



HAL
open science

A Dynamic Representation of mRNA Nucleotides Clarifies the Conundrum of Codon Redundancy

Daniele Funaro

► **To cite this version:**

Daniele Funaro. A Dynamic Representation of mRNA Nucleotides Clarifies the Conundrum of Codon Redundancy. 2023. hal-04098442

HAL Id: hal-04098442

<https://hal.science/hal-04098442v1>

Preprint submitted on 16 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Dynamic Representation of mRNA Nucleotides Clarifies the Conundrum of Codon Redundancy

Daniele Funaro

Dipartimento di Scienze Chimiche e Geologiche
Università di Modena e Reggio Emilia
Via Campi 103, 41125 Modena (Italy)
daniele.funaro@unimore.it

Abstract

The deciphering of the genetic code takes place through the reading of the nitrogenous bases, which are 4 in number. In most cases, the bases are taken three by three, thus generating 64 possible combinations with repetition. Each combination (codon) allows for the synthesis of a specific amino acid. Since the latter are only 21 in number, the codon-amino acid conversion table shows a strong redundancy. Countless efforts have been made to understand the true encryption mechanism. Here we want to add our version, which consists of associating a periodic sound based on 3 notes to each codon. RNA now becomes a dynamic object and not just a list of static instructions. In addition to a different interpretation of the genetic code, there is also a considerable reduction in redundancy, given that the number of periodic sounds that can be produced with 3 notes drops to 20 (with the addition of 4 pure frequencies). Finally we discuss the possibility of how these sounds can be generated and travel inside the double helix, and possibly emitted as biophotons.

Keywords: RNA, codon degeneracy, electric circuit, biophoton.

1 Introduction

It is well known that RNA is a string of nitrogenous bases containing the sequence of instructions for constructing amino acids, as well as other interesting information. The organization suggests a strictly deterministic procedure. Once having established the beginning of a sequence, the nucleotides are considered three by three, thus forming a series of *codons*. Each codon is therefore identified by a triplet of symbols, taken from the four available, i.e.: U (*Uracil*), C (*Cytosine*), A (*Adenine*), G (*Guanine*). There are 64 possible combinations. Each of them allows us to identify a single amino acid according to the famous

conversion chart (see table 1), obtained thanks to the pioneering efforts of M. Nirenberg and coworkers [1].

Table 1: Codon table

UUU → Phe UUC → Phe UUA → Leu UUG → Leu	UCU → Ser UCC → Ser UCA → Ser UCG → Ser	UAU → Tyr UAC → Tyr UAA → Stop UAG → Stop	UGU → Cys UGC → Cys UGA → Stop UGG → Trp
CUU → Leu CUC → Leu CUA → Leu CUG → Leu	CCU → Pro CCC → Pro CCA → Pro CCG → Pro	CAU → His CAC → His CAA → Gin CAG → Gin	CGU → Arg CGC → Arg CGA → Arg CGG → Arg
AUU → Ile AUC → Ile AUA → Ile AUG → Met	ACU → Thr ACC → Thr ACA → Thr ACG → Thr	AAU → Asn AAC → Asn AAA → Lys AAG → Lys	AGU → Ser AGC → Ser AGA → Arg AGG → Arg
GUU → Val GUC → Val GUA → Val GUG → Val	GCU → Ala GCC → Ala GCA → Ala GCG → Ala	GAU → Asp GAC → Asp GAA → Glu GAG → Glu	GGU → Gly GGC → Gly GGA → Gly GGG → Gly

The result is extraordinary but not perfectly convincing from the theoretical viewpoint. In fact, there are only 21 amino acids so there are several codons that lead to the same result. The table therefore displays some redundancy (called *codon degeneracy*) and there does not seem to exist a clear law that allows the combinations to be thinned out in a logical way. In particular, there are 8 amino acids having 4 or more codon representations (*Ser, Leu, Pro, Arg, Thr, Val, Ala, Gly*) and 12 amino acids having 1 to 3 codon representations (*Phe, Leu, Tyr, Cys, Trp, His, Gin, Ile, Asn, Lys, Asp, Glu*). Finally, there are 4 special codons. One of them corresponds to the amino acid *Met*, that determines the beginning of a protein. Each of the last 3 codons represent the command *stop* that marks instead the end of a protein. Many attempts have been made in the past to try to better formalize the table (see, e.g. [2]). Given that nature makes its resources available in a very parsimonious way, the aim is to understand if there are hidden relationships that justify the existence of 64 combinations when in the end only a third would be necessary.

The idea we propose here is to interpret the genetic code in a dynamic way. RNA is not trivially a score to be read and performed by other, more or less complex, molecules. Instead, it is itself the entire musical instrument, which decrypts and at the same time plays appropriate chords associated with each codon. The trick is to assume that the message carried by a codon can be repeated periodically so as to become a sort of sound. Different amino acids will correspond to different sounds, but there will be however less sounds than codons. Let's do some calculations.

Suppose to have a list of four symbols $\Gamma, \Delta, \Xi, \Omega$. We can consider all possible triplets that can be formed using the above characters. We know that all combinations amount to 64. We then take the periodic signal generated by the repetition of a given triplet for an indefinite number of times. For example, if we choose $\Omega\Gamma\Xi$, the corresponding "sound" is $\dots\Omega\Gamma\Xi\Omega\Gamma\Xi\Omega\Gamma\Xi\Omega\Gamma\Xi\Omega\dots$. The question is: how many different sequences can be built? First of all, there exist 4 triplets $\Gamma\Gamma\Gamma, \Delta\Delta\Delta, \Xi\Xi\Xi, \Omega\Omega\Omega$, that generate trivial constant monochromatic sequences. We are left with other 60 more interesting cases. Let us now observe that the sequence originating for example from $\Omega\Gamma\Xi$ coincides with those generated by $\Gamma\Xi\Omega$ or $\Xi\Omega\Gamma$. Indeed, the above three combinations give rise to the same sound, therefore they are redundant. This means that it is possible to group within a single class the 3 triplets that produce a specific sound. Dividing 60 by 3, we get 20 classes, i.e. 20 different possible sounds. Incidentally, with the exception of *Met*, 20 is also the number of amino acids. The coincidence is somewhat surprising. But we haven't finished yet. There are 8 sounds where the symbols are all different. It may not be a just coincidence that exactly 8 amino acids admit representations with 4 or 6 redundant codons. There are 12 sequences where a given symbol appears twice, such as for example: $\dots\Delta\Delta\Gamma\Delta\Delta\Gamma\Delta\Delta\Gamma\Delta\Delta\Gamma\Delta\Delta\dots$. Do they correspond to the other 12 amino acids? There are finally 4 monochromatic sounds. Do they correspond to the start and stop signals?

In section 2, we will show how to generate 20 + 4 different sounds starting from certain assigned frequencies. Let's say right away that we will not be able to fully solve the problem of rewriting with a minimal number of entries the conversion chart of Table 1. In practice we will not be able to associate in automatic way the generic symbols $\Gamma, \Delta, \Xi, \Omega$ with the classic ones representing the nitrogenous bases, i.e.: U, C, A, G. Indeed, it will be soon evident that a straight relation does not hold. We think that, physically, a set of three ordered codons constitutes a unique oscillating electric circuit which elaborate three resonating frequencies (described by some of the symbols $\Gamma, \Delta, \Xi, \Omega$) which in turn produce a characteristic sound. We do not yet know exactly how these circuits work. We provide some ideas in section 3. The purpose of this paper is mainly to interpret the genetic code encrypted in the DNA in an active way, which is realized through the emission of electromagnetic messages at a distance. For the reasons that will be better expressed in section 3, we believe that DNA is not only a data container but also transmits its data into the environment. At the moment this conjecture allows us to understand why the symbols U, C, A, G produce redundant combinations, while this is no longer verified if we use

symbols $\Gamma, \Delta, \Xi, \Omega$. The connection between the two sets of symbols is yet to be found. Note that we are not just sidestepping the problem with semantic speculation, we are offering a new way to understand how DNA works.

2 Sound generation

We propose the following values for our parameters:

$$\Gamma = 1 \quad \Delta = 2 \quad \Xi = 4 \quad \Omega = 5 \quad (1)$$

Such a choice is arbitrary, meaning that other choices can work in a similar way. We denote by $(\omega_a, \omega_b, \omega_c)$ the generic triplet of values obtained by choosing randomly (with possible repetition) three members from the set $\{\Gamma, \Delta, \Xi, \Omega\}$. According to what has been said in section 1, we want the triplet $(\omega_a, \omega_b, \omega_c)$ to be equivalent to $(\omega_c, \omega_a, \omega_b)$ and $(\omega_b, \omega_c, \omega_a)$ (in fact, they all generate the same sound). Instead, $(\omega_a, \omega_b, \omega_c)$ will not be equivalent in general to $(\omega_a, \omega_c, \omega_b)$, corresponding to another sound in our construction. Successively, we set $\delta = 2\pi/3$ and define the functions of the variable x :

$$f_{1,a}(x) = \frac{\sin[\omega_a(x - \delta)] \sin(\omega_a x)}{\sin(2\omega_a \delta) \sin(\omega_a \delta)} \quad (2)$$

$$f_{2,b}(x) = \frac{\sin[\omega_b(x - \delta)] \sin[\omega_b(x + \delta)]}{[\sin(\omega_b \delta)]^2} \quad (3)$$

$$f_{3,c}(x) = \frac{\sin[\omega_c(x + \delta)] \sin(\omega_c x)}{\sin(2\omega_c \delta) \sin(\omega_c \delta)} \quad (4)$$

$$f_{4,a,b,c}(x) = \frac{1}{2} \sin[\omega_a(x + \delta)] \sin(\omega_b x) \sin[\omega_c(x - \delta)] \quad (5)$$

$$f_{a,b,c}(x) = f_{1,a}(x) + f_{2,b}(x) + f_{3,c}(x) + f_{4,a,b,c}(x) \quad (6)$$

The above are periodic functions with period equal to $3\delta = 2\pi$. For a fixed a , we have the relations, $\forall x$:

$$f_{1,a}(x) + f_{2,a}(x) + f_{3,a}(x) = 1 \quad f_{a,a,a}(x) = 1 - \frac{1}{8} \sin(3\omega_a x) \quad (7)$$

$$f_{1,a}(x + \delta) = f_{3,a}(x) \quad f_{1,a}(x - \delta) = f_{2,a}(x) \quad (8)$$

We finally have:

$$f_{a,b,c}(x + \delta) = f_{b,c,a}(x) = f_{c,a,b}(x - \delta) \quad (9)$$

This last property says that an even permutation of the indices produces the same wave, up to a phase change. According to the second equation in (7), when all indices are equal to a we get a monochromatic wave of frequency ω_a/δ . Otherwise, using (1), we obtain 20 different sounds which are plotted in Figs. 1 and 2. As we said, the idea is to associate each sound with an amino acid, and vice versa. In this way there is no more redundancy. The problem remains of how a sound can be generated from a single codon, respecting the rules of the conversion table 1.

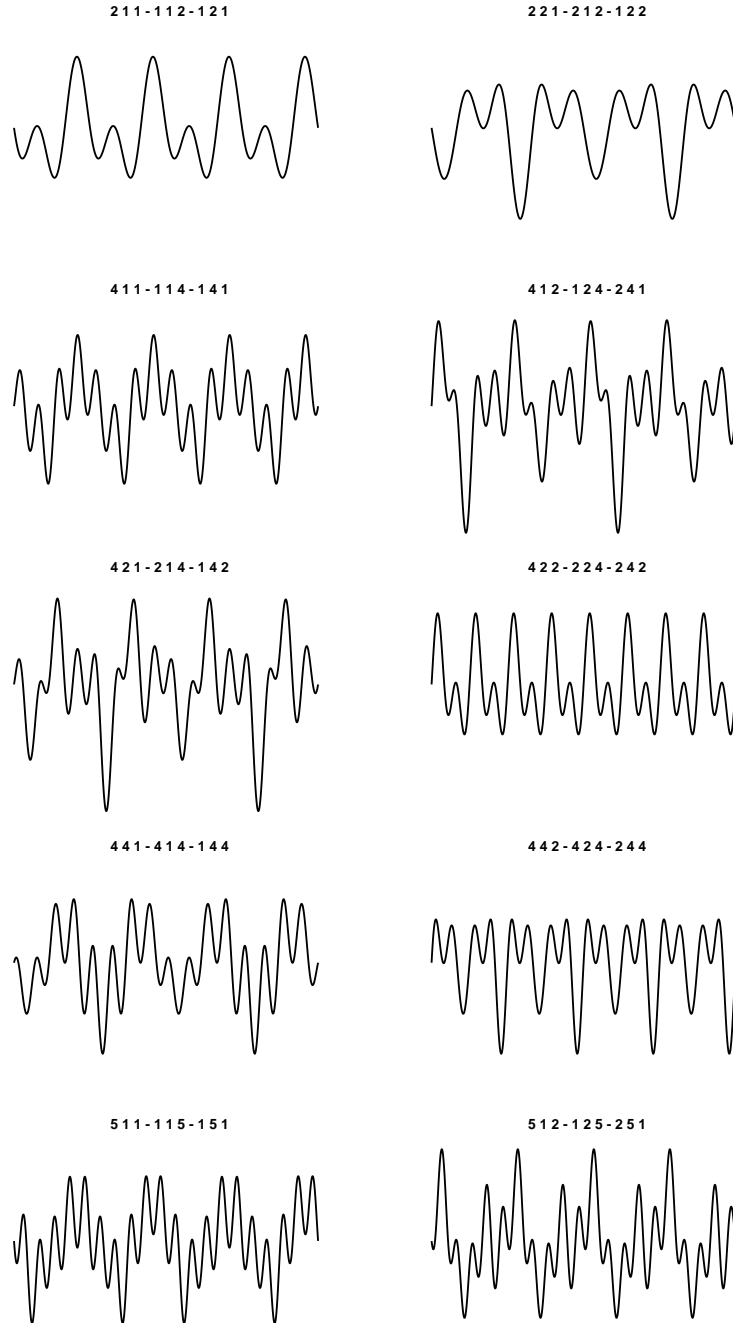


Figure 1: Sounds produced by the triplets specified at the top of each plot.

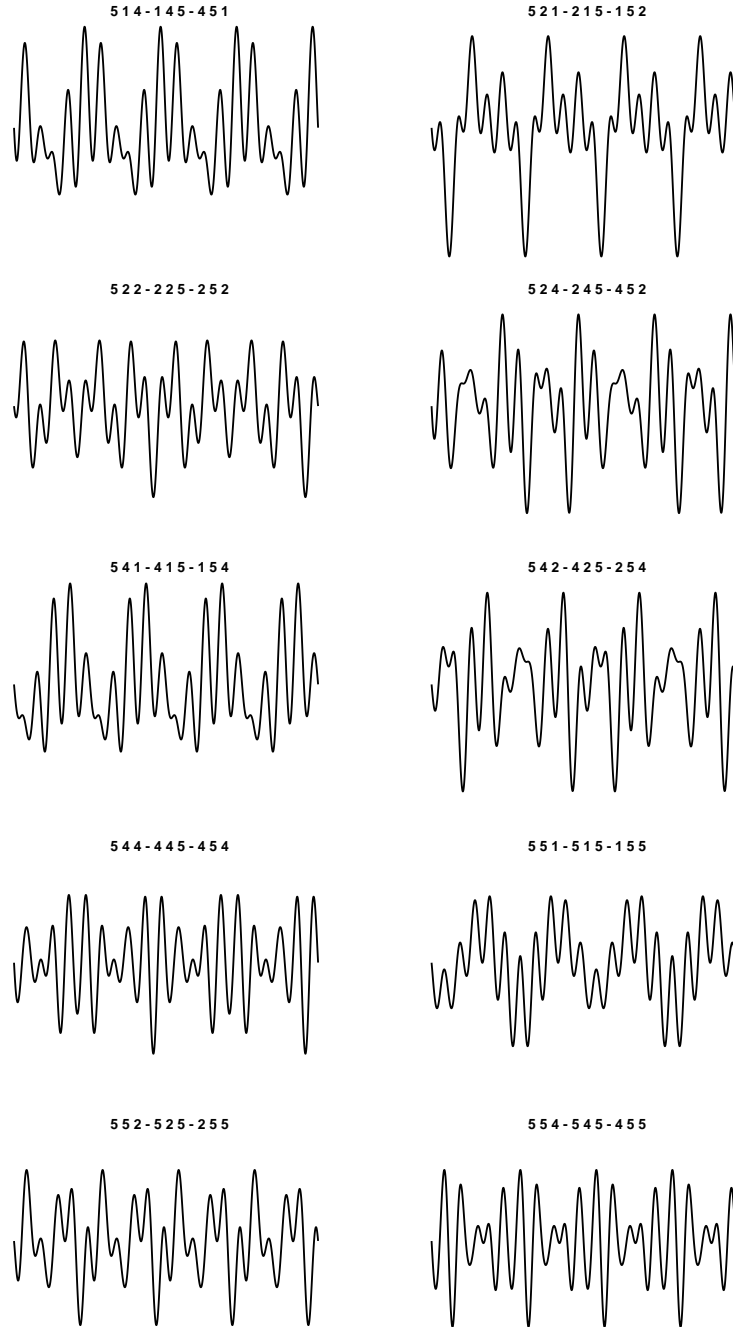


Figure 2: Sounds produced by the triplets specified at the top of each plot.

3 DNA circuitry

Electric circuits displaying a phase shift equal to $2\pi/3$ can be easily built (see Fig. 3). Signals such as those studied in the previous section can be generated by collecting the outputs of three devices of this type resonating at different frequencies. Each circuit has its own RC network and the sum of the frequencies is done by imposing suitable shifting parameters so to recover periodic composite waves.

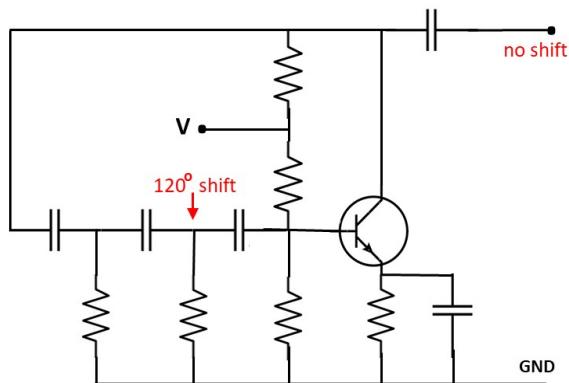


Figure 3: Typical oscillating electronic circuit with a $2\pi/3$ phase shift.

It has been ascertained under various circumstances that the DNA double helix can be associated with some electric circuit. The results are both experimental and computational. They confirm for instance the conductivity of the periodic phosphate backbone [3, 4], thanks to its one-dimensional periodic structure. Electronic properties of DNA and its nucleotides have been studied for instance in [5, 6, 7, 8]. We observe that the presence of nitrogen can give compounds a tendency to flip from one configuration to another, activating spontaneous oscillations. Concatenations of logic gates using basic molecules as switches have been taken into consideration [9, 10]. A model of DNA composed by multiple oscillatory RLC circuits is examined in [11]. These micro circuits are joined by the phosphate backbones that collect the data. All these results reinforce the idea that DNA should be interpreted in a dynamic way. Therefore, together with the intrinsically static message written in the codons, we also

must consider a high probability that this should be assigned a periodic sound, which brings us back to the content of this paper.

We have no element at this time to continue our investigation. A way should be proposed to assign to each nitrogenous base an electric circuit. The global circuit associated with each codon should then produce a periodic sound (among the 20+4 possible). In principle this is not difficult. However, it is difficult to respect the hidden rules of the codon table. For example, the generic sequence GGX (where X is any of the symbols U, C, A, G) corresponds to *Glycine* (last box in Table 1). For some unknown reason, the electric circuit associated to this codon resonates with the couple GG and inhibits the presence of the third component. We undertake to study this issue further in a future paper. Meanwhile let's add some other considerations in favor of a vision in which DNA is an instrument actually played, and not just a static collection of data.

The analysis of vibration frequency maps carried out for nucleobases [12, 13] shows strong DNA activity in producing or absorbing biophotons. Such an intense electromagnetic turmoil regulates the most fundamental biological processes. This aspect strengthens the hypothesis of signals originating from DNA, which remotely guide the main functions of a single cell, but which can extend further, transferring information between groups of cells [14, 15, 16]. These aspects have also been taken into consideration in [17] (section 3.7) and [18] in a more general context, where the chemical elements themselves arise from an agglomeration of atomic nuclei and electromagnetic manifestations that connect them in a peculiar way characterizing the compound.

Finally, we must not forget that DNA is a twisted strand that is difficult to access from a physical viewpoint. Moreover, the cell is a system overcrowded with molecules [19, 20], in which it is practically impossible to create connections except through communications that take place with the exchange of electromagnetic impulses. Of course, it is necessary to assume that suitable receiving stations exist. The RNA polymerase could be one of them, thus eliminating direct contact with DNA, which appears unrealistic in *in vivo* situations.

4 Conclusions

Recent research shows more and more that living organisms are an extended set of molecules (variably complex) that communicate through the exchange of electromagnetic messages. DNA itself is a source of biophotons. In this context it is natural to hypothesize that the genetic code should mainly be interpreted in a dynamic way, which occurs for example through the emission of periodic signals (sounds in our context). The conjecture is strengthened if one considers that the periodic signals obtainable from three (out of four) independent frequencies are 20+4 in number. This suggests that each codon is a transmitting machine, which assigns each sound the creation of a specific amino acid. Although from a circuitual point of view the codon still remains to be investigated, the representation through sounds allows us to take a step forward in understanding why codon degeneracy occurs in the static version of

DNA, but not in the dynamic one.

References

- [1] Nirenberg, M. W. The genetic code. From *Nobel Lectures, Physiology or Medicine 1963-1970*, Elsevier Publishing Company, Amsterdam, 1972.
- [2] Grosjean, H.; Westhof, E. An integrated, structure-and energy-based view of the genetic code. *Nucleic Acids Res.* **2016**, *44*, 17, 8020-40.
- [3] Brillouin, L. Giant molecules and semiconductors, in *Horizons in Biochemistry* (Kash, M.; Pullman, B.; Eds.), Academic Press, New York, 1962, p. 295.
- [4] Ikeura-Sekiguchi, H; Sekiguchi, T. Attosecond electron delocalization in the conduction band through the phosphate backbone of genomic DNA. *Phys. Rev. Lett.* **2007**, *99*, 228102.
- [5] Bakhshi, A. K.; Otto, P.; Ladik, J.; Seel, M. On the electronic structure and conduction properties of aperiodic DNA and proteins, II: Electronic structure of aperiodic DNA. *Chem. Phys.* **1986**, *108*, 2, 215-222.
- [6] Lewis, F. D.; Liu, X.; Liu, J.; Miller, S. E.; Hayes, R. T.; Wasielewski, M. R. Direct measurement of hole transport dynamics in DNA. *Nature* **2000**, *406*, 6791, 51-53.
- [7] Bixon, M.; Jortner, J. Long-range and very long-range charge transport in DNA. *Chem. Phys.* **2002**, *281*, 2-3, 393-408.
- [8] Giese, B. Electron transfer in DNA. *Curr. Opin. Chem. Biol.* **2002**, *6*, 5, 612-618.
- [9] Andréasson, J.; Pischel, U. Molecules with a sense of logic: a progress report. *Chem. Soc. Rev.* **2015**, *44*, 1053.
- [10] Zhou, Z.; Wang, J; Levine, R. D.; Remacle, F.; Willner, I. DNA-based constitutional dynamic networks as functional modules for logic gates and computing circuit operations. *Chem. Sci.* **2021**, *12*, 5475.
- [11] Xu, K. Stepwise oscillatory circuits of a DNA molecule. *J. Biol Phys.* **2009**, *35*, 223-230.
- [12] Jiang, Y.; Wang, L. Development of vibrational frequency maps for nucleobases. *J. Phys. Chem. B* **2019**, *123*, 27, 5791-5804.
- [13] Prokhorenko, V. I.; Picchiotti, A.; Pola, M.; Dijkstra, A. G.; Miller, R. J. D. New insights into the photophysics of DNA nucleobases. *J. Phys. Chem. Lett.* **2016**, *7*, 22, 4445-4450.
- [14] Mejía Morales, J.; Michaelian, K. Photon dissipation as the origin of information encoding in RNA and DNA. *Entropy* **2020**, *22*, 940.
- [15] Cifra, M.; Fields, J. Z.; Farhadi, A. Electromagnetic cellular interactions. *Prog. Biophys. Mol. Biol.* **2011**, *105*, 3, 223-246.

- [16] Scholkmann, F.; Fels, D.; Cifra, M. Non-chemical and non-contact cell-to-cell communication: a short review. *Am. J. Transl. Res.* **2013**, *5*, 6, 586-593.
- [17] Funaro, D. *From Photons to Atoms: The Electromagnetic Nature of Matter*, World Scientific, Singapore, 2019.
- [18] Funaro, D. The impact of a pervasive electrodynamical background on biological interactions, Preprint (2022), hal.science/hal-03899223v1.
- [19] Zimmerman, S. B.; Minton, A. P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 27-65.
- [20] Ellis, R. J. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **2001**, *26*, 10, 597-604.