Quantum Foundations of Resonant Recognition Model

G. Keković[a], D. Raković[b,*], B. Tošić[c], D. Davidović[c,d] and I. Cosić[d]

[a]Faculty of Electrical Engineering, University of Belgrade, P.O.B. 35-54, 11120 Serbia

[b]Vojvodina Academy of Sciences and Arts, Novi Sad, Serbia

[c]Vinca Institute of Nuclear Sciences, Belgrade, Serbia

[d]School of Electrical and Computer Engineering, RMIT, Melbourne, Australia

Biomolecular recognition is an 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), based on the finding that there is a significant correlation between spectra of the numerical presentation of amino acids in the primary structure of proteins 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 this 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.

1. Introduction

Biological processes in the living organisms are based on selective interactions between bio-molecules. 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 the 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 an electron-ion interaction pseudopotential (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:

$$E(k) = \sum_m E_m e^{-i\frac{2\pi mk}{N}}; \quad k = 1, 2, \ldots, \frac{N}{2}, \qquad (1)$$

where N is the number of amino acids in a given sequence and m is the m-th member of the original numerical series ($m = 1, 2, \ldots, N$).

The absolute values of DFT coefficients (1) of the protein sequence, $|E(k)|$ ($k = 1, 2, \ldots, N/2$), are very convenient for analysis of information contained in macromolecules, and so this energy density spectrum is called *information spectrum*.

2. 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 Fourier Transform (FFT) [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 the 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.

---

* corresponding author; e-mail: `rakovicd@etf.bg.ac.rs`

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 k-th point in the spectral function corresponds to the wavenumber k/N.

To determine the common spectral peak for a group of P protein sequences, we calculate the absolute values of multiple cross-spectral function coefficients, which are defined as follows:

$$|M(k)| = |E_1(k)||E_2(k)|\ldots|E_p(k)|$$

$$k = 1, 2, \ldots, N/2 \,. \tag{2}$$

where $|E_i(k)|$ are absolute values of DFT coefficients of the i-th sequence (protein), $i = 1, 2, \ldots, P$. Spectral peaks in such a multiple cross-spectral function denote common spectral peaks 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 spectral peak 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 spectral peak with significant signal-to-noise ratio in a consensus spectrum implies that all of the analysed sequences within the group have one spectral peak in common. This spectral peak (RRM wavenumber, and corresponding EM frequency correlated by semi-empirical dispersion relation [1]) is related to the biological function if 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.

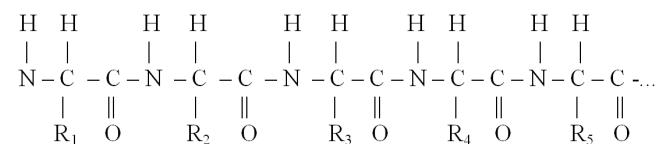3. Peak frequencies are different for different biological functions.

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.
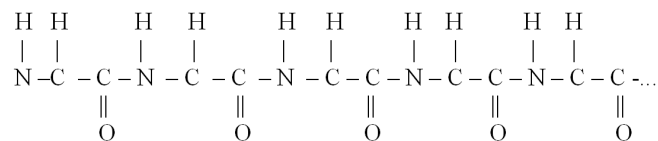
Once the characteristic frequency for a particular protein function/interaction is identified, it is possible to utilize the RRM approach to predict the amino acids and/or segments in the protein sequence, which predominantly contribute to this frequency and thus, to the observed function, as well as 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.

3. 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 the theoretical framework. The primary structure of proteins has the following shape:

```
H  H        H   H         H   H        H  H         H  H
|  |         |   |          |   |         |  |          |  |
N – C  – C – N – C  – C  – N – C – C – N – C  – C – N – C – C -...
|  ‖         |   ‖          |   ‖         |  ‖          |  ‖
R₁  O        R₂   O         R₃   O        R₄  O         R₅  O
```

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 without external (residual) attachments:

```
H  H       H   H         H   H        H  H         H  H
|  |        |   |          |   |         |  |          |  |
N –C  – C – N – C  – C  – N – C – C – N – C  – C – N – C – C -...
    ‖               ‖              ‖            ‖              ‖
    O               O              O            O              O
```

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:

$$\psi = \sum_m c_m \varphi_m; \quad (m = 1, 2, \ldots, N) \,. \tag{3}$$

By insertion of Eq. (3) into Schrödinger equation $H\psi = E\psi$ and by using linear variational method we get:

$$\sum_m c_m \int \varphi_n^* H \varphi_m \, dV = E \sum_m c_m \int \varphi_n^* \varphi_m \, dV \, ;$$
$$(n = 1, 2 \ldots, N) \tag{4}$$

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:

$$S_{nm} = \int \varphi_n^* \varphi_m \, dV = \begin{cases} 1 \, ; & n = m \, ; \\ 0 \, ; & n \neq m \end{cases} . \tag{5}$$

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

$$\begin{bmatrix} H_0 - E & H_1 & 0 \ldots\ldots\ldots\ldots H_1 \\ H_1 & H_0 - E & H_1 & 0 & 0\ldots\ldots 0 \\ 0 & & H_1 & H_0 - E & H_1\ldots 0 \\ \cdots \\ \cdots \\ H_1 & 0 & 0\ldots\ldots H_1 & H_0 - E \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_n \end{bmatrix} = 0 \,, \tag{6}$$

$$H_0 = \int \varphi_n^* H \varphi_n \, dV \, ; \quad H_1 = \int \varphi_n^* H \varphi_m \, dV \, .$$

By inspection of the Eq. (6) it has been observed the relationship:

$$c_{m-1} H_1 + c_m (H_0 - E) + c_{m+1} H_1 = 0 \,, \tag{7}$$

which can be solved if we choose the coefficients $c_m^{(k)} = e^{i \frac{2\pi m k}{N}}$, where the index $m = 1, 2 \ldots, N$ denotes number of cells and $k = 0, \pm 1, \pm 2 \ldots, \pm \frac{N}{2}$ signify molecular orbits (bonding and anti-bonding). After simple algebraic manipulations in Eq. (7) there appears expression for (single electron) energy of molecular orbits:

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

$$k = 0, \pm 1, \pm 2, \ldots, \pm \frac{N}{2} \, ; \quad l = \frac{N}{2} \, , \tag{8}$$

and corresponding wavefunctions:

$$\psi^{(k)} = \sum_m c_m^{(k)} \varphi_m \, . \tag{9}$$

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:

$$\Delta E^{(k)} = \left\langle \psi^{(k)} \left| H_{\text{int}}(k) \right| \psi^{(k)} \right\rangle \, . \tag{10}$$

By specifying the potential of interaction of aminoresidues with backbone $H_{\text{int}}(k) = \frac{1}{N} \sum_m W_m$ and $W_{m'} | \varphi_m \rangle = E_m | \varphi_m \rangle \delta_{m'm}$ (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 LCRO centered on m-th elementary cell, while $1/N$ is normalizing factor for (single electron) first order energy

corrections), it follows:

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

$$= \frac{1}{N} \sum_m E_m \, . \tag{11}$$

(for $k = 0, \pm 1, \pm 2, \ldots, \pm N/2$, see Eq. (8)).

4. 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 information spectrum is related to absolute values of DFT coefficients (1) of the sequential contributions to the first order correction of energy of macromolecule within the perturbation calculations (11), supporting 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.

The fact that there exists the RRM common peak k for all macromolecules with same or similar biological function, characterises the RRM selection rule for their particular function in the inverse space of Fourier spectra of the primary sequences of the biomolecules (compare also the ranges of k in Eqs. (1) and (8)!).

On the other hand, the proposal of the selectivity of protein-target EM interactions based on RRM resonant energy transfer between interacting biomolecules involved in key-lock biomolecular recognitions, suggests appearance of corresponding macromolecular *resonant non-radiative isomeric transitions* $(1 \rightarrow 2)$ of the protein-target complex, $E_e^{(1)}(k) + \Delta E_{\text{vib}}^{(1)}(k) = E_e^{(2)}(k) + \Delta E_{\text{vib}}^{(2)}(k)$ [16–18]. These transitions are induced by external EM excitation energy, RRM theoretically predicted and experimentally observed in the frequency range of 10^{13} to 10^{15} Hz [8] — which incorporates infrared (IR) radiation (when ground many-electron hypersurfaces (E_e) and their vibrational excitations (ΔE_{vib}) should be involved), as well as visible and a small portion of the ultra-violet (UV) radiation (when excited many-electron hypersurfaces and their vibrational excitations should be involved).

Finally, results of the RRM model imply that on the biomolecular level information process 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 significant potential psychosomatic implications as well [20].

Acknowledgments

References

[1] I. Cosic, *The Resonant Recognition Model of Macro-molecular Bioactivity*, Birkhauser, Berlin 1997.

[2] E. Pirogova, I. Cosic, in: *Proc. Biomed. Engin.*, ACTA Press 2004, p. 417.

[3] I. Cosic, E. Pirogova, V. Vojisavljević, Q. Fang, *Proc. Biomed. Engin.*, ACTA Press 2007, p. 555.

[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**, 63 (2003).

[7] E. Pirogova, I. Cosic, *Molec. Simulat.*, **28**, 845 (2002).

[8] V. Vojisavljevic, E. Pirogova, I. Cosic, *Int. J. Rad. Biol.* **83**, 221 (2007).

[9] Q. Fang, I. Cosic, *APESM* **21**, 179 (1998).

[10] E. Pirogova, M. Akay, I. Cosic, *Proc. IEEE EMBS* **21**, 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**, 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**, 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.

[17] V.I. Baranov, L.A. Gribov, *Zh. Prikl. Spektr.* **71**, 421 (2004).

[18] V.I. Baranov, F.A. Savin, L.A. Gribov, *J. Appl. Spectr.* **71**, 320 (2004).

[19] K. Pribram, Brain and Perception: *Holonomy and Structure in Figural Processing*, Eds. Lawrence Erlbaum, Hillsdale, New York 1991; K. Pribram in: *Languages of the Brain* Brandon House, New York 1971.

[20] D. Raković, M. Dugić, M. Plavšić, G. Keković, I. Cosic, D. Davidović, *Mater. Sci. Forum* **518**, 485 (2006).