medicinska revija
medical review

UDK: 615.357.015.3:577.175.72 ;
616.379-008.64-085.357.015.3
ID BROJ: 205589260

Cosic I. and Cosic D. MD-Medical Data 2014;6(1): 007-010

Originalni članci/
Original articles

BALANCING INSULIN STABILITY AND ACTIVITY – RATIONAL APPROACH*

RACIONALAN PRISTUP BALANSIRANJU INSULINSKE STABILNOSTI I AKTIVNOSTI*

Irena Cosic[1,2], Drasko Cosic[2]

[1] RMIT University, Melbourne, Victoria, Australia
[2] AMALNA Consulting, Black Rock, Victoria, Australia

Correspondence to:

Irena Cosic,
Emeritus Professor, RMIT University
Director AMALNA Consulting,
46 Second Street, Black Rock, 3193
Victoria, Australia
e-mail: irenacosic@me.com

*Invited paper/ Rad po pozivu

Key words
Insulin, Resonant Recognition Model, molecular modelling, Diebetes

Ključne reči
Insulin, Model Rezonantnog Prepoznavanja, molekularno modeliranje, diabetes

Abstract

Diabetes is one of the most prevalent diseases of the modern society. It is a group of metabolic and chronic diseases. Insulin peptide is the critical in treating patients with diabetes. However, the best treatment of different types of diabetes needs insulin variants with different balance of stability and activity. Here we present unique model, Resonant Recognition Model (RRM), which is developed by authors that can predict insulin modifications for desired balance of stability and activity. This is unique computational approach which can save a lot of time and money in developing new variations of insulin.

INTRODUCTION

Diabetes is one of the most prevalent diseases of the modern society. It is a group of metabolic and chronic diseases characterised by high blood sugar either because not enough insulin is produced in the body (type 1) or cells do not respond to the produced insulin (type 2). Insulin is peptide hormone produced by pancreas and is central to regulation of carbohydrates (sugar) and fat metabolism in the body. Insulin is one of the most investigated peptides with both structure and function well known. Insulin is stored in the body as a hexamer but is only active as monomer. Hexamer is an inactive form with long term stability, which serves to keep highly reactive insulin protected but yet readily available [(1)].

Insulin injections are the main treatment for Diabetes type 1 where there is a lack of insulin in the body. However, in case of Diabetes type 2 where the main problem is interaction between insulin and receptor the situation is different and cannot be simply solved by insulin injections[1]. In both cases slightly modified insulin is used by different pharmaceutical companies. Modifications are done to balance stability of multimers with activity of monomers[(1)]. However all these modifications are done by random mutations without understanding how the function of formation multimers as distinct to function of receptor recognition is written in the peptide sequence. Here we will use Resonant Recognition Model [(2-4)] to analyse these two functions within the insulin protein and to identify characteristic feature for each function. Once when these characteristics are identified it is possible to predict mutations which will be able to balance stability with activity of insulin in any desired proportion.

RESONANT RECOGNITION MODEL (RRM)

The RRM represents a whole new view to biomolecular interactions, in particular protein-protein and protein-DNA interactions [2-4].

The RRM is based on the finding that certain periodicities (frequencies) within distribution of energies of delocalised electrons along a protein molecule are critical for protein biological function and/or interaction with its target. The RRM enables these frequency characteristics to be calculated. These findings can be applied to the:

a) Definition of protein or DNA functions;
b) Definition of protein or DNA targets and analysis of their mutual recognition;
c) Prediction of amino acids in the protein or nucleotides in the DNA which are mostly important for the function of the macromolecule;
d) Prediction of functionally relevant mutations in proteins and/or DNA;
e) Design of a completely new peptides or DNA fragments with desired spectral characteristics and consequently corresponding biological activities been designed.

All these applications have been tested on a number of examples [(2-11)] including examples which have been already tested in biological systems, for example FGF analogues [(5)],

HIV envelope mimicking peptides [6], and Myxoma virus analogues [7] de novo designed peptides using the RRM.

The model is applied here to identify characteristics related to insulin ability to form multimers on one hand and to interact with receptor on the other hand. Once when these two differently characteristics are identified it is possible to predict mutations that will lead to desired ratio balance between these two functions.

RRM CALCULATIONS – CONCEPT OF CHARACTERISTIC FREQUENCY

All proteins can be considered as a linear sequence of their constitutive elements, i.e. amino acids. Biological function of proteins is determined primarily by the linear sequence of their constitutive elements, i.e. amino acids. The Resonant Recognition Model (RRM) interprets this linear information by transforming a protein sequence into a numerical series and then into the frequency domain using digital signal processing methods, the Fourier Transform (FFT).

Assignment of EIIP – Formation of Numeric Sequence

In the RRM, the protein primary structure is presented as a numerical series by assigning a physical parameter value relevant to the protein's biological activity to each amino acid. 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 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 this study, the energy of delocalised electrons (calculated as the electronion interaction pseudopotential (EIIP) of each amino acid residue was used. The resulting numerical series represents the distribution of the free electrons energies along the protein molecule.

Concept of Spectrum, Cross Spectrum – Characteristic Frequency

At the second stage, the numerical series are analysed by digital signal analysis methods, using Fourier Transform, in order to extract information pertinent to the biological function. The average distance between amino acid residues in a polypeptide chain is about 3.8Å and 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. 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. Therefore, 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 is 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:

$$|M_n| = |X_{1,n}|\ |X_{2,n}| \Lambda |X_{M,n}| \quad \text{K}n = 1,2, \text{K}\ N/2$$

Peak frequencies in such a multiple cross-spectral function present 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. The S/N is calculated as the ratio between the signal intensity at the particular peak frequency and the mean value over the whole spectrum. Previous research results have shown that the value of S/N ratio of at least 20 can be considered as significant. The multiple cross-spectral function of large group of sequences with the same biological function are called „consensus spectrum". The presence of a peak frequency with the 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:

a) one peak only exists for a group of protein sequences sharing the same biological function

b) no significant peak exists for biologically unrelated protein sequences

c) peak frequencies are different for different biological functions.

Above criteria have been implemented throughout the sequence databases[2-4] and the following fundamental conclusion was drawn:

Each specific biological function within protein or DNA is characterised by one frequency [2-4].

The examples shown so far, present that all protein sequences with the common biological function have common frequency component, which is a specific feature for the observed function/interaction [2-11]. This characteristic frequency is related to the protein biological function (Figure 1):

In order to understand the meaning of the characteristic frequency, it is important to clarify what is meant by the biological function of proteins. Each biological process is driv-

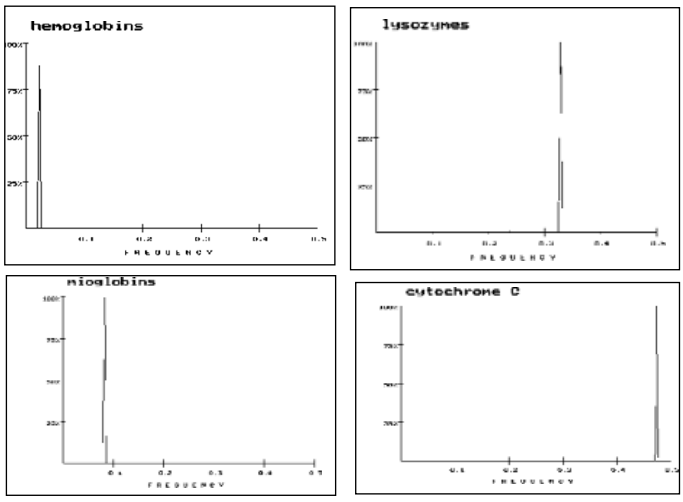

***Figure 1**: Characteristic frequencies of different biological functions*

en by proteins that selectively interact with other proteins, DNA regulatory segment or small molecules. These interactive processes that involve energy transfer between the interacting molecules are highly selective. How is this selectivity achieved? In the RRM it is assumed that the selectivity is defined within the amino acid sequence. It has been shown that proteins and their targets share the same characteristic frequency (Figure 2), but are in opposite phase (phase difference close to 3.14 radian) at this characteristic frequency for each pair of interacting macromolecules [2-4]. Therefore, we conclude that RRM characteristic frequencies represent proteins' general functions as well as mutual recognition between a particular protein and its target (receptor, ligand, etc). This results from matching of periodicities within the distribution of energies of free electrons along the interacting proteins, which can be regarded as the resonant recognition. The RRM model assumes that characteristic frequencies are responsible for the resonant recognition between macromolecules at a distance. Thus, these frequencies have to represent oscillations of some physical field which can transmit through water dipoles. One of the possibilities is that this field is electromagnetic in nature [2-4,12,13].

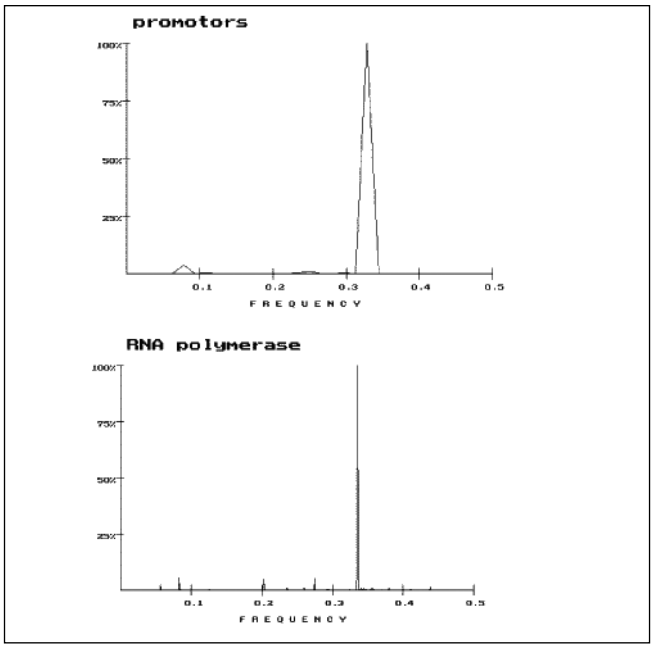

Figure 2: *Cross-spectra of promoters and RNA polymerase showing that Interacting macromolecules have the same characteristic frequency*

Therefore, it has been proposed that interacting molecules „communicate" with each other i.e. recognise each other at the distance, based on the same/similar (within the calculation error) characteristic frequency, but opposite phases at that frequency.

Applications of RRM

Once when we know that each biological function/interaction is characterised with RRM frequency and phase we can use this parameter to analyse protein biological function, predict key amino acids („hot spots") for the particular function and even design de novo peptides with the desired biological function or interaction ability.

This approach has been already used in a number of examples [2-11,14,15] .

For example, in the case of Fibroblast Growth Factors two characteristic frequencies were identified: one related to receptor recognition and another related to „growth activity". The aim of that particular project was to design peptide which can competitively bind to the FGF receptor but without inducing growth. Using only receptor recognition frequency the 16-mer peptide was designed experimentally tested and indeed had receptor recognition activity without inducing growth [5].

In the case of HIV virus the one common RRM frequency was identified for all HIV envelope protein despite their high variability. This frequency was used to design peptide that can immunologically mimic all HIV isolates and thus could be a good candidate for vaccine [6].

Similar idea as described above was used to mimic myxoma virus oncolytic function. Myxoma virus (MV) is a rabbit-specific poxvirus pathogen that also exhibits a unique tropism for human tumor cells and is dramatically oncolytic for human cancer xenografts. The RRM characteristic frequency for MV proteins was identified and used to design peptides that were experimentally shown to mimic myxoma virus oncolytic function [7].

Here, we used RRM approach to analyse insulin peptides with the aim to identify distinct characteristic frequencies for a) insulin molecules clustering and b) receptor recognition. If these characteristic frequencies are identified than it would be possible to mutate insulin molecule in such way to either increase its affinity to make clusters and thus produce slow activity or to increase receptor recognition with lower clustering ability and thus increase strength and speed of activity. Thus then it would be possible to rationally balance stability of multimers with activity of monomer.

RESULTS

Eight insulin peptides (P01317 – *insulinbovin*, P67970 – *insulinchicken*, P01321 – *insulindog*, P01308 – *insulinhuman,* P01315 – *insulinpig*, P01311 – *insulinrabit*, P01322 – *insulinrat*, P01318 – *insulinsheep*) were analysed using the RRM approach. The cross spectrum of 8 insulin peptides is presented in Figure 3. It can be noticed that there are two distinct common frequencies, f1=0.0439+-0.008 and frequency range from f2=0.36 to f2=0.39 for all analysed insulin peptides.

To identify which functions/interactions are these frequencies characterising the insulin peptides were compare with four insulin receptors (Q9PVZ4 – *insulinfrogreceptor*, P06213 – *insulinhumanreceptor*, P15208 – *insulinmousereceptor*, P15127 – *insulinratreceptor*). The result is presented

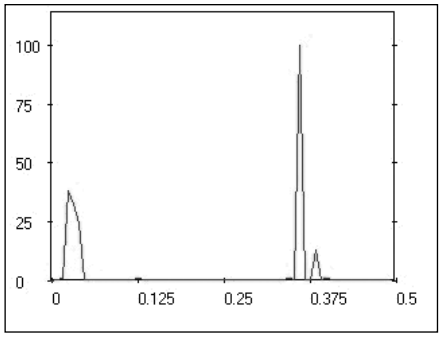

Figure 3*: 8 insulin cross spectrum*

in Figure 4. It could be observed that frequency f2=0.3828 has become much more prominent which makes it good candidate to be receptor recognition characteristic frequency.

To additionally check if this frequency is responsible for the recognition we will check phases in the interacting pairs of proteins. The phases for human insulin at frequency 0.3828 is -1.367, while the phase for human insulin receptor at frequency f2=0.3828 is 2.741. As the phase difference of

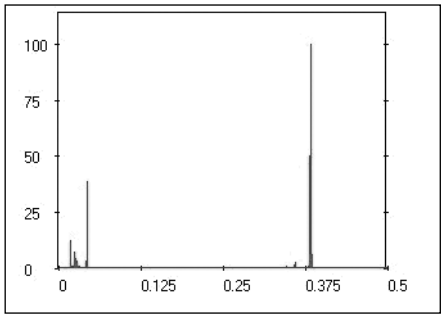

Figure 4*: 8 insulin and 4 insulin receptors cross spectrum*

3.816 at frequency 0.3828 is close enough to be considered opposite (close to 3.14) according to RRM it could be proposed that frequency of 0.3828 is good candidate to be insulin– insulin receptor recognition.

As frequency f2=0.3828 is identified as receptor recognition frequency this leaves frequency f1=0.0439 responsible for multimer formation. By changing balance of these two frequencies in human insulin it is possible computationally [2-4,13,14] to find out key („hot spot") amino acids and consequently mutations in human insulin which can produce insulin with desired balance of multimer formation and activity.

CONCLUSION

We have applied here RRM approach to identify functional characteristics within the insulin peptide. Two distinct characteristic frequencies were identified f1=0.0439 and f2=0.3828. Through comparison with insulin receptor it has been found that frequency f2=3828 is most probably receptor recognition characteristics leaving frequency f1=0.0439 to be responsible for multimer formation. With the knowledge of these two frequencies it is now possible to find out key amino acids for each of these frequencies and consequently rationally predict mutations that can influence either multimer formation or receptor recognition. This could lead to new variants of insulin with balance of stability with activity in any desired proportion.

Acknowledgment
We would like to acknowledge the support from AMALNA Consulting for this research.

Sažetak

Dijabetes je jedna od najčešćih metaboličnih hroničnih oboljenja u modernom društvu. Osnovni tretman kod dijabetesa je upotreba insulina. Za tretman različitih tipova dijabetesa je važno upotrebiti varijante insulina sa različitim odnosom stabilnosti i aktivnosti. U ovom radu Model Rezonantnog Prepoznavanja (RRM), koji su autori razvili, pokazao je mogućnost za predikciju insulinskih modifikacija sa željenom stabilnosti i aktivnosti. Ovaj kompjuterski prilaz može da uštedi vreme i novac za razvoj novih varijanti insulina.

REFERENCES

1. Diabetes mellitus-Wikipedia,the free encyclopedia en.wikipedia.org/wiki/Diabetes_mellitus?

2. Cosic I: Macromolecular Bioactivity: Is it Resonant Interaction between Macromolecules?-Theory and Applications, IEEE Trans. on Biomedical Engineering, 1994; 41:1101-14..

3. Cosic I: The Resonant Recognition Model of Macromolecular Bioactivity: Theory and Applications, Basel, Birkhauser Verlag, 1997.

4. Cosic I: Virtual Spectroscopy for Fun and Profit, Biotechnology, 1995; 13: 236-238.

5. Cosic I, Drummond AE, Underwood JR, Hearn MTW: A New Approach to Growth Factor Analogue Design: Modelling of FGF Analogues, Molecular and Cellular Biochemistry, 1994; 130:1-9.

6. Krsmanovic V, Biquard JM, Sikorska-Walker M, Cosic I, Desgranges C, Trabaud MA, Whitfield JF, Durkin JP, Achour A, Hearn MT: 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(5), pp:410-412.

7. Istivan T, Pirogova E, Gan E, Almansour NM, Coloe PJ, Cosic I: Biological effects of a de novo designed myxoma virus peptide analogue: Evaluation of cytotoxicity on tumor cells, 2011, PLoS ONE 6 (9), art. no. e24809.

8. 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.

9. Pirogova E, Akay M, Cosic I: Investigating the interaction between oncogene and tumor suppressor protein', in IEEE Transactions on Information Technology in Biomedicine, IEEE-Inst Electrical Electronics Engineers Inc, Piscataway, 2009; vol. 13, no. 1, pp:10-15 ISSN: 1089-7771.

10. Pirogova E, Vojisavljevic V, Caceres J, Cosic I: Ataxin active site determination using spectral distribution of electron ion interaction potentials of amino acids in Medical & Biological Engineering & Computing, Springer, Germany, 2010; vol. 48, no. 4, :303-309 ISSN: 0140-0118.

11. Pirogova E, Istivan T, Gan E, Cosic I: Advances in methods for therapeutic peptide discovery, design and development, Current Pharmaceutical Biotechnology, Bentham Science Publishers Ltd., Netherlands, 2010; ISSN: 1389-2010 12 (8), 1117-1127.

12. Cosic I: The Resonant Recognition Model of Bio-molecular Interactions: possibility of electromagnetic resonance, Polish Journal of Medical Physics and Engineering, 2001; vol. 7. No.1,:73-87.

13. Pirogova E, Vojisavljevic V, Istivan T, Coloe P, Cosic I: Review Study: Influence of electromagnetic radiation on enzyme activity and effects of synthetic peptides on cell transformation. MD-Medical data, 2010; 2(4): 317-324.

14. Cosic I, Hearn MTW: "Hot Spot" Amino Acid Distribution in Ha-ras Oncogene Product p21: Relationship to Guanine Binding Site, J. Molecular Recognition, 1991; 4,:57-62.

15. Cosic I, Pavlovic M, Vojisavljevic V: Prediction of "Hot Spots" in Interleukin 2, Based on Informational Spectrum, Characteristics of Growth Regulating Factors, Biochimie, 1989; 3,:333 342.

■ The invited paper was received and accepted on 03.12.2013.