Orgel, [17]).
DNA, a prototype of organic molecule, consists of strings of nucleotides bound by regularly alternating sugar (pentose) and phosphate molecules. Considering the geometry only of subunit assembly, inorganic crystals, too, grow and multiply by regular juxtapposition of chemical subunits. As regards selection, every system that is relatively small and fast reacting in relation to human history, so that mankind can meditate on its existence, has a formation rate (B) which at least matches its rate of destruction (D), that is $B \geq D$. There is nothing specifically 'organic', as far as this basic law is concerned. Mutations are assumed to random perturbations on the molecular level, the DNA, which may lead to

Ilse Walker, Department of Ecology (CPEC), Instituto Nacional de Pesquisas na Amazônia (INPA), Caixa Postal 478, 69.083–000 Manaus (AM), Brasil (fax: (092) 642–1838).

relatively large–scale effects on the higher levels of complex organismic structure. Chaos theory, applied to the origin of storms and to currents shaping coasts lines (Gleick, [8]), became fashionable in biology (May and Oster, [16]; Papa and Da Silva, [18]). In short, none of the three criteria in itself (reproduction, selection, mutation) is specific to organic, live systems. If these traditional, biological criteria are indeed specifically relevant for biology, then there must exist basic, physical principles which relate them to each other in such a way that new, systemic laws arise, which allow for the clear distinction between inorganic and organic, live systems.

The problem, then, is to isolate phase–spatial patterns of molecules and forces which are specific to biological systems. It is understood that – should this analysis be successful – such patterns took shape during precellular evolution and pertain to todays pro– and eucaryotic organisms. However, for the sake of simplicity, the Watson–Crick structure of DNA is regarded as the basic, molecular model in what follows (Crick, [5]; Watson, [25]).

2. The geometry of replication: the origin of the 'individual'

Nucleotides polymerize longitudinally by regularly alternating PO_4^- and pentose radicals. Regular alternation of ligands is also in the rule in inorganic crystals, as in the salt NaCl, for instance. However, in contrast to DNA, the salt crystal 'replicates' by homologous bonding in all three spatial dimensions, thus giving rise to continuously growing, solid bodies. Nucleotide polymerization in DNA differs by two features. Firstly, it consists of asymmetrical subunits, that is, of mono–nucleotides only, thus polymerizing in the $5' \to 3'$ C–atom direction only. Secondly, the lateral polymerization units, the complementary N–base pairs AT and GC, are different from the longitudinal ones. These two conditions result in a *closed* double string of antiparallel single strings. For repeated, lateral polymerization, a complex, molecular apparatus must open the double–string, and lateral polymerization results in two *separate* double–strings. Hence, what really distinguishes DNA–replication from crystal growth is *periodic interruption* of lateral polymerization, and absence of polymerization in the third dimension.

The result are *populations* of similar, potentially freely mobile bodies, the 'individuals'. In the course of evolution, individualization was strengthened by stabilizing proteins and by association with the enzymes for autocatalysis.

As there are no specific forces between these 'primordial individuals', they would move according to the accidental forces in their aqueous environment. For ordered processes between these 'bodies', orientation and energy for reversible interactions must be provided. We thus find that organic macromolecules are, as a rule, metastable, allowing for diverse and specific reactions, while enzymes with their specific, three–dimensional structure guide the reactant molecules to specific places. Consequently, chemical processes that would take several hundred years under conditions of random movement, proceed in micro– and nano–seconds (Karlson, [11]). *Spatial information determines* the order between potentially

Figure 1. **A.** Chemical structure of the DNA– helix. P = phosphate group; Z = sugar (deoxy– ribose); A, T, C and G: symbols for the four bases adenine, thymine, guanine and cytosine respectively. 3', 5': indicate the 3.rd and the 5.th C– atom in the ribose and as such the polarity of the single strand. Copy– reproduction and reading of the genetic code proceeds from the 5' to the 3' – end (arrow), thereby A links to T and C to G only (dotted lines). Only one strand (+) serves as genetic code. **B.** Three– dimensional structure of the double– helix.

freely mobile bodies, between macromolecules and/or autocatalytic 'individuals', while separate inputs of energy (usually via ATP) allow for the chemical metabolism.

3. The necklace principle of coding

In single–stranded DNA a sequence of nucleotides is stabilized by the pentose–phosphate bonds. Laterally, each sugar binds to either of four different nitrogen bases, adenine (A), thymine (T), guanine (G) and cytosine (C). There are no chemical bonds between these bases, and whatever the forces between them, they are not sufficient to force them into a specific sequence. This disposition results in two parallel and *essentially different* string patterns: a force–free sequence of different subunits, the bases, which constitute the code, and the 'backbone' of regularly alternating phosphate and pentose radicals, which fix the particular string pattern of the bases (which, in addition, is stabilzed by the stacking forces along the DNA–double helix).

This is the necklace principle: The pattern of forces that link a code is not related to the pattern of the code, necessarily so, because for coding, codons must be freely combinable. The spatial pattern of the code, the information, and the forces that permit the constitution of this pattern, are two separate inputs.

Suppose we take the various precious stones of a torn necklace to a jeweler to repair it. He will want to know the former pattern. If we cannot remember, he cannot reproduce the original pattern. If he is competent, he will invent a new pattern, and the new necklace may be somewhat shorter or longer. From this little scene we derive the following principles:

1. Loss of a unique, force–free string–pattern is irrevocable, it is death. Owing to its complexity (number and types of subunits) it will not form *de novo* as a result of random movement within historically meaningful time.
2. Maintenance of the force–free string–pattern necessitates the presence of redundant sets for copy–reproduction in case of accidents.
3. As accidents are inevitable, the rate of copy–reproduction (B) must at least balance the rate of accidents (D), that is $B \geq D$.
4. The force–free condition of the string–pattern allows for variation, including variation of the chain length, provided the copy mechanism is not incapacitaded by the variation. Under the condition $B > D$, the new pattern can establish itself.

The necklace principle of coding, that is, the separation of phase–spatial information (sequence of codons) from the forces allowing for the synthesis and stability of the code, results automatically in death and in variation, and in the necessity of copy–reproduction for continued existence, in short: in *mutation* and *selection*. Salt crystals cannot 'die'. Under suitable conditions of temperature and concentration they form at any time in any place, because the specific vectors between their ions determine both, the order in space and the intensity of force, that is the shape and the solidity of the crystal.

The level of redundancy, that is, the necessary minimal number of spare copies, is set by the length of the coding string and by the range of the accident rates, which may include catastrophs. $B > D$ allows for the recovery of a critically reduced number of copies, and it allows for new varieties to attain the minimal necessary density for long–term maintenance ($B = D$). Chemical resources and space permitting, this does not necessarily entail extinction of other species. The principle of selection, $B > D$, means and foremost selection against accidents, i.e. repair.

Repair implies a hierarchical system: subunits are replaced within a 'super–unit', such as nucleotides and genes in chromosomes, chromosomes in cells (fertilization and meiosis), cell organelles in cells, cells per tissues, body segments in lower animals and plants, individuals in populations with social structure. The hierarchy as a whole as integrated dynamical system is constituted and maintained by the high–precision mechanism of gene regulation, which includes negative feedback regulation of structure and function.

Two mathematical laws impose repair for the maintenance of force–free constellations and for their codes in particular: Firstly, the relation between copy errors and selection, and secondly, the relation between redundancy and phenotype.

3.1. **Copy errors and selection.**

Apart from any other types of mutations, copy errors in DNA– and RNA–replication are inevitable. The longer the code, and the higher the rate of replication, the higher the error rate per time. Hence, the process of selection itself, i.e. faster reproduction, increases the error rate. Based on realistic, chemical data of RNA–stability and copy error rates in a small virus, Eigen and Schuster ([7]) calculated the relation between maximum genome length (number of nucleotides), error rate and selective superiority of the optimized wild type. For the maintenance of the ca 4500 nucleotides of the virus genome, the wildtype must grow at least 20 times faster than the average of the rest of the population with one to several errors per genome, and this allows for a maximum of only 6% optimized wildtype in the population (= necessary redundancy for the maintenance of the genome). The authors concluded that only powerful mechanisms of repair could explain the evolution of the large genomes of the eucaryotes with their several billions of nucleotides. They proposed a model of post–replication repair between two homologous chromosomes, as occurs during sexual reproduction.

In their model, D (in the condition $B > D$) includes loss to sources allowed for the emergence of ever new 'wild–types', i.e. species, provided they obeyed *their* condition $B \geq D$. Competition may be a relevant type of accident, but competition is not a necessary condition of evolution (Walker, [24]). Regarding the evolution of genomes under these considerations, we find that the genomes of today's species range from shortest in procaryotes to longest in the higher animals and plants, and that the continuous conversion of inorganic matter to organic macromolecules allowed for the coevolution of innumerable species. Higher organization is not the result of competitive selection against lower species.

3.2. **Redundancy and phenotype.**

The apparently contradictory term 'necessary redundancy' refers to the time structure of the system: the redundant set is presently redundant with regard to its actual function, but is necessary in future for repair. For this reason, selection for redundancy via the actual functional state, that is, via the phenotype, is not possible. There is no difference of fitness between the redundant and non–redundant types (between homozygous and heterozygous dominant wildtype in mendelian genetics, for instance). Mathematical analysis shows that it is impossible to stabilize the number of individuals and the redundancy in their substructure by selection at the level of individuals. Stabilization of hierarchical systems with redundant sub–structure needs repair–restitution of the redundant sub–units (Williams and Walker, [27]; Steiner and Walker, [21]). This means that for long–term persistence of species ($B \approxeq D$), selection must operate on the level of the repair system as such. Beside the complex repair apparatus of the genetic code itself, this includes asexual reproduction

as well as regeneration of tissues, organs and body segments after mutilation in animals and plants. It should thus not come as a surprise, that molecular biology reveals ever new and more sophisticated repair mechanisms. Sentences like: "Cell cycle checkpoints ensure that chromosomal DNA is replicated and repaired before nuclear division" (Boddy et al., [4]) show that by now DNA repair mechanisms are a matter of course. On the level of tissue maintenance there is the process of genetically regulated cell death (apoptosis) which eliminates damaged and excessive cells (Hengartner, [9]). What remains as an issue of controversy is the evolution of sexual reproduction as a mechanism of DNA–repair (Walker, [23]; Bernstein and Bernstein, [2]; Avise, [1]). Within–population organization is possible to the extent individuals emit, perceive and understand the *same language*, that is, the same physico–chemical signals. Bisexual reproduction means that only individuals with reasonably similar genomes can leave offspring, and continuous crossbreeding with recombination and post–replication repair during meiosis (Howell and Stern, [10]) maintains vital old, and distributes favourable new, information within the population/species as a whole. Bisexual reproduction thus *homogenizes the information within species*. Not surprisingly, *obligatory bisexual reproduction* (with obligatory somatic mortality!) is the rule in higher animals with complex neural systems (Walker, [23]; Williams and Walker, [27]).

4. The code–determined system

The 5' → 3' sequence of the four bases (A, T, G, C) in the DNA–strand determines the sequence of amino acids in the proteins. Specific base triplets symbolize specific amino acids in the polypeptide string, the primary protein structure. There are no specific forces between the 20 odd different amino acids, so that they are freely combinable according to the base sequence in the DNA strand; string stability is provided by identical peptide bonds between any two neighbour amino acids. Furthermore, there are no specific forces between the triplet codon and the corresponding amino acid, the 'analogon' (Walker, [22]). Amino–acid–specific enzymes (t–RNA–ligases) with their specific 3–dimensional structure ascertain the correct association between codon and analogon in a highly complex translation process, which includes an error correction mechanism. The energy necessary for the various chemical reactions is provided by separate ATP–inputs.

In short, the necklace principle, that is, the separation between force and phase spatial information (structure and assembly sequence) extends from the code to the code–determined system. This principle relies on informed location (locomotion) of potentially independent bodies: transport of metabolites through membrane pores, through ducts and circulatory systems, cell movement during embryogenesis, and autolocomotion of lower algae and of animals, to mention some examples.

A word of caution, though, is necessary in order to avoid misunderstandings. The 'necklace principle' of coding and its extension to the code–determined 'target systems', that

is, *separate inputs of causal information and of permissive energy supplies* refers to specific processes of subunits assembly, and it does not contradict the principles of classical physics and chemistry. The permissive energy supplies follow these laws, and absence of bonds between 'neighbour' molecules is the usual situation in liquids or gases. The necklace principle of coding combines these two features by a simple geometrical trick: a string of similar chemical bonds (pentose–phosphate–backbone) fixes a random sequence of molecules from "the liquid" by lateral bonding, the coding nitrogen bases A, T, G, C. DNA–replication, in turn, depends on specific chemical forces between the complementary base pairs AT and GC, and the spatial information provided by enzymes depends on their tertiary and quaternary structure, which is determined by chemical bonds between specific amino acids in the peptide string (sulfur bridges for instance). Evidence is accumulating that direct forces constrain the freedom of codon sequence on higher levels of organization: comparative chromosome analysis shows that gene order along the chromosomes is not random, this led to the 'chromosome field theory' (Lima–de–Faria et al., [15]; Lima–de–Faria, [14]). Furthermore, gene regulation and transcription needs a precise nuclear architecture (Scherrer, [20]; Wei et al., [26]). It may be argued that during pre–cellular evolution any one singular nucleotide string may have arisen by chance, as discussed by Küppers ([13]) for instance. However, weak forces are likely to bias random probabilities. Whatever such bias, what needs explanation is the multiplicity of these improbable structures, and this is only possible by means of positional information in combination with copy–reproduction.

4.1. **Neural codes and artefacts.**

One of the products of the primary genetic code is the neural code, neural coding implying that the perception of more or less complex, environmental patterns takes on a symbolic significance for repeated actions, such as remembering a particular path to a particular food source, collection of nest materials, signal transduction between individuals and language in the most ample sense.

It is still unknown how the genetic code tranlates into neural codes. Changes in the environment may alter gene expression in neurons of experimental animals, however, it is not known, which particular genes are involved in the consolidation of specific memories (Birnbauer and Schmidt, [3]). Certain is that the resulting animal and human artefacts obey the principles of separate inputs of force and information: spider webs, birds' nests and car engines are not the product of the forces between their structural subunits. After destruction, these artefacts can only be reconstructed via specific neural information. Noteworthy are the systematized, human–cultural codes which were invented during the last five millenia: written language, mathematics and music, and their translation into the binary code of modern informatics. In all cases, defined sets of symbols without specific forces between them combine to linear sequences with a given reading direction which contains the information, while any kind of of stable matrix allows for free composition and maintenance of the codes: stone for hieroglyphs, leather, papyrus, paper, tapes, and disks etc. Long–term maintenance needs redundancy and copy–reproduction, accidents and errors result in irrevocable loss, and new inventions accelerate evolution, the History

of mankind. Space and resources permitting, the condition B $\geq$ D leads to growth and expansion, while the competitive component in D affects the rate of diversification (which may be negative, as during imperialistic periods, for instance 'globalization'). Thus, while the code–dependent sciences of mankind largerly determine its history, the complementary – but by no means contradictory – argument is equally valid, namely that History is Biophysics.

5. Conclusions: non–additivity of force in informational determinism

The problem phrased in the Introduction was "to isolate phase–spatial patterns of molecules and forces which are specific to biological systems, from which systemic laws arise which allow for a clear distinction between inorganic and organic systems".

The necklace principle of coding consists in the force free association of codons and in the stabilization of these association patterns by non–specific linkage to an auxiliary substrate. That is, *causal phase–spatial information and permissive energy supply are two separate inputs*. Codon–specific forces, on the other hand, allow for code replication, i.e. for repetition in space and/or time of information and of information–determined structures and processes. These principles extend from the genetic code to the information systems on higher levels of organization.

The *criterion* for the distinction of biological systems from inorganic systems is therefore: multiple repetition in space and/or in time of complex structures and processes the genesis and/or maintenance of which cannot be determined by means of addition of force vectors between the constituent components. This *principle of organization* includes artefacts of biological systems: multiplicity of right angles in amorphous substrates, for example, allows for the conclusion that the structures are information–determined, as we deduce from the ruins of ancient cities. This brings it to the mind that human history is biophysics.

In organic systems, high levels of complexity, unpredictable and irreversible variation and destruction are the result of vector components of zero–force between specific subunits, while non–zero force vectors between relevant subunits allow for reproduction (i.e. for order in the sense of repetition) of these complex systems, provided informational redundancy of the zero–force coding pattern (B $\geq$ D) can be maintained. Thus, the irreversible emergence and extinction of *complex order* in ontogeny and phylogeny is the physical consequence of the species–specific patterns of zero–force–vectors in the infrastructure of biological systems.

The *systemic alternation* between zero–force and information–determined force vectors in biological structure and function results in time *a*symmetry of the cognitive process of the observer: retrospectively, causal, newtonian physics is invariably confirmed, yet, vector analysis does not allow for the prediction of future, highly repetitive patterns. This situation seems to meet Pauli's ([19]) characterization of possible complementarity as regards the physics of living systems, mamely "...the possibility of specific biological laws which

are not contained in our present–day physics, yet present–day physics being invariably valid where biological systems are examined by its methods".

References

[1] Avise, J.C., *The evolutionary biology of aging, sexual reproduction and DNA repair*. Evolution 47, 1293–1301, 1993.

[2] Bernstein, C. and Bernstein, H, *Aging, sex and DNA–repair*, Academic Press, New York, 1991.

[3] Birnbauer, N. and Schmidt, R.F., *Allgemeine Physiologie der Grosshirnrinde*. In: Schmidt, R.F. and Thews, G. (eds), *Physiologie des Menschen*. Springer, Berlin, Heidelberg, New York,1995.

[4] Boddy, M.N., Furnary, B., Mondesert, O. and Russel, P., *Replication checkpoint enforced by Kinases Cdsl and Chkl*. Science 280: 909–912, 1998.

[5] Crick, F.H.C., *Codon–anticodon pairing: the wobble hypothesis*. J. Mol. Biol. 19: 548, 1966.

[6] Eigen, M., *Self–organization of matter and the evolution of macromolecules*. Naturwissenschaften 58: 465–523, 1971.

[7] Eigen, M. and Schuster, P., *The hypercycle*. Springer, Berlin, Heidelberg, New York, 1979.

[8] Gleick, J., *Chaos, making a new science*. Viking, New York, 1990.

[9] Hengartner, M.O., *Apoptosis and the shape of death*. Genetics 21: 245–248, 1997.

[10] Howell, S.H. and Stern, H., *The appearence of DNA–breakage and repair activities in the synchronous meiotic cycle of Lilium*. J. Mol. Biol. 55: 357–378, 1971.

[11] Karlson, P., *Kurzes Lehrbuch der Biochemie*. Georg Thieme Verlag, Stuttgart, New York, 1980.

[12] Koolman, J. and Röhm, K.H., *Taschenatlas der Biochemie*. Georg Thieme Verlag, Stuttgart, New York, 1998.

[13] Küppers, B.O., *Der Ursprung biologischer Information*. Piper, München, Zürich, 1986.

[14] Lima–de–Faria, A., *Chromosome gradient and chromosome field in Agapanthus*. Chromosoma 6: 330–370, 1954.

[15] Lima–de–Faria, A., Arnason, U., Widegren, B., Isaksson, M., Essenmüller, J. and Jaworska, H., *DNA cloning and hybridization in deer species, supporting the chromosome field theory*. BioSystems 19: 185–212, 1986.

[16] May, R. and Oster, G.F., *Bifurcations and dynamic complexity in simple, ecological models*. The Amer. Naturalist 110: 573–599, 1976.

[17] Miller, S.L. and Orgel, L.E., *The origin of life on the earth*. Prentice Mall Inc. Englewood Cliffs, New Jersey, 1974.

[18] Papa, A.R., *Avalanches no cérebro*. Ciência Hoje 135: 18–21, Rio de Janeiro, 1998.

[19] Pauli, W., Paragraph from a letter to Prof. E. Hadorn, University of Zurich.

[20] Scherrer, K., *A unified matrix hypothesis of DNA–directed morphogenesis, protodynamism and growth control*. Bioscience Reports 9/2: 157–188, 1989.

[21] Steiner, A. and Walker, I., *The pattern of population growth as a function of redundancy and repair*. Acta Biotheoretica 38: 83–90, 1990.

[22] Walker, I., *Biological Memory*. Acta Biotheoretica 21: 203–235, 1972.

[23] Walker, I., *The evolution of sexual reproduction as a repair mechanism. Part I. A model for self repair and its biological implications*. Acta Biotheoretica 27: 133–158, 1978.

[24] Walker, I., *Competition and information*. Acta Biotheoretica 41: 249–266, 1993.

[25] Watson, J.D., *Molecular biology of the gene*. W.A. Benjamin Inc., Menlo Park, California, 1970.

[26] Wei, X., Samarabandu, J., Devolhar, R.S., Siegel, A.J., Acharia, R. and Brezney, R., *Segregation of transcription and replication sites into higher order domains*. Science 281: 1502–1505, 1998.

[27] WILLIAMS, R.M. AND WALKER, I., *The evolution of sexual reproduction as a repair mechanism. Part II. Mathematical treatment of the wheel model and its significance for real systems.* Acta Biotheoretica 27: 159–184, 1978.