Permissible Base Sequences for non Self Complementary Messenger RNAs


Departamento de Matematicas * y Departamento de Fisica, Biofisica, Facultad de Ciencias
Fisicas y Matematicas, Universidad de Chile


It is a well accepted assumption that the transcription of DNA is done along a single strand (Tocchini-Valentini and cols. 1, Speigelman and Hayashi 2, Marmur 3 and Tohá C. 4) and that the biosynthetic mRNA obtained is monotonic and is read in a linear fashion during the process of translation. In order to have a good interaction with the ribosome and an easy reading of the message the mRNA should remain extended (Okamoto and Takanami 2; Khorana 6). We assume that this state is due to a lack of complementary zones in the strand of mRNA. If the assumption is valid, we can imagine that the sequences of basis in the messenger is substantially limited by this restriction that excludes the existence of complementary zones.

We have studied with the help of a computer * program the possibility of building up messengers that will not present complementary zones longer than six neighbouring basis. This number has been chosen taking into account that the temperature of fusion of oligonucleotides of three to four basis is less than 37 °C (Lipsett, Heppel, and Bradley 7, and that the double strand of a sRNA is stable at that temperature. On this account we have assumed that strands with complementary zones longer than six contiguous basis would present a double structure which could impede or at least make quite difficult a process of translation. The complementarity of the mRNA molecule could be parallel or anti-parallel (see fig. 1). This two types of complementarity can restrict the possible mRNAs corresponding to heavy proteins, in such a way, that their determination could be utilized to find out the sequence of basis in the DNA.

The number of different sextupletes that can be obtain with four different basis is 4^6 (approximately 4,000). It is easy to see that half of them are complementary with the other half. In this manner we can forecast that it is possible to build an mRNA of about 2,000 basis without having complementary sextupletes assuming only antiparallel configurations. If we take into account also a parallel configuration this maximum length will be reduced approximately to one half given a number of basis of the order of 1,000. A messenger of this length can codify a protein consisting of about 300 aminoacids.

To test the validity of our assumptions we have explored the probability of obtaining messengers without complementarity for natural and ficticious

* A FORTRAN PROGRAM is available for those who may be interested.
** Fellow of Comisión Nacional de Investigaciones Científicas y Tecnológicas.

4 J. Tohá C. Z. Naturforschg. 20b, 868[1965].
5 T. Okamoto and M. Takanami, Biochim. biophysica Acta [Amsterdam] 76, 266 [1963].
proteins. The corresponding probabilities are shown as curves A, B and C in fig. 2. Natural proteins are those that exist in reality and we have used Human Hemoglobin α chain, Horse Heart Cytochrome, Half Human Haemoglobin α chain, Beef Insuline, Human ACTH and Human β melanocyte S. Hormone to built curve A.

Ficticious proteins have been generated at random by us using an aminoacid pool (curve B). Each point in curves A and B shows the number of messengers, free of complementary zones, obtained in hundred trials designed to generate the mRNA corresponding to a given protein. This generation was simulated with a computer program.

The curve C corresponds to the calculated probabilities of obtaining mRNAs without complementarity for the same fictitious proteins used for building the curve B.

It is very interesting to observe from fig. 2 that the probability of success in the case of natural proteins is significantly higher ($P < .01$) than in fictitious proteins. Apparently non-complementarity has played a role in the natural selection process of proteins, specially the big ones. Even if this restriction is not absolute it seems to us that non-complementarity statistically favours the process of translation and gives a trend to the selection of mRNAs in cellular evolution.

Besides it is possible that the existence inside the cell of messengers with and without complementarity could serve as a regulation mechanism of protein synthesis and cellular differentiation and could explain the existence in virus infected cells of different types of messengers: “early messengers” (free of complementarity) and “late messengers” (with complementarity). In the same way mRNA with complementarity may operate in cells with threshold temperature or, in hibernating cells, or could be responsible for protein synthesis in non-nucleated cells. If a mutation creates a complementary zone in an mRNA the difficulties in translation involved can well be an important control mechanism against genetic errors, specially taking into account that the transcription involves in many cases a polymer of long chain formed by a set of mRNAs.

In this way in order to have for big proteins a surviving spontaneous mutation two or more contemporary changes in the mRNA would be needed to maintain non-complementarity. However, if complementarity is produced by a single mutation in two neighbouring sextuplets, the corresponding mRNA can still serve for purposes of translation, due to the steric inability of those two neighbouring sextuplets for interacting.

It is interesting to observe (see fig. 3) that the proportion of aminoacids that correspond to the more degenerated part of the code (three or more codons per aminoacid), is greater in bigger proteins ($r = .7; P < .01$). As all individual codons have rather similar probability of complementarity, this bigger proportion must be related to some requirements of the chain of the RNA, as it grows.

---

Fig. 2. Curves of probabilities of obtaining messenger RNAs without zones of self complementarity of sextuplets or bigger than sextuplets from proteins of different sizes. A: Natural proteins. B: Ficticious proteins, generated from a random aminoacid pool. C: Calculated curve of probabilities for ficticious protein. Each point of curves A and B corresponds to computed hundred trials.

---

8 B. Keil, Annu. Rev. Biochem. 31, 152 [1962].
This could be a better protection against mutation and or the avoidance of complementarity.

To show in a more direct manner the need of non complementarity in mRNA, we are working, with a model of subcellular protein biosynthesis, in which tRNA and ribosomes are previously heated beyond the melting point ($T_m$) of tRNA. In this conditions tRNA apparently acts as mRNA.

** This work has been partially supported by a grant of the "Comisión Nacional de Energía Nuclear" y "Comisión Nacional de Investigación Científica y Tecnológica".
14 B. Keil, Annu. Rev. Biochem. 34, 176, 178 [1965].
15 V. Tomasek, O. Mikes, V. Holeysorsky, and F. Sorm, Collect. czechoslow chem. Commun. 29, [1964].