**THEORY OF MOLECULAR GENETIC REGULATORY SYSTEMS (MGRS): KEY IDEAS AND RESULTS.**

RATNER V.A.

Institute of Cytology and Genetics, (Siberian Branch of the Russian Academy of Sciences), 10 Lavrentieva ave., Novosibirsk, 630090 Russia

Keywords: molecular genetic regulatory systems, regulaton of metabolism, genetic language

The theory of molecular genetic regulatory systems aroused in the mid-60s as an application of the ideas and approaches of cybernetics (J. von Neuman, A.A. Lyapunov, I.A. Poletaev) to rapidly developing molecular biology and genetics. By that time, the informational cybernetic approach had yielded the first fruits. In 1964-1966, the foundation of the theory of MGRS was laid (V.A. Ratner). In the years that followed, various directions of the theory were being developed (M. Eigen et al., V.A. Ratner et al., #Т. Kauffman, R. Thoma, M. Savageau#, etc.). Computer genetics and the theory of molecular evolution (TME), which are actually based on the ideas of MGRS theory, progressed dramatically in 80s-90s.

**Concept** (Ratner 1966, 1975, 1993). The totality of irregular cell biopolymers (DNA, RNA, proteins, and their complexes) along with the systems carrying out various processes involving these biopolymers are considered as a **MGRS** that performes various operations with **genetic information.** Sequences of monomer's symbols are considered as **genetic texts;** the rules according to which the information is written down, as the **genetic language;** and schemes of regulatory interactions between genes, enzymes, metabolites, etc., - as **genetic networks.** **Genetic information** is determined as an aggregate of molecular and more complex biological properties and functions encoded in the genetic texts. The genetic texts considered from the standpoint of long-term storing of genetic information form the **genetic memory.** We are applying this approach ubiquitously.

**Block-modular organization of MGRS**(Ratner, 1992, 1993). Actual MGRS of the cell display a hierarchical block-modular organization formed in the course of evolution. The major modules of a MGRS are: (a) the archive of genetic information (hereditary memory); (b) block of self-reproduction; (c) block for control of metabolism; (d) for control of development; (e) for control of physiological reactions; etc. The modules (a) and (b) form the **central subsystem of MGRS;** the rest, **peripheral subsystem.** The block of self-reproduction (syser) forms the core of MGRS organization. Most known dynamic modules of lower organizational levels are codons, marks of punctuation, genes, proteins, operons, replicons, mobile genetic elements (MGE), etc.

**Self-reproduction.** M. Eigen (1979, 1979) developed the theory of self-reproducing ensembles of macromolecules basing on the **catalytic hypercycle** with translation and nonuniversal replication. V.A. Ratner and V.V. Shamin (1980-1983) studied the role of replication universality and demonstrated that the necessary condition for stability of ensemble's structure demanded the limitation of the number of fractions of nonlinked matrices **n >= k+1,**where **k** is the number of different replication processes. As the real replication processes in cells are universal (**k = 1**), a model of cyclic ensemble with universal replication---**syser**---was proposed. Construction of the syser formed the basis of the self-reproduction module of MGRS, which includes the blocks of basic genetic processes: replication, transcription, translation, and segregation.

**Archive of genetic information** is a module for storing, changing, and operating with genetic information, that is, the hereditary memory. It contains the blocks of the following basic genetic processes: reparation, recombination, transposition of MGE, rearrangements, etc. It provides for noise immunity and variation of genetic information. Mathematical and computer-assisted analyses of the principles of the archive noise immunity and occurrence of various boundaries of **"the catastrophe of errors and losses"** have been carried out by J. #Y.B.S. Haldane (1957), M. Eigen (1970, 1979), Batchinsky and Ratner (1976), and Ratner et al. (1985, 1996).

**Block for regulation of metabolism** is composed of hundreds and thousands autonomous systems of metabolic control, working almost independently and in parallel or organized into complex cascade systems. The complexity of this block is close to the complexity of metabolic pathways themselves. Simplest models of operons and operon systems have been described by #J. Monod and F. Jacob (1961), B. Goodwin (1966), T. Kauffman (1974), R. Thomas (1979),# Ratner, Tchuraev et al. (1966-1998). They took into account a diversity of direct interactions and the feedbacks of the genes, enzymes, regulatory proteins, metabolites, signaling agents, namely: repression, induction, retroinhibition, activation, etc. In a more up-to-date form, such modeling is based on employment of a computer-assisted method of **gene networks** (Kolchanov et al.). A detailed mathematical analysis of the kinetics of metabolic chains under enzymatic and genetic control was carried out by #M. Savageau (1976) and #H. Kacser et al. (1980-1985). They developed the method of control coefficients, allowing the limiting metabolic links in the genetic control of the products to be revealed. The idea of I.A. Poletaev on the **limiting factor** (1970, etc.) is very efficient and advantageous in this filed as a method of approximation.

**Block for development regulation** controls the temporal dynamics and spatial topography of ontogenetic events. It is most essential for the MGRS of multicellular eukaryotes, although individual cells and even viruses inside the cells undergo development. Organization of this module is suggested to center around the idea that the program of ontogenesis is not recorded in the DNA archive as a continuous text (Altan, Koppel, 1991) but represented indirectly through the interaction of molecular components of MGRS. As a result, a dynamic scheme of MGRS function and development is formed as a gene network to realize this program. The program for development also has a block-module structure containing sequential and alternative subprograms. One of the first models of the dynamic system controlling and realizing the intracellular development of  phage was constructed by Tchuraev, Kananyan, and Ratner (1974-1983). It was actually the first case of knowledge base application to a particular field.

**The modules absent in MGRS** (Ratner 1975, 1993). MGRS lack the block for recording the external information into genetic memory, in particular, the units for (1) recording of genetic information from the alphabet of the processes proceeding with time into the alphabet of genetic texts; (2) encoding the external images into the alphabet of genetic texts; and (3) recoding the polypeptide texts in polynucleotide texts. No new information can be written down into the hereditary memory, although inclusion of the already recorded information is not forbidden. The reason is in the peculiarities of the genetic language.

**Genetic language of MGRS** is a system of rules and regularities for encoding the genetic information in the genetic texts (Ratner, 1974, 1975, 1993). The key property is: the rules are not recorded in the genetic memory but in a way inherent to the material carriers of this memory. This language is the language of formation of the MGRS macromolecular components, their interaction, realization of the information contained, perception of the signals, etc. In other words, this is the language of MGRS in all the diversity of their manifestation. Genetic language is specified by (1) the **alphabet,** a set of symbols of canonical monomers or text modules; (2) **grammar,** a set of rules for creating genetic texts; (2) a set of **marks of punctuation and regulation;** and (4) **semantics,** a list of elementary molecular properties and functions and the rules of their combining and correspondence in the genetic texts. We have described a variety of general rules of the grammar, punctuation, semantics, synonymy, and correspondence of genetic texts and functions (Ratner, 1993) according to the model "Meaning <=> Text", developed by I.A. Mel'chuk (1974).

Genetic language is organized hierarchically and forms a **linguistic system** (Ratner 1966, 1975, 1993): codons, cistrons, scriptons, replicons, linkage groups as well as a number of intermediate sublevels. Typical of this system is a number of general linguistic regular patterns. That is why, despite the suggestions of many researches to introduce new languages for description of other regularities, I believe that they should be mandatory included into the linguistic system. A profound analogy can be traced between the linguistic systems of the genetic language, natural human languages, and computer languages.

**Block-modular reorganization of MGRS via MGE** (Ratner et al., 1992, 1996). MGE amount up to 10% of the eukaryotic genome, maintain the "engine" for their own transposition, and contain a variety of functional site motifs. They can be considered as **"movable cassettes of functional sites"**, which change the regulation of genes and polygenes by changing their location. They can be induced by stress external and physiological internal effects. A plenty of examples (Britten, 1996) demonstrate the involvement of MGE in regulation of the adjacent genes.

**MGRS are the central object of the theory of molecular evolution (TME)** (Ratner et al., 1985, 1992, 1996). MGRS are an apparent object of molecular evolution. We have attempted successfully to develop a Unified TME based on the concept of MGRS. This approach allows the unification of essentially heterogeneous methods, approaches, objects, facts, hypotheses, and conceptions. MGRS are the first regulatory systems, and the genetic language is the first linguistic system having originated and being preserved in nature. The block-modular principle of MGRS organization reflects the stages of the evolutionary process.