Computational and Structural Biotechnology Journal 15 (2017) 265–270

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Fourier Analysis of Conservation Patterns in Protein Secondary Structure

Ashok Palaniappan[a,*], Eric Jakobsson[b,*]

[a]Dept of Biotechnology, Sri Venkateswara College of Engineering, Post Bag No. 3, Pennalur, Sriperumbudur 602117, India
[b]University of Illinois at Urbana–Champaign, IL 61820, USA

ARTICLE INFO

Article history:
Received 8 November 2016
Accepted 12 February 2017
Available online 22 February 2017

Keywords:
Periodicity
Secondary structure
Evolution
Moment of conservation
Fourier transform
Potassium channel

ABSTRACT

Residue conservation is a common observation in alignments of protein families, underscoring positions important in protein structure and function. Though many methods measure the level of conservation of particular residue positions, currently we do not have a way to study spatial oscillations occurring in protein conservation patterns. It is known that hydrophobicity shows spatial oscillations in proteins, which is characterized by computing the hydrophobic moment of the protein domains. Here, we advance the study of moments of conservation of protein families to know whether there might exist spatial asymmetry in the conservation patterns of regular secondary structures. Analogous to the hydrophobic moment, the conservation moment is defined as the modulus of the Fourier transform of the conservation function of an alignment of related protein, where the conservation function is the vector of conservation values at each column of the alignment. The profile of the conservation moment is useful in ascertaining any periodicity of conservation, which might correlate with the period of the secondary structure. To demonstrate the concept, conservation in the family of potassium ion channel proteins was analyzed using moments. It was shown that the pore helix of the potassium channel showed oscillations in the moment of conservation matching the period of the α-helix. This implied that one side of the pore helix was evolutionarily conserved in contrast to its opposite side. In addition, the method of conservation moments correctly identified the disposition of the voltage sensor of voltage-gated potassium channels to form a 3_{10} helix in the membrane.

1. Introduction

Amino-acid conservation is an evolutionary property. Physical properties of amino acid side-chains exhibit a higher-order moment (also known as periodicity) in the context of repetitive secondary structures, such as the α-helix and β-sheet. A notable physical property whose moments turned out to be significant is the hydrophobicity [1]. The disposition of structured domains in the protein is strongly correlated with the overall hydrophobicity and the amphiphilicity of the domains. These properties stabilize the structure of the protein and, for membrane proteins, the protein's association with the membrane. For α-helical membrane proteins, the strength of the hydrophobic moment is maximal at the period of the helix (i.e, 100°). Similarly, for beta-barrel membrane proteins, the hydrophobic moment is maximal at the period of the beta sheet (i.e, 160°–180°). In both cases, the surface of the secondary structure element which is in contact with lipid exhibits a strong hydrophobicity to allow for partitioning into the membrane. Domains exposed to the electrolyte on either side of the membrane exhibit strong amphiphilicity, or a high hydrophobic moment. The periodicity of residue properties of α-helical proteins can be visualized using the helical wheel representation. Spatial asymmetry in the distribution of hydrophobicity, say, on the helical wheel would imply the fine-tuning of protein function via the achievement of amphiphilicity. Amphiphilicity has turned out to be key to the activity of antimicrobial peptides. Most native and engineered antimicrobial peptides face the amphiphilicity requirement to successfully insert into and permeabilize the bacterial membrane [2].

The use of sequence profiles improved the ability of hydrophobicity to predict the formation of α-helices [3]. Analogous to hydrophobicity, we consider that the residue conservation in a protein alignment displays a first-order moment. Residue conservation is directly correlated with general functional importance. The moment of residue conservation would likely contain information not captured by a linear residue-by-residue conservation. The evolutionary basis of the moment of conservation is as follows: one face of an α-helix involved in critical interatomic interactions must be

* Corresponding authors.
E-mail addresses: aplnppn@gmail.com,ashok@svce.ac.in (A. Palaniappan), jake@illinois.edu (E. Jakobsson).

http://dx.doi.org/10.1016/j.csbj.2017.02.002

conserved, while the diametric face might not be equally constrained and evolve with neutral drift. In order to detect and quantify this spatially oscillatory constraint in the protein secondary structure, we introduce a measure called 'conservation moment' and illustrate its applications.

2. Material and Methods

2.1. Calculation of the Zeroth Moment of Conservation

The zeroth moment of conservation is the sum of the conservation values of the residues based on a profile of homologous sequences. The profile is built using homology detection methods and multiple sequence alignment. The conservation c_n of each column n of the alignment could then be computed using, for e.g., Shannon entropy:

$$c_n = -\sum_i p_i \ln p_i \tag{1}$$

where the p_i's are the probabilities of finding residue i in column n and the summation is over all the 20 amino acids. The c_n's are scaled from 0 to 1, 0 denoting a column of all different residues and 1 denoting a column of all identical residues. The resulting one-dimensional function of conservation values over the length of the alignment is called the conservation vector. The zeroth conservation moment C_0 of an alignment segment of length N is equal to the sum of the c_n's of the columns of the alignment segment.

$$C_0 = \sum_{n=1}^{N} c_n \tag{2}$$

C_0 is a measure of the net conservation of an alignment segment. A contiguous sequence of conserved residues in a protein family would give rise to a high C_0.

2.2. Calculation of First-order Conservation Moment

To detect an asymmetry in the conservation pattern of an alignment segment, we search for periodicities in the corresponding conservation vector. The moment of the conservation vector at a given periodicity is a measure of the signal strength at that periodicity, and is known as the first-order conservation moment, $C_1(\theta)$. For a given period θ,

$$C_1(\theta) = \left\{ \left[\sum_{n=1}^{N} C_n \sin(\theta n) \right]^2 + \left[\sum_{n=1}^{N} C_n \cos(\theta n) \right]^2 \right\}^{\frac{1}{2}} \tag{3}$$

where N is the length of the alignment segment, and the period θ is measured in radian. An evolutionary asymmetry in the α-helix structure would be manifested as a strong conservation moment at the period of the α-helix. This corresponds to $\theta = 2\pi/100^\circ = 3.6$ rad. Similarly, an evolutionary moment in the β-sheet structure would give rise to a maximal signal at the period of the β-sheet ($=160^\circ$–180°). Eq. (3) could be rewritten as the modulus of the fourier transform of the conservation vector.

$$C_1(\theta) = \left| \sum_{n=1}^{N} C_n e^{i\theta n} \right| \tag{4}$$

3. Results and Discussion

When the protein secondary structure is known, from a crystal structure or otherwise, $C_1(\theta)$ could be calculated for each secondary structure element at its respective period to detect any spatial asymmetry in evolutionary pressure. Periodicity in evolutionary pressure is valuable for transmembrane structures which accommodate hydrophobic constraints to be stable in the lipid bilayer. This might enable the transmembrane structure to achieve a higher-order functional specificity. An illustrative secondary structure element is the pore helix of the potassium ion (K^+) channel.

Potassium channels are tetrameric transmembrane (TM) structures with two TM helices per subunit [6]. In addition, each subunit has a pore helix that spans half the membrane before looping back. These pore-helices are under an interesting evolutionary constraint. By virtue of scaffolding the 'selectivity filter' of potassium channels, their packing interfaces are evolutionarily constrained. This sidedness of conservation could be detected using the first-order conservation moment. A profile of all the human potassium channel sequences was constructed in the following manner. A representative sequence of each potassium channel subfamily was chosen, and used as a query in PSI-BLAST with an E-value of 0.001 until convergence [7]. After eliminating duplicates, each hit was screened for the presence of selectivity filter characteristic of a potassium-selective channel, to obtain 123 channels (available as a supporting information).

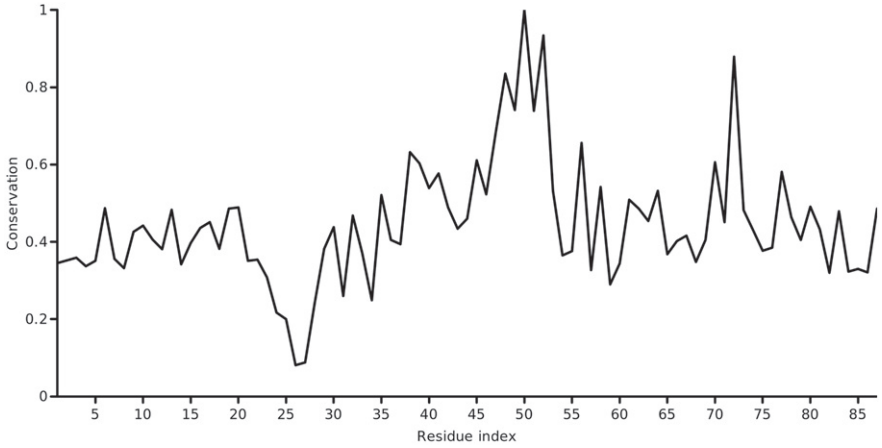

Fig. 1. Profile of the conservation of each position in the KcsA potassium channel sequence, as calculated using Scorecons. A peak (conservation = 1.0) corresponding to the selectivity filter of the K^+-channel could be observed at position 50.