PIOTR LENARTOWICZ SJ

ARE WE FULLY SHAPED AND DETERMINED BY OUR GENES?

Opublikowano w: ***Genethik, Internationale Mediziner Arbeitsgemeinschaft,*** 41 Internationales Karwochenseminar, 9-14 April 1997, St. Virgil, Salzburg 1997, pp. 67-80.

Our adult body is composed of miniature cells, millions and billions of them. Each cell contains the identical copy of a relatively very long nucleopolymer molecule, divided into 46 uneven fragments, which are tightly packed into structures called chromosomes. In other words, our body contains billions of the absolutely identical polymer molecules of the DNA. According to a current, widespread belief, any single set of 46 fragments of the DNA molecule carries all forms of encoded information capable to build the adult human body and to determine the behavior of a concrete specimen of *Homo sapiens*. The subunits of this molecular information are called genes.

Figure 1 depicts the complex, highly ordered way a fragment of the DNA molecule is coiled within the body of a single chromosome during the metaphase period of cell division.

The general perspective – spirits and genes

At the beginning of my talk I will try to show and to discuss the ancient prototype of the modern idea of the *genetic information.* We are kind of animals, I have no doubt about it. Yet we seem to possess „something", a property, a capacity which traditionally is called „soul". Before the discovery of the molecular biosynthetic codes and before the DNA became a symbol of the inner, hereditary determination, it was the *Aristotelian soul* which was supposed to determine the shape and the behavior of our material adult bodies. The reader of this text has to be warned that the word *„soul"* is rather ambiguous. In the European and Christian culture we can distinguish two, quite different ideas of the human soul.

On one extreme we have the platonic soul which reminds me a poor Arabian spirit, incarcerated in a bottle and buried in the sands of the Arabian Desert. This powerful spirit cannot bear this bottle. His life is crumpled there. One has to open the bottle or to break it. Only then this spirit can manifest its full power, full activity. The death of the bottle (read „body") means liberation and a triumph of the *platonic soul.*

On the opposite pole we have the *Aristotelian* idea of soul. This invisible spirit lives in a bottle and cannot survive without this bottle. He produces it from the relatively raw matter, he equips it with different tools (members) and machines (organs) which help him to master the environment. The embryological development constitutes this period

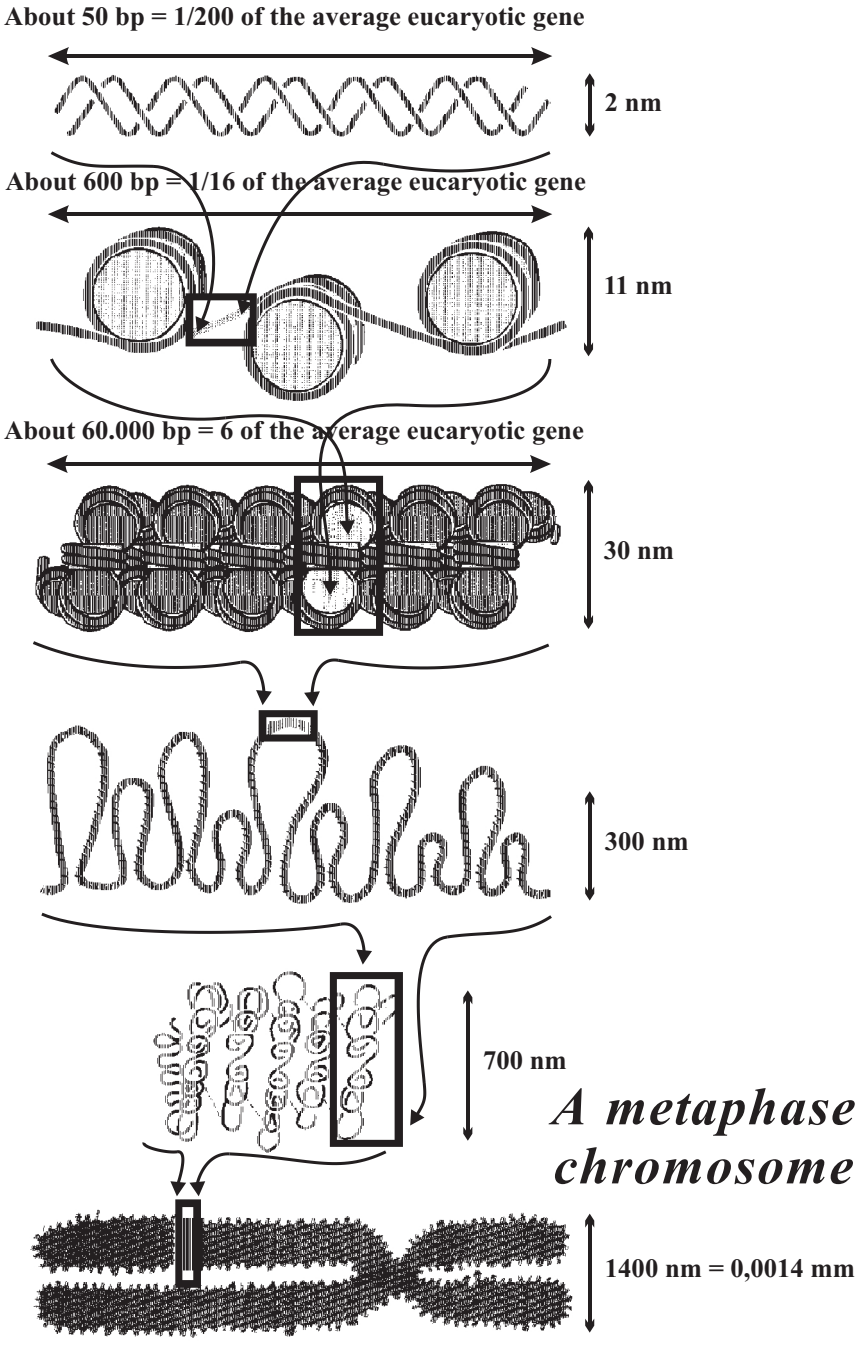
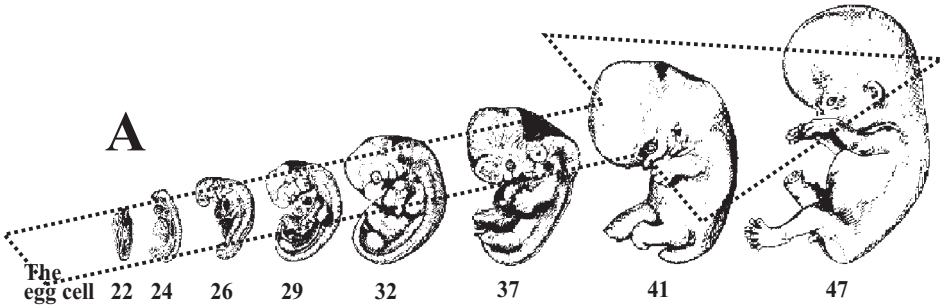


Fig. 1. **C**omplex, highly ordered way a fragment of the DNA molecule is coiled within the body of a single chromosome during the metaphase period of cell division (after Freifelder, 1987/183 and Alberts *et al*., 1994/354).

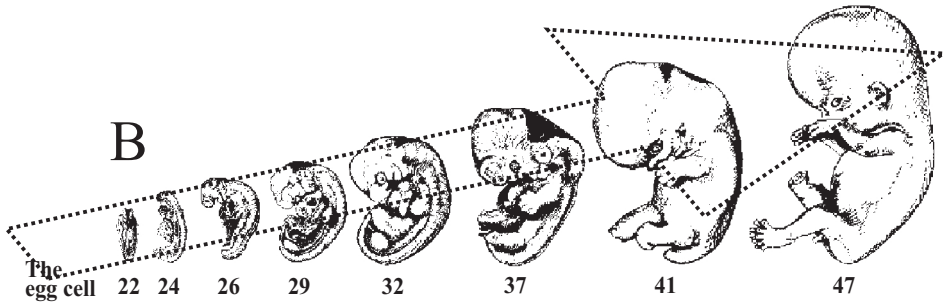
of life in which the power of the Aristotelian soul is seen at its best. His spiritual bottle - called „the body" – has a nice head, skillful hands, a good little computer between the ears. On the whole the Aristotelian spirit feels comfortably in his bottle, tries to keep it in a good fit and to repair it in case of a mutilation.

The „selfish gene"

Apart from those two completely different theories of the *spiritual* soul we have the theory of a *material* soul. Dawkins named it „the selfish gene". This selfish gene is believed to shape our body, like the Aristotelian soul, and to determine our behavior. Figure 2 portrays the conceptual identity of the role ascribed to the Aristotelian soul on one hand and the DNA encrypted molecular messages[1]. However, the idea of a molecular genetic program differs from the Aristotelian spirit in several important points.



Assumed activity of the Aristotelian soul



Assumed activity of the molecular DNA messages

Fig. 2. Conceptual identity of the role ascribed to the Aristotelian soul and the DNA encrypted molecular messages.

[1] Max Delbrück has noticed this essential similarity between the Aristotelian idea of „soul" and the modern ideas on the genetic program (see *„Aristotle-totle-totle"* in: *On microbes and life*, ed. by J. Monod and A. Borek, New York, 1971). Ernst Mayr writes: „... Aristotle more than 2.000 years ago already understood remarkably well that a program of instruction is needed for the development of an egg ..." *(„Toward a new philosophy of biology"*, Harvard UP, 1988, p. 261, see also p. 249).

Perfect and improvable

First, the Aristotelian soul was conceived as something absolutely integrated and indivisible. Its masterpiece, the „body" can change from an imperfect into a perfect form. New body parts may appear or disappear, the body can be mutilated. But the soul itself is not a spatial being, therefore it does not occupy a dimension and cannot be divided, dismembered into spatial parts. According to the Aristotelian doctrine, plant and animal souls are destructible *per accidens*, that means, the extensive damage inflicted on the „body" they build and control, might – somehow – irreversibly stop their activity. The human soul, according to that doctrine, was believed to be absolutely indestructible, i. e. the death of its body does not stop its activity. The „selfish DNA gene", on the contrary, is not indivisible, it occupies space, a part of it can be added, and a part can be removed. It is subject to changes, improvements and mutilations (mutations). Mendelian segregation can operate on it at random.

Active and passive

Second, the Aristotelian soul was thought to be immanently active. Actually is was presumed to exist – in a close union with „matter" – as a *pure, immanent activity* (pure act). Its activity differs from the material, causal activity of the tools it shapes (like eyes or hands) and the machines it constructs (like enzymes and neural wiring), yet its activity is able to exploit the raw resources of the matter and energy to produce those tools and those machines we call animal or human body. So the activity of the Aristotelian soul is essentially different from the dynamism of raw matter and the raw energy on one hand, and from the dynamism of a machine on the other.

The Aristotelian soul acts in a selective way, but in contradistinction to a machine, it is not determined in its activity by a material structure – it is capable to shape structures, to impose a form on the passive matter. The selfish gene, to the contrary, has no activity on its own. If an activity might be attributed to it, it has *a priori* to be of the same kind as the dynamism observed in the chemical and physical surroundings.

Complexity and functionality

Third, the activity of the Aristotelian soul consists in *integration,* that means in a permanent and undeviating tendency to shape the matter not in a random form but into the functional, efficient „tools" or „machines", i. e. *the functional structures*. *A propos* of this third aspect: many of us, I believe, would say – „here I see no difference". The idea of a selfish gene and the idea of the Aristotelian soul are identical – just in this particular aspect of integration. We may leave this problem aside.

There are two reason why I am leaving out this question: (1) The „selfish gene" (or „genome") is usually conceived as something like a sophisticated computer superprogram. According to my knowledge, no computer program was ever created by the raw matter. The abiotic origin of the genetic information – be it complete or incomplete – is an old, respectful scientific dream, quite similar to the desire to build a *perpetuum mobile*. (2) In the next part of my talk I will try to show, how deficient the DNA messages are during biosynthesis. So at the moment there is no need to enter in a row on the supposedly miraculous capacities of the DNA to shape, on its own, the integrated structures of living body.

Raw matter – precursors – final products

The shaping of our body takes place during the embryological stage of our life. It is a gradual process and consists in building a hierarchy of structures, starting from the molecular level. Biomolecular level, organellar level, cellular level, organ level are just steps on the path to reach the complete adult body's anatomical level.

The biomolecular level has also a hierarchical structure. For instance the biosynthesis of an enzyme (a molecular machine) involves the process of putting together different single aminoacid units into a nonrandom polymerized file, then folding it in the 3D space in a nonrandom structure. Usually it is not the end of the process. The cell has also to mount into it a cofactor (coenzyme), or to position it in a non-radom way among other elements of cell's machinery.

Our biochemical knowledge is still rather fragmentary, but some details of biosynthesis are fairly well understood. What we already know about the processes of biosynthesis seems incompatible with the current belief that the information enciphered on the DNA molecule is sufficient to control the processes of biosynthesis alone, not to mention cyto-, morpho- and organogenesis.

A „soul" or a „crib"?

Now I am turning to the main part of my talk. We know fairly well the structure of some DNA genes and we know the general outline of the biosynthesis in the case of macromolecules. This fragmentary knowledge proves quite convincingly, that the messages enciphered along the DNA molecule are not a „soul" (the ultimate principle driving and controlling the integrated development) but just a molecular crib. The regular crib, Ladies and Gentlemen, is a small sheet of paper, containing the most difficult parts of knowledge necessary during an examination. The crib is *fragmentary*. There is no use to copy the whole textbook. The crib has to be small. The crib is *enciphered* and the information is written in an *abbreviated* form. Only its author can fully understand those abbreviations. Finally, the information written on the crib is usually *chaotic*. For an author it doesn't matter where a particular piece of information is written. He may find it easily without putting it in a particular order. Besides, because of the fragmentary nature of those messages no real, objective order is possible. The informational gaps prevent any logical order.

Now I will show you, Ladies and Gentlemen, an example of the process of shaping the molecular structures within the cell. You will see, how the living cells utilizes its crib (I mean the DNA genes) and you will see how fragmentary is the informational content of those messages. But first we have to make another distinction, namely between the molecular signals and the molecular instructions.

Signals and instructions

In our conscious life we may have to do either with an instruction (which informs us *how* to do something) or with signals which remind us about something, but do not instruct us exactly what to do. An alarm clock gives me a signal – its sound provides me with no idea what I have to do.

In the DNA molecule such *quasi* alarm clock signals were discovered. They determine a given succession of different changes, but do not instruct a cell what has to

be done, or how. Those signals became famous with the discovery of the so called *homeoboxes*. The homeoboxes are small genes, determining an aminoacid sequence of a small polypeptide, which has no causal power to act, but when it appears in the cytoplasm, it evokes a coordinated reaction, just like the traffic lights coordinate the movement of cars on the street. A traffic light has no causal power to force a change in the direction, to speed up, or to slow down the traffic. It has to be interpreted in the right way, otherwise it is meaningless. Actually, the homeobox products do not differ essentially from the well known hormones.

The DNA homeoboxes have surprisingly similar chemical structure in different animal groups, for instance in mammals and in insects. But the reaction of the organism to such a gene product is different in a fly and quite different in a mouse.

tRNA biosynthesis

But in addition to the homeoboxes the DNA molecule carries the *instructions* which are absolutely necessary to produce enzymes, different nucleotide complexes, and smaller protein molecules used in the processes of transcription and translation.

One of such important tools are tRNA molecules. Each one living cell constantly utilizes some 40 different forms of the tRNA molecule.

We will concentrate now on the production of the transporting RNA for the tyrosine aminoacid in yeasts.

The cellular DNA molecule carries a gene which determines a rough, crude, incomplete shape of the $tRNA^{Tyr}$. To produce a functional $tRNA^{Tyr}$ molecule the following steps are necessary:

(1) **The $tRNA^{Tyr}$ gene has to be found in one of the chromosomes**. Here that gene is represented with the aid of four symbols *A*, *C*, *G*, *T*, substituted for the rather complex chemical formulas of four different deoxyribonucleotides:

the yeast DNA gene for $tRNA^{tyr}$

GTTATCAGTTAATTGACTCTCGGTAGCCAAGTTGGTTTAAGGCGCAAGACTGTAATT
TACCACTACGAAATCTTTGAGATCGGGCGTTCGACTCGCCCCCGGGAGATT

(2) **The DNA gene has to be copied (transcribed)**. The gene remains intact in the structure of the given chromosome. The subunits of the copy are made from a different kind of sugar (not deoxyribose, but ribose), and one of the original organic bases (T) is substituted by another one (U):

the early RNA transcript of the gene

GUUAUCAGUUAAUUGACUCUCGGUAGCCAAGUUGGUUUAAGGCGCAAGACUGUA
AUUUACCACUACGAAAUCUUUGAGAUCGGGCGUUCGACUCGCCCCCGGGAGAUU

Both the $tRNA^{Tyr}$ gene and its faithful RNA copy consist of four different units (specific, complex chemical molecules; mol. m. about 300) linked in a sequence 104 subunits long. It may seem to us quite random, but in fact it is absolutely non-random, i.e. this specific sequence is closely correlated with the final, functional properties of the $tRNA^{Tyr}$ molecule. Virtually any change in this sequence makes it unfit to fill properly its job in the cell.

This nonrandomness may be rephrased as *selectivity.* We may measure it in a very crude, approximate way, by calculating how many different sequences one could make from the same 104 subunits of four different kinds. This calculation based on a simple

combinatorial equation tells us, that there are approximately 10^{63} different sequences composed of the 104 subunits of four different kinds. So, owing to the DNA gene which stores the proper sequence, the cell can build a right precursor – eliminating the wasted material, wasted energy and wasted time. In other words the cost of the successful biosynthesis is cut by *sixty two orders of magnitude!*

I said *precursor*. The RNA copy dictated by the structure of the $tRNA^{Tyr}$ *gene* undergoes many further complex modifications before it changes into the functional primary structure of the $tRNA^{Tyr}$ molecule.

(3) **Deletions**. Parts of the precursor copy have to be removed. The removed elements are written in lowercase:

the yeast $tRNA^{Tyr}$ precursor I

guu auc agu uaa uug aCU CUC GGU AGC CAA GUU GGU UUA AGG CGC AAG ACU GUA Auu
uac cac uac gaa AUC UUU GAG AUC GGG CGU UCG ACU CGC CCC CGG GAG Auu

(4) **Additions**. A couple of nucleotides have to be added. The added elements are shown in bold characters:

the yeast $tRNA^{Tyr}$ precursor II

CUC UCG GUA GCC AAG UUG GUU UAA GGC GCA AGA CUG UAA AUC UUU GAG AUC
GGG CGU UCG ACU CGC CCC CGG GAG A**CAA**

(5) **Transformations.** Several single units have to undergo a transformation. The transformed elements are shown in bold characters:

the final, functional primary structure of the yeast $tRNA^{Tyr}$

CUC UCG GUA **Gm**C CAA G**DD Gm**G **DDD** AAG GC**G m**CAA GAC UG **Ps**A **Ai**A **Ps**C UUU
GAG A**DC m**GG GCG **TPs** CG**A m**CU CGC CCC CGG GAG ACCA

Again let us have a look at the early RNA transcript of the tRNA gene (1) dictated by the sequence of the DNA gene

the early RNA transcript of the gene – the supposedly unique source of information

GUUAUCAGUUAAUUGACUCUCGGUAGCCAAGUUGGUUUAAGGCGCAAGACUGUA
AUUUACCACUACGAAAUCUUUGAGAUCGGGCGUUCGACUCGCCCCGGGGAGAUU

The final, functional primary structure of the yeast $tRNA^{Tyr}$ differs considerably from the original DNA gene. It is shorter than the gene – just 78 subunits long, but it is composed not from *four* but from *eleven* different subunits. We may apply to it the same approximate measure of selectivity we used above. To produce such a sequence from 78 random subunits a selectivity $1:1{,}6x10^{81}$ is necessary.

So the process of modifying the raw precursor involves *a considerable increase in the selectivity* of the product. This increase is 10^{18} fold. In other words the selectivity guaranteed by the DNA gene is insignificantly small in comparison with the selectivity really needed and actually somehow provided.

$$\text{Selectivity of the final product} = 1:1{,}69 \times 10^{81}$$

$$\text{Selectivity imposed by the DNA gene} = 1:4{,}11 \times 10^{62}$$

$$\textbf{The informational deficit} = 1 : (1{,}69 \times 10^{81} - 4{,}11 \times 10^{62}) = \mathbf{1:1{,}68 \times 10^{81}}$$

Where this enormous amount of the „selective power" does reside? At the moment nobody can answer this question. One seems evident. If we are searching for structures capable to carry and to impose such a selectivity on the complex chemical structures, we

really don't know where to look for them. The present, almost completed inventory of the cell certainly does not contain any complex set of structures (of the unknown role) comparable to the system of transcription or translation. In the case of the posttranscriptional modifications we just analyzed, we would have to postulate a system unimaginably more precise, and consequently structurally unimaginably more complex. I don't see any chance of finding such a system. So what? The cell, no one could doubt it, does utilize a mysterious source of information.

Do we have – in our search – to enter the subatomic sphere of biostructures? I really don't know. We see the problem but we certainly don't know the answer. However, something seems clear enough. It is premature to claim that the DNA molecule was shown to carry all the information needed to shape our cells, our organs, our minds, our intellect ... and so on. Actually, we don't even know where the instructions needed to shape the common biomolecules are hiding.

Artificial synthesis of protein molecules (Merrifield)

My second example concerns the process of protein synthesis. As I have already said, this synthesis involves the production of the right, nonrandom sequence of the different aminoacid units and then the proper folding of the resulting polypeptide chain in 3D space.

Some people go on repeating that given the right sequence of the aminoacids all the remaining problems of structure are solved. This, in my opinion, is another example of a pseudoscientific, unwarranted generalization.

In sixties, the American chemist Robert B. Merrifield (Nobel Prize 1989) invented a machine-like, chemical laboratory system capable of producing any arbitrarily determined sequence of the aminoacids. So, since almost 30 years we are capable to produce any sequence of aminoacids we like. Yet, to fold it properly in space is still beyond our capacity. To be precise, during a random folding of hundreds and thousands of the identical molecules a number of them folds in the proper way. But the statistical yield is a far cry from the virtually 100% efficiency of the cell[2].

---

[2] Cfr. Tsou Chen-Lu (1988) *Folding of the Nascent Peptide Chin into a Biologically Active Protein*. Biochemistry 27, 1809-1812; Wright P. E., Dyson H. Jane, Lerner R. A. (1988) *Conformation of Peptide Fragments of Proteins in Aqueous Solution: Implications for Initiation of Protein Folding*. Biochemistry 27, 7167-7175. „The greatest challenge remains the folding of the nascent polypeptide chain in vivo."; Rothman J. E. (1989) *Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells*. Cell 59, 591- 601; Pain R. H. (1990) *Shuffling on this mortal coil*. Nature 344, 198-199; Sander Ch. (1990) *Inverting the protein-folding problem*. Biochem. Soc. Symp. 57, 25-33. „Today, the protein-folding problem as a structure prediction problem remains fundamentally unsolved."

„... pure proteins that have been unfolded by chemical denaturing agents will often refold spontaneously to their correct functional conformations on dilution or removal of the denaturing agent. Because no molecules other than the solvent were present, these experiments led to the important conclusion that all the steric information for polypeptide chains to refold correctly is contained within their primary structures. However, success in such refolding experiments is favored by protein concentrations much lower than those found inside cells. At concentrations similar to those of the cell, proteins often aggregate

So again, if it is true that a DNA gene provides the cell with the exact instructions how to make a proper aminoacid sequence, it is not true that this gene is sufficient to control the shaping of the final, functional product.

mRNA editing

About ten years ago the mRNA editing was discovered. That means that the DNA genes, in some cases, do not determine the proper sequence of aminoacids. The information provided by DNA messages is – in these cases – simply wrong, from the functional, biological point of view, and it has to be *edited* during a separate stage (posttranscriptional modifications). As a result of this editing the original, primary molecular meaning of a given gene can be radically changed. Suppose an author brings a text to be printed, but the text has no spaces and some letters are missing. The editorial board has to decide where to put a space and where to insert an additional letter to make the text not only legible but to save its presumably integrated sense[3].One has to add that in some mitochondria thousands of different (heterogenous), small RNA molecules (gRNA for „guiding RNA") were found, and that those molecules contain the fragmentary information needed in the process of mRNA editing. However, the utilization of those dispersed pieces of information entails an orchestrated activity of many different molecular agents in many different compartments of the cell[4] . The conductor remains invisible. So the mystery of the mRNA editing – and of its origin – remains essentially unsolved.

Other empirical data on the DNA informational deficit

If one considers DNA as a top agent, the main ruling agency of the body mechanisms,

---

because partially folded chains can interact with one another incorrectly through transiently exposed hydrophobic areas. Misfolded chains can also arise when partially folded intermediate states become trapped and are unable to proceed toward functional conformations at suitable rates." (Ellis R. J. (1996) *The „Bio" in biochemistry: protein folding inside and outside the cell*. Science 272, 1448-1449).

[3] Cfr. Simpson L., Shaw Janet (1989) *RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa*. Cell 57, 355-366; Volloch V., Schweitzer B., Rits Sophia (1990) *Uncoupling of the synthesis of edited and unedited COIII RNA in Trypanosoma brucei*. Nature 343, 482-484; Weissmann Ch., Cattaneo R., Billeter M A. (1990) *Sometimes an editor makes sense*. Nature 343, 697-696.; Simpson L. (1990) *RNA editing - a novel genetic phenomenon?* Science 250, 512-513.; Echols H., Goodman Myron F. (1991) *Fidelity mechanisms in DNA replication*. Ann. Rev. Biochem. 60, 477-511; Cattaneo R. (1992) *RNA editing: in chloroplast and brain*. TIBS 17, 4-5; Benne R., van der Spek H. (1992) *L'editing des messages génétiques*. La Recherche 23, 846-854; Stuart K., Allen T., E., Heidmann S., Seiwert S. D. (1997) *RNA editing in kinetoplastid protozoa*. Microbiol. and Molec. Biol. Revs, March 1997, p. 105-120. „Recognition of this RNA editing process reveals the existence of a previously unrecognized level for the control of gene expression". K. Carr enumerates several specific mechanisms of control which participate in determining the fate of the raw information provided by the DNA genes (1994, *Life after transcription*, Nature 369, 440-441).

[4] See Stuart K. *et al*. (1997)

then any mutilation of the DNA molecule should be fatal and unrecoverable. In reality DNA can be repaired, and it is constantly repaired. Somehow the organism knows how to detect a change in the genetic message, an informational mutilation (a mutation), and utilizes many different, complex procedures to repair different forms of the biochemical wound. It is also rather interesting that those parts of the DNA which carry no specific genetic message are seldom, if ever, repaired. So the system of rescue seems to „know" the *meaning* of a structure in its broader context, not just its chemical structure. From the purely chemical point of view the idea of a „mutilation" is meaningless.

Conclusions

Many efforts have been undertaken to explain away the facts concerning the crib-like properties of the DNA molecule. Usually they consist in an attempt to find the missing part of the information in the other parts of the DNA molecule. But the logical problem of ordering those dispersed pieces of information in a single, functional dynamic whole does remain. According to my knowledge the constant progress in biochemical sciences reveals new details of the exquisitely precise machinery of living cell, and so the tremendous problem of its integration is more and more pronounced. The solution of this problem seems to be more and more difficult.

Therefore, in my opinion, it is absolutely premature to put all our faith in the DNA molecule and its informational capacity. For the moment the DNA molecule seems to be a quite efficient crib, and, to be sure, a necessary crib. The future discoveries will decide whether the immense gap between the information actually needed to produce the biochemical, cytological and anatomical structures of the living body on one hand, and the information carried by the DNA molecule on the other, will become narrower or, to the contrary, wider and wider.