Supplementary Materials: RRM Prediction of Erythrocyte Band3 Protein as Alternative Receptor for SARS-CoV-2 Virus

Irena Cosic, Drasko Cosic and Ivan Loncarevic

Resonant Recognition Model

Here, we present our own nonconventional, biophysical, theoretical Resonant Recognition Model (RRM), which is based on the findings that certain periodicities within the distribution of energy of delocalized electrons along protein molecules are critical for protein biological functions and/or interactions with their targets [1–12]. The RRM model has been extensively published and/or experimentally successfully tested [1–29]. The RRM model has been presented and well explained in our previous publications.

All proteins can be considered as a linear sequence of their constitutive elements, i.e., amino acids and the biological function of proteins are determined primarily by this linear sequence. The RRM [1–3] interprets this linear information by transforming the protein sequence into a numerical series and then into the frequency domain using the digital signal processing method Fast Fourier Transform (FFT).

The primary structure of proteins can be presented as a numerical series by assigning a relevant physical parameter value to each amino acid. Our investigations have shown that the best correlation can be achieved with parameters that are related to the energy of the delocalized electrons of each amino acid (calculated as Electron Ion Interaction Potential (EIIP)), as electrons delocalized from the particular amino acid have the strongest impact on the electronic distribution of the whole protein [1–4]. The resulting numerical series represents the distribution of the free electron energies along the protein molecule.

Such numerical series are then analyzed by digital signal analysis methods, using FFT, to extract information pertinent to the biological function. As the distance between amino acid residues in a polypeptide chain (protein backbone) is 3.8 Å, it can be assumed that the points in the numerical sequence are equidistant. For further numerical analysis, the distance between points in these numerical sequences is set at an arbitrary value of d = 1. Therefore, 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 an 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.

To extract the common spectral characteristics of sequences with the same or similar biological functions, the multiple cross-spectral function is used. Peak frequencies in such a multiple cross-spectral function present common frequency components for all sequences analyzed. Such common frequency components are found to be related to the common biological function of the analyzed proteins, leading to the conclusion that each specific biological function within the protein is characterized by one frequency [1–3,5,6].

Each biological function and/or process is driven by proteins that selectively interact with other proteins, DNA regulatory segments or small molecules. Using the RRM, it has been shown that proteins and their targets share the same matching characteristic frequency (periodicity) [1–3,5,6]. The matching of periodicities within the distribution of energies of free electrons along the interacting proteins can be regarded as the resonant recognition and is highly selective. Thus, the RRM frequencies characterize not only protein function, but also recognition and interaction between a protein and its targets: receptors, binding proteins and inhibitors. In addition, it has also been shown that interacting proteins have opposite phases at their characteristic recognition frequency [1–3,5–7]. Every frequency can be presented by one sinusoid characterized with its frequency, amplitude and phase. The phase is presented in radians and can be between $-\pi$ and $+\pi$ (−3.14 and +3.14). A phase

difference of or about 3.14 is considered an opposite phase. The phase value can be presented in the phase circle, where it is easier to observe graphically opposite phases.

It has been proposed that the RRM frequencies characterize not only a general function, but also a recognition and interaction between the macromolecule and its target, which then can be considered as a resonant recognition. This could be achieved with resonant energy transfer between the interacting macromolecules through oscillations of a physical field that is electromagnetic in nature. Since there is evidence that proteins have certain conducting or semi-conducting properties, a charge moving through the macromolecular backbone and passing different energy stages, caused by different amino acid or nucleotide side groups, can produce sufficient conditions for specific electromagnetic radiation or absorption. The frequency ranges of this field depend on the charge velocity. The RRM proposes, based on solid state physics principles, that the charge is travelling through the macromolecular backbone at an estimated velocity of 7.87×10^5 m/s [1–3,8,9]. For this velocity, and with the distance between amino acids in a protein backbone of 3.8Å, the frequency of protein interactions was estimated to be in the range between 10^{13}Hz and 10^{15}Hz. Therefore, the estimated frequency range for both amino acid and nucleotide macromolecules includes far infra-red, infra-red, visible and ultraviolet light. To support this idea, we compared our RRM computational predictions for variety of biological functions with a number of published experimental results [1–3]:

- Laser light growth promotion of cells, by using the frequencies of light to produce a similar effect to that of growth factor proteins;
- Chymotrypsin activation (increase of enzyme activity) achieved by laser light radiation in the range of 850–860nm;
- Activation of highly homologous plant photoreceptors, which, although being very homologous, absorb different wavelengths of light;
- Photoactivated proteins, e.g., rhodopsin, flavodoxin, etc.

These comparisons have empirically shown a strong linear correlation between frequencies, as calculated using the RRM method and experimentally measured characteristic frequencies, with a slope factor of K = 201 [1–3,10]. This finding parallels with the frequency range previously associated with the RRM numerical frequency spectrum that has been calculated from the charge velocities through the protein backbone. This correlation can be represented as follows:

$$\lambda = K / f_{rrm} \quad (1)$$

where λ is the wavelength of light irradiation in nm, which can influence particular biological process, f_{rrm} is RRM numerical frequency and K is coefficient of this linear correlation.

We applied this concept to a number of proteins and DNA examples [8–12]. The concept has also been experimentally tested by predicting the electromagnetic frequencies for the activation of L-Lactate Dehydrogenase [13]. In this experiment, by radiating L-Lactate Dehydrogenase with RRM-predicted electromagnetic frequencies, a significant change in enzyme activity was achieved. The concept has also been tested independently on experimental measurements of photon emission from dying melanoma cells [14], on photon emission from lethal and non-lethal Ebola strains [15], as well as on the classic signaling pathway JAK-STAT, traditionally composed of nine sequential protein interactions [16]. Moreover, the RRM model, for the first time, explains how and why external blue light can be used in the treatment of Crigler–Najjar syndrome [11].

Keeping all this in mind, we propose that the RRM concept is an excellent predictor of proteins' selective interactions, biological processes and pathways in living cells. In our previous work, we have calculated large number of specific frequencies for different protein and DNA biological functions and interactions [1–13,17–24].

Bioactive Peptide Design

Once the characteristic biological function of the protein is identified, it is possible to design new proteins with the desired frequency components and, consequently, with the desired biological functions [1–3,7,18–24]. The process of bioactive peptides design is as follows:

- Determination of RRM characteristic frequency using multiple cross-spectral functions for a group of protein sequences that share common biological function (interaction);
- Determination of phases for the characteristic frequencies of a particular protein, which is selected as the parent for the agonist/antagonist peptide;
- Calculation using an Inverse Fourier Transform of the signal with characteristic frequency and phase. The minimal length of the designed peptide is defined by the characteristic frequency f as 1/f;
- Determination of the resulting amino acid sequence using tabulated EIIP parameter values.

This approach has already been successfully applied and experimentally tested in the design of FGF [1–3,19], HIV envelope protein analogue [1–3,7,20,21], mice IL-12 analogue [17,22] and a peptide to mimic the oncolytic function of the myxoma virus [23,24].

It is interesting to note that de novo designed peptides do not have any significant homology with the original protein, but, when their 3D structure is predicted, it has been shown that this 3D structure is very similar to the 3D structure of the original active protein site [19]. As the 3D structure is defined by the protein's primary structure, it seems that RRM, by analyzing frequencies within the primary structure, can decipher the rules of how the 3D structure is formed.

Prediction of the Key Amino Acids—"Hot Spots"

Knowing the characteristic frequency of particular protein functions creates the possibility to predict which amino acids prevail in the sequence and predominantly contribute to this frequency and, consequently, to the observed function. This could be achieved by small alternations in the amplitude of the single protein spectrum at the characteristic frequency and an observation of which amino acids are most sensitive to this alternation [1–3,25–28]. These sensitive amino acids ("hot spots") are related to the characteristic frequency and, consequently, to the corresponding biological function. The "hot spots" predictions, using the RRM, have already been applied to a number of protein and DNA examples, including interleukin-2, SV40 enhancer, epidermal growth factor (EGF), Ha-ras p21 oncogene product, glucagons, hemoglobins, myoglobins and lysozymes [1–3,25–28]. It has been experimentally documented via the example of influenza virus that such predicted amino acids denote residues crucial for protein function [29].

In addition, these amino acid "hot spots", although not sequentially linked, are found to be spatially clustered in the protein tertiary structure and to be positioned in and around the protein active site [25–28]. This could mean that the 3D protein structure is formed as a resonant box for the characteristic frequency.

References

1. Cosic, I. Macromolecular Bioactivity: Is it Resonant Interaction between Macromolecules? Theory and Applications. *IEEE Trans. Biomed. Eng*. **1994**, *41*, 1101–1114.
2. Cosic, I. Virtual spectroscopy for fun and profit. *Biotechnology*, **1995**, *13*, 236–238
3. Cosic, I. In The Resonant Recognition Model of Macromolecular Bioactivity: Theory and Applications. Birkhäuser Verlag: Basel, Switzerland, 2012.
4. Pirogova, E.; Cosic, I. Examination of amino acid indexes within the Resonant Recognition Model. In Proceedings of the *Conference of the Victorian Chapter of the IEEE EMBS*. IEEE: Melbourne, Australia, 19–20 February 2001, pp. 1–4.
5. Cosic, I.; Cosic, D.; Lazar, K. Analysis of Tumor Necrosis Factor Function Using the Resonant Recognition Model. *Cell Biochem. Biophys.* **2015**, *11*, 175-180.
6. Cosic, I.; Paspaliaris, V, Cosic, D: Analysis of Protein-Receptor on an Example of Leptin-Leptin Receptor Interaction Using the Resonant Recognition Model. *Appl. Sci.*, **2019**, *9*, 5169.
7. Krsmanovic, V.; Biquard, J.M.; Sikorska-Walker, M.; Cosic, I.; Desgranges, C.; Trabaud, M.A.; Whitfield, J.F.; Durkin, J.P.; Achour, A.; Hearn M.T. Investigation into the Cross-reactivity of Rabbit Antibodies Raised against Nonhomologous Pairs of Synthetic Peptides Derived from HIV-1 gp120 proteins. *J. Peptide. Res.* **1998**, *52*, 410–412.
8. Cosic, I.; Lazar, K.; Cosic, D. Prediction of Tubulin resonant frequencies using the Resonant Recognition Model (RRM). *IEEE Trans. NanoBioscience* **2015**, *12*, 491–496.

9. Cosic, I.; Cosic, D.; Lazar, K. Is it possible to predict electromagnetic resonances in proteins, DNA and RNA? *Nonlinear Biomedical Physics*, **2015**, *3*, 5.
10. Cosic, I.; Cosic, D.; Lazar, K. Environmental Light and Its Relationship with Electromagnetic Resonances of Biomolecular Interactions, as Predicted by the Resonant Recognition Model. *Int. J. Environ. Res. Public. Health* **2016**, *13*, 647.
11. Cosic, I.; Cosic, D. The Treatment of Crigler-Najjar Syndrome by Blue Light as Explained by Resonant Recognition Model. *EPJ Nonlinear Biomed. Physics.* **2016**, 4, 9.
12. Cosic, I.; Paspaliaris, V.; Cosic, D. Explanation of Osteoblastic Differentiation of Stem Cells by Photo Biomodulation Using the Resonant Recognition Model. *Appl. Sci*, **2019**, *9*, 1979.
13. Vojisavljevic, V.; Pirogova, E.; Cosic, I. The Effect of Electromagnetic Radiation (550nm-850nm) on I-Lactate Dehydrogenase Kinetics. *Intern. J. Radiat. Biol*. **2007**, *83*, 221–230.
14. Dotta, B.T.; Murugan, N.J.; Karbowski, L.M.; Lafrenie, R.M.; Persinger, M.A. Shifting wavelength of ultraweak photon emissions from dying melanoma cells: their chemical enhancement and blocking are predicted by Cosic's theory of resonant recognition model for macromolecules. *Naturwissenschaften* **2014**, *101*, 87–94.
15. Murugan, N.J.; Karbowski, L.M.; Persinger, M.A. Cosic's Resonance Recognition Model for Protein Sequences and Photon Emission Differentiates Lethal and Non-Lethal Ebola Strains: Implications for Treatment. *Open J. Biophys.* **2014**, *5*, 35.
16. Karbowski, L.M.; Murugan, N.J.; Persinger, M.A. Novel Cosic resonance (standing wave) solutions for components of the JAK-STAT cellular signalling pathway: A convergence of spectral density profiles. *FEBS Open Bio.* **2015**, *5*, 245–250.
17. Pirogova, E.; Istivan, T.; Gan, E.; Cosic, I. Advances in methods for therapeutic peptide discovery, design and development. Curr. Pharm. Biotechnol. **2011**, *12*, 1117–1127.
18. Cosic, I.; Pirogova, E. Bioactive Peptide Design using the Resonant Recognition Model. *Nonlinear Biomedical Physics* 2007, 1(7), doi: 10.1186/1753-4631-1-7.
19. Cosic, I.; Drummond, A.E.; Underwood, J.R.; Hearn, M.T.W. In vitro inhibition of the actions of basic FGF by novel 16 amino acid peptides. *Molecular and Cellular Biochemistry*, **1994**, *130*, 1–9.
20. Krsmanovic, V.; Cosic, I.; Biquard, J.M.; Hearn, M.T.W. Peptides immunologically related to known viral protein. U.S. Patent No. 6,294,174, 25 Sep. 2001.
21. Achour, A.; Biquard, J.M.; Krsmanovic, V.; M'Bika, J.P.; Ficheux, D.; Sikorska, M.; Cozzone, A.J. Induction of Human Immunodeficiency Virus (HIV-1) Envelope Specific Cell-Mediated Immunity by a Non-Homologus Synthetic Peptide. *PLoS one*, **2007**, *11*, 1–12.
22. Pirogova, E.; Istivan, T.; Gan, E.; Cosic, I. Advances in Methods for Therapeutic Peptide Discovery, Design and Development. *Curr. Pharm. Biotechnol.* **2011**, *12*, 1117–1127.
23. Almansour, N.; Pirogova, E.; Coloe, P.; Cosic, I.; Istivan, T. Investigation of cytotoxicity of negative control peptides versus bioactive peptides on skin cancer and normal cells: a comparative study. *Future Med. Chem.* **2012**, *4*, 1553–1565.
24. Istivan, T.; Pirogova, E.; Gan, E.; Almansour, N.; Coloe, P.; Cosic, I. Biological effects of a De Novo designed myxoma virus peptide analogue: Evaluation of cytotoxicity on tumor cells. *PLoS one*, **2011**, *6*, 1–10.
25. Cosic, I.; Hearn, M.T.W. "Hot Spot" Amino Acid Distribution in Ha-ras Oncogene Product p21: Relationship to Guanine Binding Site. *J. Mol. Recognit.* **1991**, *4*, 57–62.
26. Cosic, I.; Hearn, M.T.W. Studies on Protein-DNAInteractions Using the Resonant Recognition Model: Application to Repressors Transforming Proteins Eur. *J. Biochem*. **1992**, *205*, 613–619.
27. Cosic, I.; Hodder, A.N.; Aguilar, M.I.; Hearn, M.T.W. Resonant Recognition Model protein topography: model studies with myoglobin haemoglobin lysozyme Eur. *J. Biochem*. **1991**, *198*, 113–119.
28. Caceres, J.L.H.; Cosic, I.; Cosic, D. Application of the Resonant Recognition Model to the Study of Plasmodium Proteins Involved in Malaria Infection. *MD Med Data* **2015**, *7*, 7–14.
29. Schmier, S.; Mostafa, A.; Haarmann, T.; Bannert, N.; Ziebuhr, J.; Veljkovic, V.; Dietrich, U.; Pleschka, S. In silico prediction and experimental confirmation of HA residues conferring enhanced human receptor specificity of H5N1 influenza a viruses. *Sci. Rep.* **2015**, *5*, 11434.