

Committee 6  
Science and Music: A Unifying Concept

Draft – February 1, 2000  
For Conference Distribution Only



Music Patterns in Molecular Spectra and Structures

Roland Lefebvre  
Professor Emeritus of Chemical Physics  
University of Paris  
Bures Sur Yvette, France

The Twenty-second International Conference on the Unity of the Sciences  
Seoul, Korea February 9-13, 2000

## I. Introduction

In his book "The Physics and Psychophysics of Music" [1], Roederer tries to answer the question: "What is Music? ...Why do we respond emotionally to complex musical messages, which seem to contain no information of any survival value?...This motivation may well be the result of our inborn drive to train at an early age in the highly sophisticated auditory analysis operations expected for speech perception". This interesting hypothesis is further expressed into the statement that "imitating environmental sounds has never been the prime force driving the development of a musical culture".

In this note, we will examine some of the schemes which have been proposed for a musical transcription of scientific data. Since scientific information is, in some sense, an "environment", at least for scientists, are we watching a development of musical inspiration going beyond the previous rather restrictive statement? And are we going to respond emotionally to such music? These two questions are certainly going to provoke widely different answers, since it is to a large extent a matter of personal taste to appreciate the value of any "new music".

There are many reasons to associate physics and music, beyond the obvious ones related to the way sounds are produced by different instruments. The analysis and construction of musical scales have soon revealed the omnipresence of integers in the rules for the determination of consonances. This is basically due to the fact that multiples of a fundamental frequency are present, through Fourier theorem, in any periodic signal. However integers are also present in many fields of physics, for instance in the quantization of the energy. The energy levels of the hydrogen atom are given (with the so-called atomic units) by the simple formula  $E_n = -\frac{1}{2n^2}$ . The energies of an harmonic oscillator are  $E_n = (n + \frac{1}{2})\omega$ , etc...We must remember the permanent fascination of humans for the whole numbers. Pythagoras declares: "Everything is Number and Harmony". Helmholtz [2] gives other examples. "...According to Tso-kiu-ming, a friend of Confucius (B.C. 500), the five tones of the old chinese scale were compared to the five elements of their natural philosophy: water, fire, wood, metal and earth. The whole numbers 1, 2, 3, and 4 were described as the source of all perfection. At a later time the twelve semitones of the octave were connected with the twelve months of the year". Are we really far from this attitude when we adopt for the definition of the second the time it takes to produce *exactly* 9192631770 transitions between the two hyperfine levels of the Cesium atom in its ground state?

We are going to look at two sources of data which have suggested a musical tran-

scription. The first one is taken from the field of molecular spectroscopy. Since the spectral data can be given in the frequency scale, this could appear to be immediately translatable into sounds. However in most cases a processing of the data is needed, because the frequencies in the spectra are completely out of scale to suggest anything connected to music.

The other example is based on the consideration of data characterizing the biomolecules. The deciphering of the genetic code is one of the major advances in the last half-century: the recurrent sequences of the code have suggested to many people that a "transcription" into music would be of interest, both for mnemotechnic reasons or for aesthetic reasons. Another source is the arrangement of amino-acids in proteins. The abundance of sites on the web on both subjects is a mark of the strong impact of such possibilities.

We will not consider in this brief review a more controversial topic which is the possible effects (good or bad) of such music on living species (plants, animals or human beings).

## II. Musical Transcription of Molecular Spectra

The spectrum of a molecule is a kind of identity card of the system. It says what energies the system can absorb or emit when interacting with an electromagnetic field. These energies can be given on the frequency scale. It is therefore tempting to make a transcription into audible sounds. However two difficulties are on the way:

(1) The frequencies are generally completely outside the audible range even in the spectra of lower frequencies such as that related to transitions between atomic hyperfine lines. Table I indicates in Hertz what are the frequencies which are typically related to the different degrees of freedom of a molecular species. It is immediately clear that some rescaling is needed.

(2) If a set of molecular lines are associated to audible sounds after some kind of rescaling, this would be without aesthetic interest, since the common way to present this information consists in ordering them either in decreasing or increasing order. Some additional processing is therefore needed.

Table I. The various frequencies which are associated to typical molecular transitions. The molecule is made of electrons and nuclei. The transition is said to be electronic when under the influence of radiation the change in structure takes place essentially among electrons. Vibrational and rotational transitions affect the nuclear skeleton. Both types of particles have also a spin, that is to say a freedom somewhat like that of a rotating body. All data with \* concern the hydrogen molecule  $H_2$ . The data with \*\* correspond to an electron or a proton spin in a magnetic field of 0.1 tesla or 1000 gauss.

Degree of freedom	$\Delta E$ (eV)	wavelength ( $\text{\AA}$ )	frequency (Hz)
electronic *	$\sim 11$	1128.	$2.66 \times 10^{15}$
vibrational *	$\sim 0.52$	$2.38 \times 10^4$	$1.26 \times 10^{14}$
rotational *	$\sim 0.015$	$8.26 \times 10^9$	$3.63 \times 10^{12}$
electron spin **	$\sim 0.58 \times 10^{-5}$	$2.14 \times 10^9$	$1.4 \times 10^9$
nuclear spin **	$\sim 3.14 \times 10^{-9}$	$3.95 \times 10^{12}$	$0.76 \times 10^6$

Sydney Leach [3], in his musical transcription of the information contained in the photoelectron spectrum of phosphabenzene (or phosphoridine) has suggested to base the operation on differences between successive frequencies. In this way fluctuating numbers are obtained. We analyze now in detail how this operation is performed. We take from [3] the example of the experimental photoelectron spectrum of this molecule. When a molecule is subjected to a flux of photons having an energy large enough to detach an electron from the molecule, the kinetic energy of the free electron depends on the way it was bound to the system. For a fixed photon energy a weakly bound electron will come out with a comparatively larger kinetic energy than in the case of a strongly bound electron. The analysis of the kinetic energies of the outgoing electrons allows for a measure of their binding energies. These quantities are called ionization potentials and denoted  $I_i$ . Figure 1 is showing the experimental photoelectron spectrum of phosphabenzene [2].

There is a series of twelve peaks which are assigned to the detachment of electrons with increasing binding energies from left to right. These quantities are reported in column 1 of Table II in eV (the energy that an electron acquires when it goes through a potential difference of 1 Volt). To transform these data into a series with fluctuations, Leach defines the quantity  $\Delta_i = I_{i+1} - I_i$ . These are given in millielectron volts (meV) in column 2 of Table II. We have the equivalence  $1meV = 2.418 \times 10^{11} Hz$ . It is thus clear that such frequencies are completely outside the audible range which goes at most to 20000 Hz. With the factor  $k = 2.026 \cdot 10^{-12}$ ,  $\Delta_i$

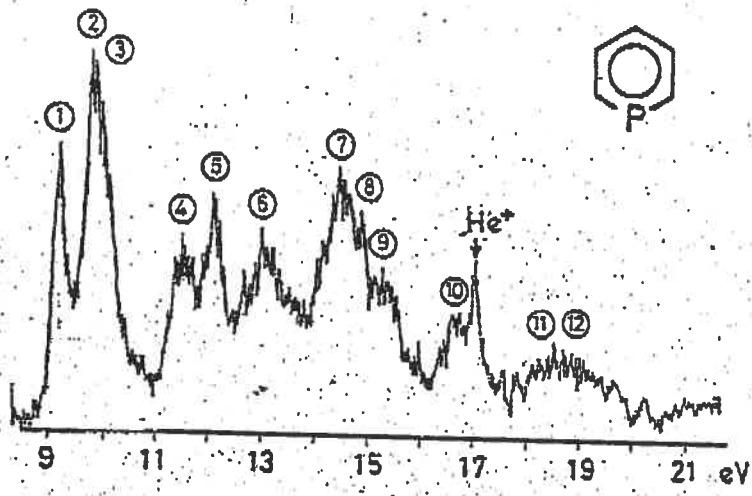


Figure 1: The photoelectron spectrum of phosphabenzene. In the upper right corner the structure of the molecule. Excitation radiation is from a discharge lamp containing Helium. Twelve peaks are obtained, with various intensities. They make it possible to estimate the twelve ionization potentials given in Table II (after A. J. Ashe et al., *Helv. Chim. Acta*, 59, 1944 (1976)).

is converted into the frequency given in column 3. By looking at the frequency of the equal temperament scale closest to a calculated frequency, we get the notes given in columns 4 and 5, depending on the convention for notation. The transcription is not yet achieved. In order to assign a duration to each note S. Leach makes it proportional to the intensity of peak number *i* of the spectrum for note *i*. The sequence of notes obtained in this way is given in Figure 2. Reference [3] contains a composition written by G. Englert based on the three "muispectra" derived by Leach from the experimental spectrum and from two theoretical spectra of references [4] and [5]. We give in Figure 3 a variation made by Marja Rantanen-Maruani on the theme given in Figure 2.

Table II. Transformation of the data from the photoelectron spectrum of phosphabenzene into notes of the musical scale. Column 1 gives the measured ionization potentials in eV. Columns 2 and 3 give successive differences of ionization potentials either in meV or in Hz. They are transformed (column 4) into audible frequencies  $\nu_i$  by using the factor  $k = 2.026 \times 10^{-12}$ . The frequency  $\nu'_i$  of the closest note in the equal temperament scale is given in column 5. Column 6 gives the notes in two notations (CDEFGABC or do ré mi fa sol la si do). s stands for sharp and d for dièze. The frequency of the note A3 (or la3) is assigned the value 440 Hz. The durations (column 7) are proportional to the intensities of the successive peaks of the spectrum. The unit is the sixteenth.

Exp. $I_i$ (eV)	$\Delta I_i$ (meV)	$\Delta I_i$ (Hz)	$\nu_i$ (Hz)	$\nu'_i$ (Hz)	note	duration
9.2	600	$1.451 \times 10^{14}$	294	294	D3 (ré3)	8
9.8	200	$4.836 \times 10^{13}$	98	98	G1 (sol1)	12
10.0	1500	$3.627 \times 10^{14}$	735	740	F4s (fa4d)	12
11.5	600	$1.451 \times 10^{14}$	294	294	D3 (ré3)	10
12.1	1000	$2.418 \times 10^{14}$	490	494	B3 (si3)	8
13.1	1400	$3.385 \times 10^{14}$	686	698	F4 (fa4)	9
14.5	400	$9.672 \times 10^{13}$	196	196	G2 (sol2)	12
14.9	400	$9.672 \times 10^{13}$	196	196	G2 (sol2)	12
15.3	1500	$3.627 \times 10^{14}$	735	740	F4s (fa4d)	6
16.8	1900	$4.594 \times 10^{14}$	931	932	A4s (la4d)	5
18.7	600	$1.451 \times 10^{14}$	294	294	D3 (ré3)	3
19.3	—	—	—	—	—	—

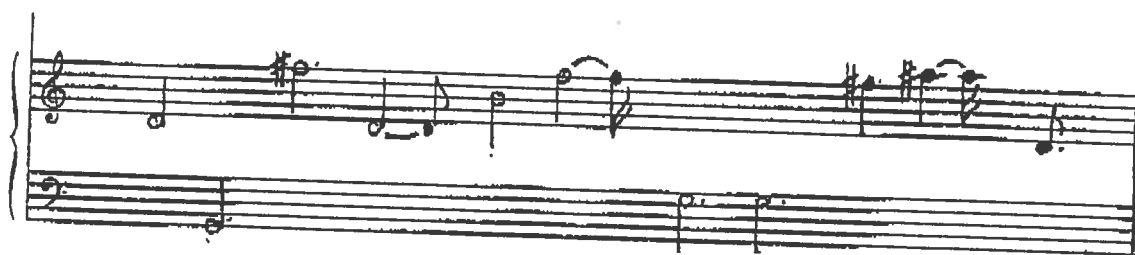


Figure 2: Musical Transcription of the Experimental Photoelectron Spectrum of Phosphabenzene. The third note differs from that given in Reference [3]

The image displays three systems of handwritten musical notation for piano. Each system consists of a grand staff with a treble clef on the upper staff and a bass clef on the lower staff. The first system begins with a treble clef and a key signature of one sharp (F#). The first measure of the first system is marked with a piano-piano dynamic (*pp*), and the second measure is marked with a piano (*p*) dynamic. The notation includes various rhythmic values, including eighth and sixteenth notes, and rests. The second system continues the piece with similar rhythmic patterns and dynamics. The third system concludes the piece with a double bar line and a repeat sign.

Figure 3: Variation made by Marja Rantanen-Maruani, Composer, on the musical theme given in Figure 2, based on the composition made by G. G. Englert in reference [9].

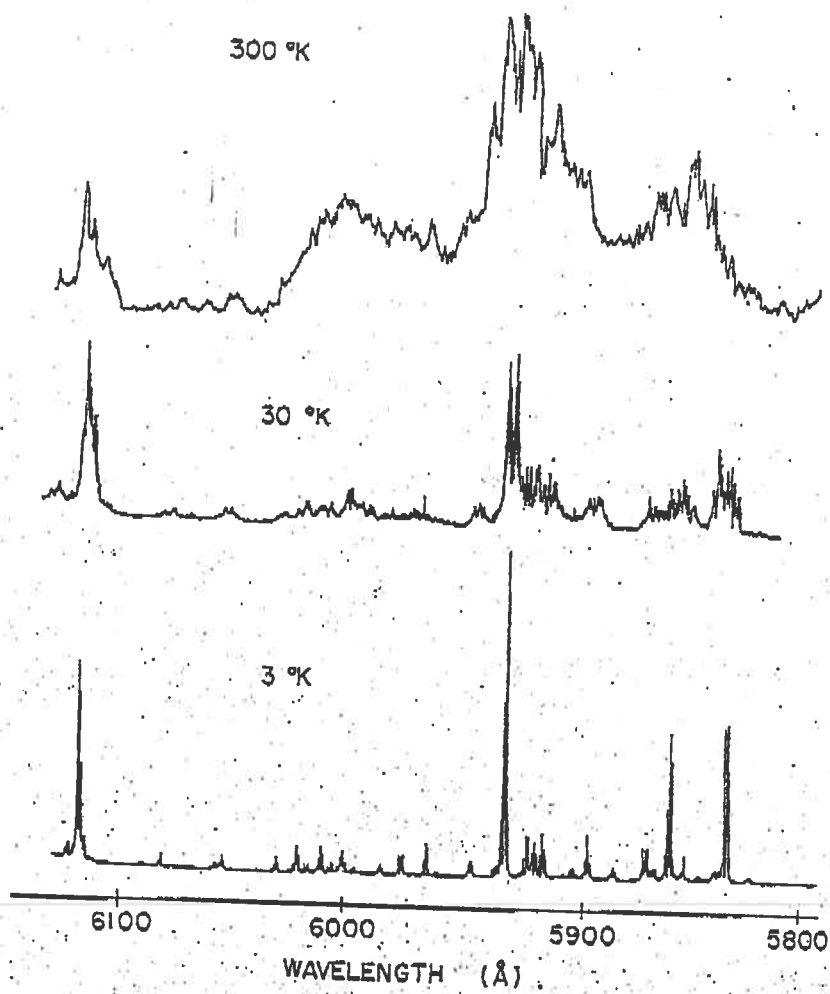


Figure 4: Improvement of experimental techniques transforms a spectrum with very poor information content into a rich set of data awaiting a "musical" transcription (after R. E. Smalley et al., *J. Chem. Phys.*, 61, 4363 (1974)), with permission.



The scientific literature is full of atomic or molecular spectra with very detailed information. Recent experimental techniques have considerably improved our knowledge of transition energies. Figure 4 gives an example of the evolution of spectra, starting from very badly resolved profiles to extremely rich content. Such spectra and many others are awaiting for a musical transcription.

### III. From Biomolecules to Music

This review is very far from being exhaustive. A recent book by Y. Fukagawa [6] describes some of the proposals made to translate into music the structural informations taken either from DNA or from protein molecules. We will examine these proposals which were not all made for musical purposes, but also as a simple memorizing technique.

The first attempt appears to be that of Hayashi and Munakata [7]. These authors start from the well-known Watson-Crick structure of DNA, with its succession of pairs of bases linking the two chains of the double helix. The four bases adenine, guanine, cytosine and thymine, symbolized by the letters A, G, C and T are always associated in the same way, A with T and G with C. This is shown in Figure 5. DNA sequencing is producing long sequences such as CAGCCTGACTG... One sequence is enough to tell us what is the sequence of the adjoining chain, because of the associative law. In order to "minimize the distress of handling such information", these authors propose to associate the note D (ré) with G, E (mi) with C, G (sol) with T and A (la) for A. The experience of the authors is that the stretch that they could memorize increased at least threefold. They do not pretend that this music has aesthetical value.

The following proposal for musical transcription is that of Ohno and Jabara [8]. They adopt a very simple rule shown in Figure 6. Each base is assigned two successive values of the musical scale. This is not enough to produce a musical score. The key, the mode and the rhythm are suggested by comparison with well-known pieces of classical music. Figure 7 gives an example of a transcription.

We come now to the most ambitious enterprise, that of J. Sternheimer [9]. The material, here, is the structure of proteins. These molecules are synthesized in a way which reflects the structure of DNA, after the double helix has split into two chains. They are constituted of chains of amino acids, twenty in total. An example of such a chain is given in Figure 8.

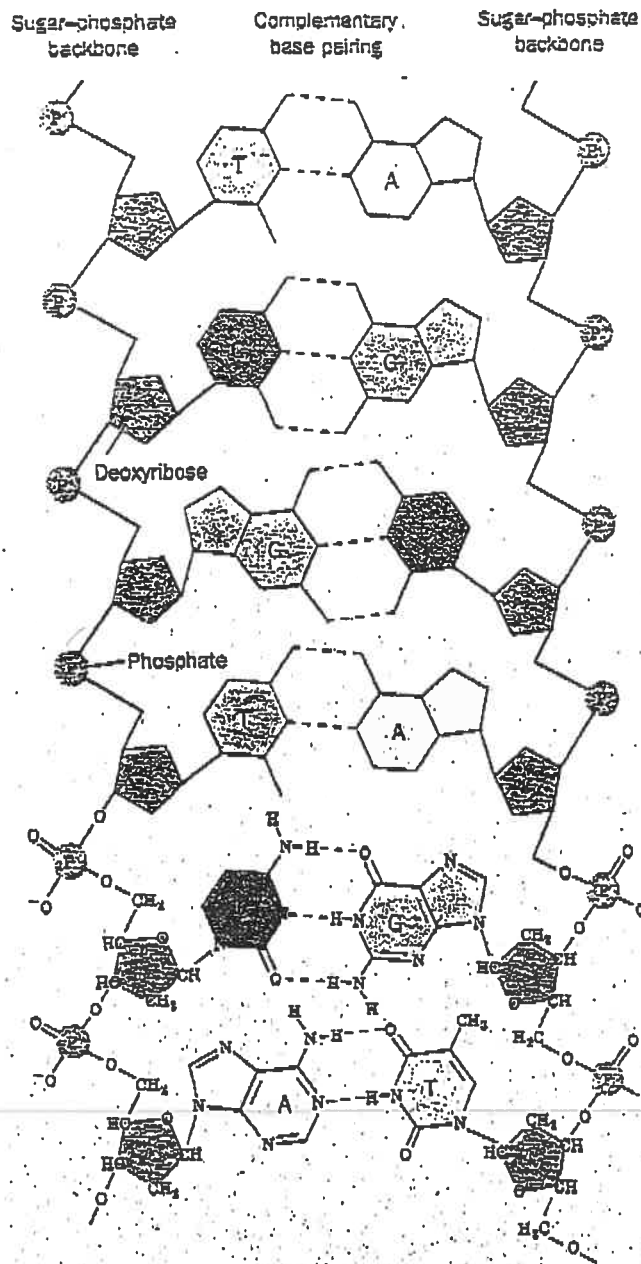


Figure 5: *The structure of the DNA double helix, with association of complementary base pairs A-T and G-C (after D. Voet and J. G. Voet, "Biochemistry", Wiley, New York 1996, with permission).*



Figure 6: *The notes assigned to the four bases of DNA by Ohno and Jabara [8].*

ANTI- $\mu\text{P}^{\beta}$  Ig VH VERSION II (A)

LEGATO ♩ = 100

mf

-2 HIS SER GLN 1 VAL GLN LEU GLN GLN PRO  
CAC TCC CAG G TCC AAC TGC A CAG CCT G

p

GLY 10 ALA GLU LEU VAL LYS PRO  
G G G C T GA G C T T G T GA A G C C

Figure 7: The beginning of a musical composition—extracted from a coding sequence by Ohno and Jabara [8]. Each group of three bases makes up a "codon" which is unambiguously associated to an amino acid abbreviated as shown in Figure 9.

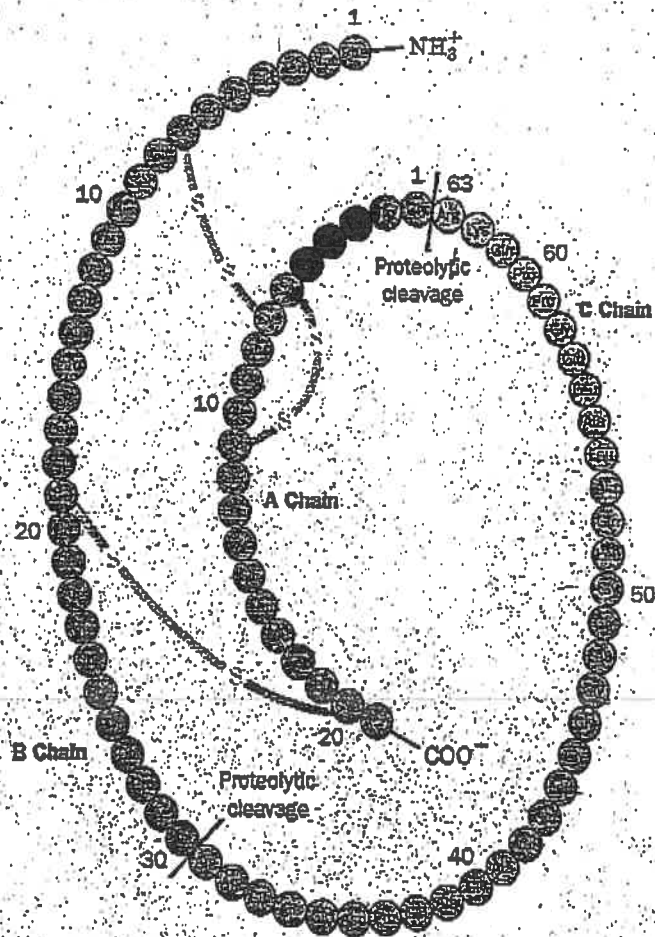


Figure 8: An example of the primary structure of a protein. The notation for the amino acids is given in Figure 9 (after R. E. Chance, R. M. Ellis and W. W. Brommer, *Science* 161,165 (1968), with permission).

Group I Amino Acids with Apolar R Groups	Group II Amino Acids with Uncharged Polar R Groups	Group III Amino Acids with Charged R Groups
Alanine Ala A $M_r$ 89 <sup>a</sup> $\begin{array}{c} \text{H} \\   \\ \text{CH}_3 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	Glycine Gly G $M_r$ 75 $\begin{array}{c} \text{H} \\   \\ \text{H} - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	Aspartic acid Asp D $M_r$ 133 $\begin{array}{c} \text{O} \\ // \\ \text{C} - \text{CH}_2 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Valine Val V $M_r$ 117 $\begin{array}{c} \text{CH}_3 \\   \\ \text{CH} - \text{C} - \text{COO}^- \\   \quad   \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$	Serine Ser S $M_r$ 105 $\begin{array}{c} \text{H} \\   \\ \text{HO} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	Glutamic acid Glu E $M_r$ 147 $\begin{array}{c} \text{O} \\ // \\ \text{C} - \text{CH}_2 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Leucine Leu L $M_r$ 131 $\begin{array}{c} \text{CH}_3 \\   \\ \text{CH} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \quad   \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$	Threonine Thr T $M_r$ 119 $\begin{array}{c} \text{H} \\   \\ \text{CH}_3 - \text{CH} - \text{C} - \text{COO}^- \\   \quad   \\ \text{OH} \quad \text{NH}_3^+ \end{array}$	Lysine Lys K $M_r$ 146 $\text{H}_2\text{N}^+ - (\text{CH}_2)_4 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+$
Isoleucine Ile I $M_r$ 131 $\begin{array}{c} \text{H} \\   \\ \text{CH}_3 - \text{CH}_2 - \text{CH} - \text{C} - \text{COO}^- \\   \quad   \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$	Cysteine Cys C $M_r$ 121 $\begin{array}{c} \text{H} \\   \\ \text{HS} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	Arginine Arg R $M_r$ 174 $\text{H}_2\text{N} - \text{C} = \text{NH} - (\text{CH}_2)_3 - \text{C} - \text{COO}^- \\    \quad   \\ \text{NH}_2 \quad \text{NH}_3^+$
Proline Pro P $M_r$ 115 $\begin{array}{c} \text{H}_2 \\   \\ \text{H}_2\text{C} - \text{C} \\   \quad   \\ \text{H}_2\text{C} \quad \text{C} - \text{COO}^- \\   \quad   \\ \text{H} \quad \text{N} - \text{H} \end{array}$	Tyrosine Tyr Y $M_r$ 181 $\text{HO} - \text{C}_6\text{H}_4 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+$	Histidine (at pH 6.0) His H $M_r$ 155 $\begin{array}{c} \text{H} \\   \\ \text{HC} = \text{C} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \quad   \\ \text{HN} \quad \text{NH} \\   \\ \text{H} \end{array}$
Phenylalanine Phe F $M_r$ 165 $\text{C}_6\text{H}_5 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+$	Asparagine Asn N $M_r$ 132 $\begin{array}{c} \text{NH}_2 \\   \\ \text{C} = \text{O} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	
Tryptophan Trp W $M_r$ 204 $\text{C}_8\text{H}_7\text{N} - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+$	Glutamine Gln Q $M_r$ 146 $\begin{array}{c} \text{NH}_2 \\   \\ \text{C} = \text{O} - \text{CH}_2 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$	
Methionine Met M $M_r$ 149 $\text{CH}_3 - \text{S} - \text{CH}_2 - \text{CH}_2 - \text{C} - \text{COO}^- \\   \\ \text{NH}_3^+$		

Figure 9: The twenty amino acids found in the structure of proteins, with their chemical denominations and two abbreviations used in the text. The molecular weights are also given (after D. Voet and J. G. Voet, "Biochemistry", Wiley, New York 1996, with permission).

## Cytochrome oxydase

Dernière protéine  
de la chaîne respiratoire humaine

Début de la sous-unité 3 mitochondriale

M T H Q S H A Y H M V K P S P W P L T G A L S A L L

M T S G L A M W F H F H S M I L L M L G L L T N I L T M Y Q W W

R D V T R E S T Y Q G H H T P P V Q K G

Figure 10: Transcription made by J. Sternheimer [9] of a protein belonging to the human respiratory chain. According to the author, music not only reflects the inner structure of proteins, but can also act on protein synthesis at the cellular level. Therefore it should not be played without care or expert assistance.

Figure 9 gives the chemical formulas of these twenty molecules, with their names and some common abbreviations to designate them. An important information is the molecular weight, also given in Figure 9.

A feature of the molecular weights of amino acids was noticed by Sternheimer: they cluster into groups. Frequency and mass can be associated through de Broglie's relation:  $\lambda = \frac{h}{mv}$ , where  $\lambda$  is the wavelength of the wave associated to a particle, while  $m$  is its mass and  $v$  its velocity. The frequency results from the relation  $\nu = \frac{c}{\lambda}$ , where  $c$  is the velocity of light. An estimate of the characteristic frequencies associated to amino acids shows that they are far above the audible range. This is a circumstance which was met before in the transcription of photoelectron spectra. The remedy is rescaling, here by 76 octaves! Due to the clustering of masses, only ten notes are retained by Sternheimer. Adopting the shortest notation of Figure 9 for each amino acid, we have the following correspondence (mw stands for molecular weight):

G (mw: 75) A2 (la2)  
 A (mw: 89) C3 (do3)  
 S (mw: 105) E3 (mi3)  
 P, V, T, C (mw: 115, 117, 119, 121) F3 (fa3)  
 L, I, N, D (mw: 131, 131, 132, 133) G3 (sol3)  
 Q, E, K, M (mw: 146, 146, 147, 149) A3 (la3)  
 H (mw: 155) B3 flat (si3 flat)  
 F (mw: 165) B3 (si3)  
 R, Y (mw: 174, 181) C4 (do4)  
 W (mw: 204) D4 (ré4)

There are some additional rules to determine the key, the mode and the rhythm. Figure 10 gives an example of transcription of a protein. We refer to the book of Fukagawa [6] for a description of the alleged effects of "protein music" on plant growing. This subject is outside the present review which only pretends to outline the technical procedures which are at work in musical transcription from scientific data.

#### IV. Conclusions

We have reviewed several schemes proposed to produce music from the spectral or structural information of molecules. It is clear that this is possible because such data are, in most cases, not distributed at random. There are some hidden regularities which are amplified by a musical transcription. This activity can be considered as

taking place in a wide area called "pattern recognition". It is sometimes easier for us to recognize a pattern from an auditive transcription, than from a visual one.

Another recent development [10] linking biomolecules to music is worth mentioning. This is the possibility for a computer to identify a song or a piece of music just from listening to a few notes: an algorithm translates the notes according to pitch and duration and compares the information to a bank of data kept in memory. This algorithm is very similar to that used to identify a DNA structure starting from the knowledge of a particular sequence of bases.

### Acknowledgments

I thank Sydney Leach for a discussion about his method of building a "muispectrum". Mrs Marja Rantanen-Maruani has kindly written a variation on the musical composition given by G. G. Englert in reference [3].

### References

- [1] J. G. Roederer, *The Physics and Psychophysics of Music* (Springer-Verlag, New York, 1995).
- [2] H. Helmholtz, *On the Sensation of Tone* Trad. E. J. Ellis (Dover, New York, 1954).
- [3] S. L. Leach and G. G. Englert, *J. Electr. Spectr. and Rel. Phen.*, **41**, 181 (1986).
- [4] E. Heilbronner and A. Schmelzer, *Nouv. J. Chim.*, **4**, 23 (1980).
- [5] W. von Niessen, G. H. F. Dierksen and L. S. Cederbaum, *Chem. Phys.*, **10**, 345 (1975).
- [6] Y. Fukagawa, "Protein Music", (Ed. Y. Fukagawa, Tokyo, 1999).
- [7] K. Hayashi and N. Munakoto, *Nature*, **310**, 96 (1984).
- [8] S. Ohno and M. Jabara, *Chem. Scripta*, **26B**, 43 (1986); see also A.-L. Lecocq and A. Danchin, *La Recherche*, **17**, 977 (1986).
- [9] J. Sternheimer, Brevet d'Invention n<sup>o</sup>9206765, INPI, Paris (1992). See also the article by Andy Coghlan in *New Scientist*, **142**, 10 (1994).
- [10] Interview of Professor J.-G. Ganascia in *Le Monde*, January 26 (2000).