YUCOMAT 2009, MRS Serbia, Herceg Novi, Montenegro, 2009.

QUANTUM FOUNDATIONS OF RESONANT RECOGNITION MODEL

G. Keković[1], D. Raković[1], B. Tošić[2], D. Davidović[3,4], I. Cosic[4]

[1]Faculty of Electrical Engineering, Belgrade, Serbia
[2]Vojvodina Academy of Sciences and Arts, Novi Sad, Serbia
[3]Vinca Institute of Nuclear Sciences, Belgrade, Serbia
[4]School of Electrical and Computer Engineering, RMIT, Melbourne, Australia

ABSTRACT. Biomolecular recognition is open scientific problem, which has been investigated in many theoretical and experimental aspects. In that sense, there are encouraging results within Resonant Recognition Model (RRM), which is based on theory of information spectrum of macro-molecules. The RRM concept is based on the finding that there is a significant correlation between spectra of the numerical presentation of amino acid and their biological activity. It has been found through an extensive research that proteins with the same biological function have a common frequency in their numerical spectra. This frequency was found then to be a characteristic feature for protein biological function or interaction The RRM model proposes that the selectivity of protein interactions is based on resonant energy transfer between interacting biomolecules and that this energy, electromagnetic in its nature, is in the frequency range of 10^{13} to 10^{15} Hz, which incorporates infra-red (IR), visible and a small portion of the ultra-violet (UV) radiation. In the paper, the quantum mechanical basis of the RRM model will be investigated using the solution in the simplified framework of Hückel-like theory of molecular orbits.

Keywords. ***Resonant Recognition Model; Hückel Theory; Isomeric Transitions; Macromolecules.***

INTRODUCTION

Biological processes in living organisms are based on *selective interactions* between biomolecules. These interactions are very specific and selective. This specificity is driven by the proteins, but it is still a *puzzle* where and how this specificity is written in the *protein structure*.

Currently accepted explanation is that the *specificity* of protein interactions is written in the protein *3-D structure* and it is based on *"key-and-lock" fit* between 3-D structure of the *protein active side* and *interactive target.* However, this fit in most cases is *very "loose",* and it is difficult to believe that this is the solely important parameter for the extremely selective and specific recognitions/interactions between biomolecules.

The ***RRM model*** is based on representation of the protein *primary structure* as a *discrete signals* by assigning to each *amino acid*, the electron *excitation energy* E_m [1-4], which is calculated as a *electron-ion interaction pseudo-potential* (*EEIP values*). Consequently, these *numerical series* are converted into *Fourier spectrum* by using Discrete Fourier transform (DFT).

The *coefficients* in the discrete Fourier transform are defined as:

$$E(n) = \sum_m E_m e^{-i\frac{2\pi mn}{N}} \quad ; \quad n = 1,2,.... \frac{N}{2} \tag{1}$$

where *N* is the number of amino acids in a given sequence and *n* is the *n*-th member of the original numerical series.

In practice, the *energy density spectrum* is defined as:

$$S(n) = E(n)E^*(n); \qquad n = 1,2,.... \frac{N}{2} \tag{2}$$

which is very convenient for analysis of information contained in macro-molecules and which is called *information spectrum*.

The above result can be *generalized* in case when there exists a *number of proteins*:

$$M(n) = \prod_{i=1}^{m} S_i(n); \qquad n = 1,2,..... \frac{N}{2} \tag{3}$$

where in the *energy density spectrum* appears ***peak*** which describes the *same or similar biological function* of these *macromolecules*.

RESONANT RECOGNITION MODEL (RRM)

All *proteins* can be considered as a *linear sequence* of their constitutive elements, i.e. *amino acids*. The RRM model interprets this ***linear information*** using signal analysis methods by transforming protein into a *numerical series* and then into the *frequency domain* using the *Discrete Fourier Transform* (DFT) [1,5]. The RRM is based on the representation of the protein primary structure as a numerical series by *assigning* to each amino acid a *physical parameter value* relevant to the *protein's biological activity*.

Although a number of amino acid indices have been found to correlate in some ways with the biological activity of the whole protein, our investigations [6,7] have shown that the *best correlation* can be achieved with parameters which are related to the energy of *delocalised electrons* of each amino acid. These findings can be explained by the fact that the electrons delocalised, from the particular amino acid, have the *strongest impact* on the electronic distribution of the *whole protein*.

In our extended studies, the energy of delocalised electrons (calculated as the ***electron-ion interaction pseudopotential*** (EIIP) [1,5-7]) of each amino acid residue was used. The resulting numerical series then represents the distribution of the free electrons energies along the protein.

Numerical series obtained this way are then analysed by *digital signal analysis methods* including Fourier Transform and Wavelet Transform in order to extract information pertinent to the biological function. As the average distance between amino acid residues in a polypeptide chain is about 3.8 Å, it can be assumed that the points in the numerical sequence derived are equidistant. For further numerical analysis the distance between points in these numerical sequences is set at an arbitrary value $d = 1$. Then the maximum frequency in the spectrum is $F = 1/2d = 0.5$. The total number of points in the sequence influences the resolution of the spectrum only. Thus for N-point sequence the resolution in the spectrum is equal to $1/N$. The n-th point in the spectral function corresponds to the frequency $f = n/N$.

In order to extract common spectral characteristics of sequences having the same or similar biological function, the *cross-spectral function* was used. Peak frequencies in the amplitude cross-spectral function define common frequency components of the two sequences analysed. To determine the common frequency components for a group of protein sequences, we have calculated the absolute values of *multiple cross-spectral function coefficients* M, which are defined as follows:

$$|M_n| = |X_{1n}|.|X_{2n}|...|X_{Mn}| \qquad n = 1,2,...,N/2 \qquad (4)$$

Peak frequencies in such a multiple cross-spectral function denote common frequency components for all sequences analysed. *Signal-to-noise ratio* (*S*/*N*) for each peak is defined as a measure of similarity between sequences analysed. *S*/*N* is calculated as the *ratio* between *signal intensity* at the particular peak frequency and the *mean value* over the whole spectrum. The extensive experience gained from previous research [1-3,5,8-10] suggests that a *S*/*N* ratio of *at least 20* can be considered as *significant*.

The multiple cross-spectral function for a large group of sequences with the same biological function has been named *"consensus spectrum"*. The presence of a ***peak frequency*** with significant signal-to-noise ratio in a consensus spectrum implies that all of the analysed sequences within the group have one frequency component in common. This frequency is related to the ***biological function*** provided the ***following criteria*** are met:

1. *One peak* only exists for a group of protein sequences *sharing* the same *biological function*
2. *No significant peak* exists for *biologically unrelated* protein sequences
3. *Peak frequencies* are *different* for *different biological functions*.

In our previous extensive studies, the above criteria have been implemented and the following fundamental conclusion was drawn: Each specific biological function within the protein or DNA is characterised by one frequency. It has been shown in previous research that all protein sequences with the common biological function have common frequency component, which is a specific feature for the observed function/interaction [1,5,9]. This characteristic frequency is related to the protein biological function as it was found in our previous investigations [1-3,5-10]. Furthermore, it was shown that the ***proteins and their targets*** have the ***same characteristic frequency*** in common. Thus, it can be postulated that the *RRM frequencies characterise not only a general function* but also a ***recognition and interaction*** between the particular *protein and its target*.

Once the *characteristic frequency* for a particular *protein function/interaction* is identified, it is possible then to utilize the RRM approach to ***predict*** the ***amino acids and/or segments*** in the protein sequence, which *predominantly contribute to this frequency* and thus, to the ***observed function***, as well as to *design de novo peptides* having the *desired periodicities*. As was shown in our previous studies of FGF peptidic antagonists [1,11] and HIV envelope agonists [12,13] such *de novo* designed peptides express the desired biological function.

THE POSSIBLE QUANTUM BACKGROUND OF RRM

Protein amino acids sequences determine their structure and biological function. Thus investigation of this sequence is very important for understanding biomolecular recognition. In that sense, *Hückel-like theory* [14,15] of *molecular orbits* could be relevant and *simplified theoretical framework*.

The ***primary structure of proteins*** (Fig. 1) has the following shape:

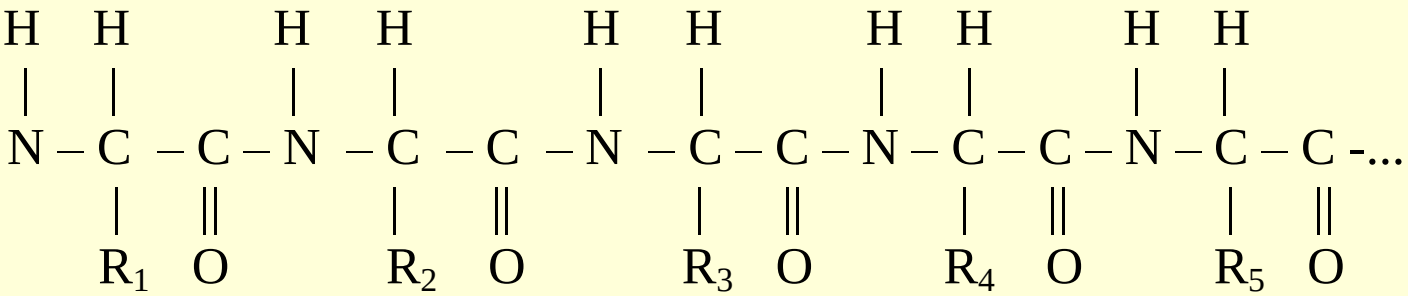

Fig. 1. Primary structure of a protein

where the terms R_1, R_2, … R_N identify *residues of amino acids* which carry *relevant information* about biological *function of protein* and C, N, H, O signify atoms of carbon, nitrogen, hydrogen and oxygen, respectively.

Within further procedure, we shall consider ***backbone of a protein*** (Fig. 2) without external (residual) attachments:

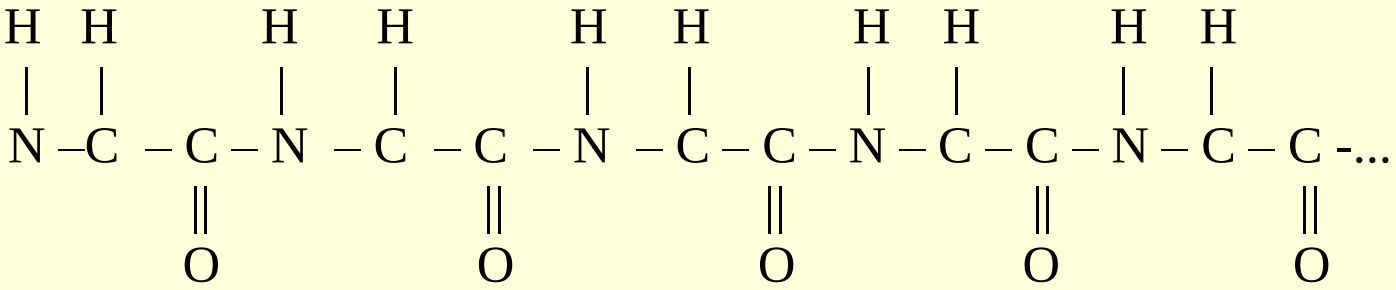

Fig. 2. Backbone of primary structure of a protein

We can simplify the problem if the structure motif – NH – CH – CO = consider as an *elementary cell*, described by *wavefunction* φ_m (as a *molecular orbit* (MO) in the form of *linear combination of atomic orbits* (LCAO) *centered on* *m*-th elementary cell). In this case ***wavefunction of molecular chain*** is:

$$\psi = \sum_m c_m \varphi_m ; \qquad (m = 1,2,....N) \tag{5}$$

By insertion of expression (5) into *Schrödinger equation* $H\psi = E\psi$ and by using linear variational method we get:

$$\sum_m c_m \int \varphi_n^* H \varphi_m dV = E \sum_m c_m \int \varphi_n^* \varphi_m dV; \quad (n = 1,2..., N)$$
$$\int \varphi_n^* \varphi_m dV = \delta_{nm}. \tag{6}$$

If we employ *approximation of the nearest neighbors* $H_{mn} = \int \varphi_n^* H \varphi_m dV \neq 0$, ($m = n \pm 1$), *cyclic boundary conditions* and *condition of weak overlapping* of wavefunctions of the nearest neighbors:

$$S_{nm} = \int \varphi_n^* \varphi_m dV = \begin{cases} 1; & n = m; \\ 0; & n \neq m \end{cases} \tag{7}$$

similarly to *Hückel's theory* we arrived to the following expression in *matrix form*:

$$\begin{bmatrix} H_0 - E & H_1 & 0 & \cdots & 0 \\ H_1 & H_0 - E & H_1 & 0 & 0 \cdots 0 \\ 0 & H_1 & H_0 - E & H_1 & .0 \\ . & & & & \\ . & & & & \\ H_1 & 0 & 0 & \cdots & H_1 \quad H_0 - E \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix} = 0 \quad , \qquad (8)$$

$$H_0 = \int \varphi_n^{*} H \varphi_n dV; \; H_1 = \int \varphi_n^{*} H \varphi_m dV$$

By inspection of the expression (8) it has been observed the *relationship*:

$$c_{m-1} H_1 + c_m (H_0 - E) + c_{m+1} H_1 = 0; \qquad (9)$$

which can be *solved* if we choose the *coefficients* $c_m^{(k)} = e^{i\frac{2\pi mk}{N}}$, where the index $m = 1,2...,N$ denotes *number of cells* and $k = 0,\pm 1,\pm 2.....,\pm N/2$ signify *molecular orbits* (bonding and anti-bonding).

After simple algebraic manipulations in Eq. (9) there appears expression for (single electron) *energy of molecular orbits*:

$$E(k) = H_0 + 2H_1 \cos\frac{k\pi}{l};$$

$$k = 0,\pm1,\pm2,.....\pm\frac{N}{2}; \quad l = \frac{N}{2}; \tag{10}$$

and corresponding *wavefunctions*:

$$\psi^{(k)} = \sum_m c_m{}^{(k)} \varphi_m \ . \tag{11}$$

The wavefunctions specified above allow us to calculate the *correction of energy* induced by *interactions of aminoresidues with backbone* in the framework of the *perturbation theory* (the first order correction) *via* formula:

$$\Delta E^{(k)} = \left\langle \psi^{(k)} \middle| H_{\text{int}}(k) \middle| \psi^{(k)} \right\rangle . \tag{12}$$

By ***specifying the potential of interaction*** of aminoresidues with backbone $H_{\text{int}}(k) = \frac{1}{N}\sum_m W_m$, $W_{m'}\left|\varphi_m\right\rangle = E_m\left|\varphi_m\right\rangle\delta_{m'm}$ (where for the *interaction energy* of *m*-th aminoresidues R_m and *m*-th elementary cells we can take the *electron excitation energy* E_m of the *m*-th amino acid, which is specifically determined by *local electronic structure of the aminoresidues* R_m [16], with corresponding *eigenfunctions* φ_m as MO LCAO *centered on m-th elementary cell*, while $1/N$ is normalizing factor for (single electron) first order energy corrections), it follows:

$$\Delta E^{(k)} = \frac{1}{N}\sum_{mnm'} c_m^{(k)} c_n^{(k)*} e^{i\frac{2\pi m'k}{N}} E_{m'}\left\langle\varphi_n\left|\varphi_m\right\rangle\delta_{m'm} = \frac{1}{N}\sum_m E_m \,. \qquad (13)$$

By analyzing formula given above, it can be seen that *DFT* (1) of the *sequential contributions to the first order correction of energy* (13), related to primary sequence of aminoresidues, is connected with *energy density spectrum* according to the formula (2).

DISCUSSION AND CONCLUSION

The important consequence of our model of quantum foundation of the RRM biomolecular recognition, based on simplified Hückel-like theory of molecular orbits, is that *energy density spectrum* (2) is the product of Discrete Fourier Transform and complex conjugate of the *sequential contributions to the first order correction of energy* of macromolecule *within the perturbation calculations* (13), confirming the RRM findings that ***primary sequence of aminoresidues is essential*** for bioinformation coding in proteins.

This fully *approves RRM possibility* ***to predict*** the *amino acids and/or segments* in the protein sequence, which *predominantly contribute* to the *observed function*, as well as ***to design*** theoretically *de novo* peptides having the *desired biological function*.

Also, the fact that *condition* $S_1(k) = S_2(q)$ for $k = q$ is *fulfilled* for *macromolecules* with *same or similar biological function*, as well as for *proteins and targets* involved in *key-lock biomolecular recognitions*, suggests *appearance* of corresponding macromolecular *resonant non-radiative isomeric transitions* (1→2): $E_e^{(1)} + \Delta E_{vib}^{(1)} = E_e^{(2)} + \Delta E_{vib}^{(2)}$ [16-18].

It should be noted that alongside with vibrational excitations (ΔE_{vib}), both *ground* and *excited* many-electron hypersurfaces (E_e) might be involved in these processes, as suggested by RRM theoretical predictions and experimental findings that the selectivity of protein interactions is based on *resonant energy transfer* between *interacting biomolecules* and that the *external excitation energy*, electromagnetic in its nature, is in the frequency range of 10^{13} to 10^{15} Hz, which incorporates ***infra-red*** (IR) radiation (when ***ground many-electron hypersurfaces*** and *their vibrational excitations* are involved), as well as ***visible*** and a small portion of the ***ultra-violet*** (UV) radiation (when ***excited many-electron hypersurfaces*** and *their vibrational excitations* are involved) [8].

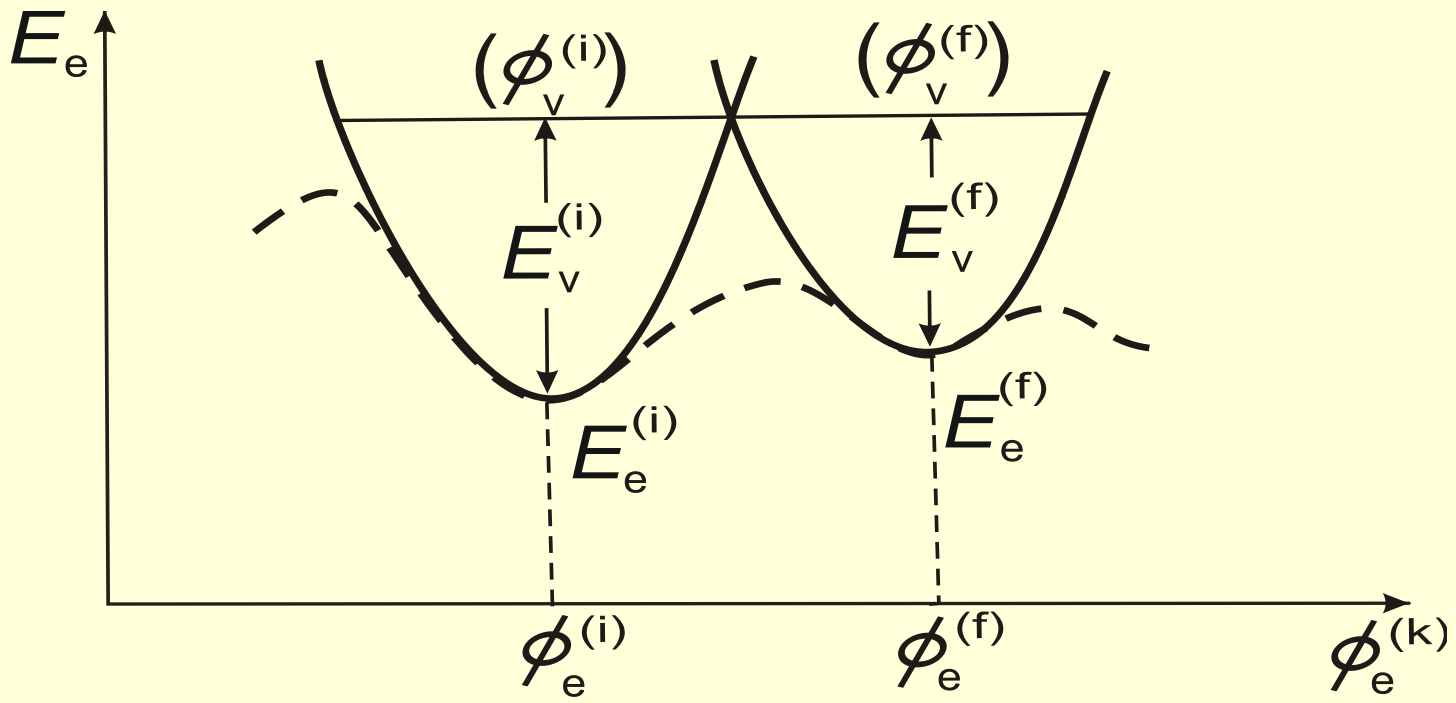

Figure 1. (Quasi)classical problem of *many-electron hypersurface* $E_e(\phi^{(k)})$ - not adiabatically well-defined when traversing between two adjacent local minima - is *replaced* by better defined problem of two (virtually intersecting) *isomeric many-electron hypersurfaces* (hyper-paraboloids) serving as potential hypersurfaces for two *vibrational (isomeric) problems*.

Finally, results of the RRM model imply that ***biomolecular information processing*** is going on in the *inverse space* of Fourier spectra of the *primary sequences of biomolecules*, bearing *resemblance to quantum-holographic ideas* that ***cognitive information processing*** is going on in the *inverse space* of Fourier spectra of the *perceptive stimuli* [19], *tentatively suggesting* possible ***quantum-holographic fractal coupling*** of various *hierarchical levels in biological species*, with potentially *significant psychosomatic implications* as well [20].

REFERENCES

1. I. Cosic, *The Resonant Recognition Model of Macromolecular Bioactivity*, Birkhauser, Berlin (1997).
2. E. Pirogova, I. Cosic, *Proc.* (417) *Biomed. Engin.*, ACTA Press (2004).
3. I. Cosic, E. Pirogova, V. Vojisavljević, Q. Fang, *Proc.* (555) *Biomed. Engin.*, ACTA Press (2007).
4. D. Raković, M. Dugić, M. Plavšić, *Mater. Sci. Forum*, 494, 513 (2005).
5. I. Cosic, *IEEE Trans. Biomed. Engin.*, 41, 1101 (1994).
6. E. Pirogova, G. P. Simon, I. Cosic, *IEEE Trans. NanoBiosci.*, 2(2), 63 (2003).
7. E. Pirogova, I. Cosic, *Molec. Simulat.*, 28(8-9), 845 (2002).
8. V. Vojisavljevic, E. Pirogova, I. Cosic, *Int. J. Rad. Biol.*, 83(4), 221 (2007).
9. Q. Fang, I. Cosic, *APESM,* 21, 179 (1998).
10. E. Pirogova, M. Akay, I. Cosic, *Proc. IEEE EMBS,* 21(2), 1215 (1999).
11. I. Cosic, A. E. Drummond, J. R. Underwood, M. T. W. Hearn, *Molec. Cell. Biochem.*, 130, 1 (1994).
12. V. Krsmanovic, J. M. Biquard, M. Sikorska-Walker, I. Cosic, C. Desgranges, M. A. Trabaud, J. F. Whitfield, J. P. Durkin, A. Achour, M. T. Hearn, *J. Peptide Res*., 52(5), 410 (1998).
13. B. Krsmanović, I. Cosić, J. M. Biquard, M. T. W. Hearn, Artifical peptides and induced antibodies, *Patent Appl. PCT/FR 93/00171*.
14. W. Kutzelnigg**,** *J. Comput. Chem.*, 28(1), 25 (2006).
15. J. P. Lowe, K. Peterson, *Quantum Chemistry*, Elsevier, Amsterdam (2005).
16. L. A. Gribov, *From Theory of Spectra Towards Theory of Chemical Transformations*, URSS, Moscow (2001), in Russian.
17. V. I. Baranov, L. A. Gribov, *Zh. Prikl. Spektr.*, 71(4), 421 (2004).
18. V. I. Baranov, F. A. Savin, L. A. Gribov, *J. Appl. Spectr.*, 71(3), 320 (2004).
19. K. Pribram, Brain and Perception: Holonomy and Structure in Figural Processing, Lawrence Erlbaum, Hillsdale, NJ (1991); K. Pribram, Languages of the Brain: Experimental Paradoxes and Principles in Neuropsychology, Brandon House, New York (1971).
20. D. Raković, M. Dugić, M. Plavšić, G. Keković, I. Cosic, D. Davidović, *Mater. Sci. Forum*, **5**18, 485 (2006).