# ScanMoment: a web server for combinatorial analysis of basic residues in nucleic acid binding sites

**Steven E. Massey**

Department of Biology, University of Puerto Rico – Rio Piedras, PO Box 23360, San Juan, Puerto Rico 00931; Fax: (787) 7642610; Tel: (787) 764000 x 2037; Steven E. Massey **-** E-mail: stevenemassey@gmail.com

**Abstract:**
ScanMoment is a webserver designed to identify the presence of the basic faced α-helix (BFAH) motif in the nucleic acid binding sites of proteins. The program calculates the 'Basic Moment', a parameter that quantitizes the distribution of basic residues on the surface of an α-helix. A sliding window is used to generate a plot displaying regions of the protein sequence that possesses a high Basic Moment and thus likely to possess a BFAH motif. The user may vary the periodicity from that of an alpha-helix (100$^{o}$), to those of other secondary structures such as beta sheets and 3$_{10}$ helices. The program can also plot the periodicity of basic residues in a protein sequence using a Fourier transformation. The procedure has been used to characterize the presence of BFAHs in the N-terminal extensions of the eukaryotic aminoacyl-tRNA synthetases and to indicate the presence of a BFAH in the tRNA binding site of alanyl-tRNA synthetase.

**Availability:** www.scanmoment.org

**Keywords:** basic moment, Fourier transformation, periodicity, alpha helix, Aspartyl-tRNA synthetase

**Description:**
Protein functional motifs are usually recognized by amino acid sequence similarity. The approach depends on the one-dimensional order of residues in a sequence. An alternative perspective is to consider the combination of residues in a sequence. Such periodicity is not expected to be reflected in primary sequence homology, making such patterns difficult to detect and generally poorly characterized. α-helices are the most common secondary structural elements present in proteins and the residues in an α-helix display a periodicity of approximately 100$^{o}$, corresponding to 3.6 residues per turn of the α-helix **[1]**. BFAHs were originally detected from primary sequences using Schiffer-Edmundson (helical wheel) representations **[2, 3, 4, 5, 6]**. The helical wheel method is subjective and labor intensive, therefore a quantitative method of measuring the distribution of basic residues in a protein sequence, the 'Basic Moment', was developed. The Basic Moment is derived as follows. An α-helix may be visualized as a cylinder, with the functional groups of the amino acids pointing outwards from the surface of the cylinder. When the cylinder is viewed end on, a resultant vector (μ), points outwards from the center of the α-helix and is composed of two perpendicular components as given by equation 1 (supplementary material). The perpendicular is taken arbitrarily as the line made by the first residue with the center of the α-helix. For a window of N =18 residues, the angle of θ is incremented in 18 steps of 5Л/9 radians (5Л/9 = 100$^{o}$). The function was squared in order to improve the signal to noise ratio. The resulting parameter (μ$_B$) is termed the Basic Moment as given in equation 2 (supplementary material).

The Basic Moment is a finite Fourier expansion of the distribution of basic residues in an amino acid sequence. ScanMoment has a user defined choice of a 12 or 18 residue sliding window for the calculation of the Basic Moment and plots the resulting value in the centre of the window. An ideal BFAH,

with all of the basic residues arrayed on one side of the α-helix, has a Basic Moment of 33.2 for a window of 18 residues (**Figure 1a).** A Basic Moment plot of yeast cytoplasmic aspartyl-tRNA synthetase is displayed in **Figure 1b.**

Aspartyl-tRNA synthetase aminoacylates asp-tRNA with aspartic acid. The peak at residues 30-47, indicated by arrows, is the site of a BFAH. The presence of a BFAH at this precise position has been demonstrated experimentally [7]. The BFAH has functional significance as it is responsible for non-specific binding of the enzyme to asp-tRNA **[8]** and is present in a region of the enzyme involved in specifically binding to aspartyl-tRNA synthetase mRNA in an autoregulatory mechanism **[9]**.

In addition to generating a graphical plot, ScanMoment tabulates the ordinate and abscissa values in a flat file. ScanMoment is also capable of plotting the Basic Moment for periodicities corresponding to other protein secondary structures: 87$^{o}$ (π helix), 120$^{o}$ (3$_{10}$ helix) and 160$^{o}$ (beta sheet). The Basic Moment periodicity, μ$_B$ (θ), calculates the Basic Moment at different values of θ, and is defined by equation 3 (supplementary material). For each value of θ, the Basic Moment was calculated by incrementing the angle of Φ by 18 steps of θ. ScanMoment calculates the periodicity for protein segments 18 amino acids long (N =18), calculating values for θ in increments of 1$^{o}$, from 80-180$^{o}$ (18 residues corresponds to one full rotation of the α-helix). The resulting function is a Fourier transform and is equivalent to the periodicity of basic residues in an amino acid sequence. The basic residue periodicity of the region encompassing the 18 amino acids from position 30 to 47 of yeast cytoplasmic aspartyl-tRNA synthetase is displayed in **Figure 1c.** A maxima at 97.5$^{o}$ is observed, consistent with the periodicity of an α-helix. Indeed, it has been demonstrated elsewhere **[10]** that the average periodicity of α-helices is closer to 97.5$^{o}$, as opposed to the figure of 100$^{o}$, calculated elsewhere **[1]**.
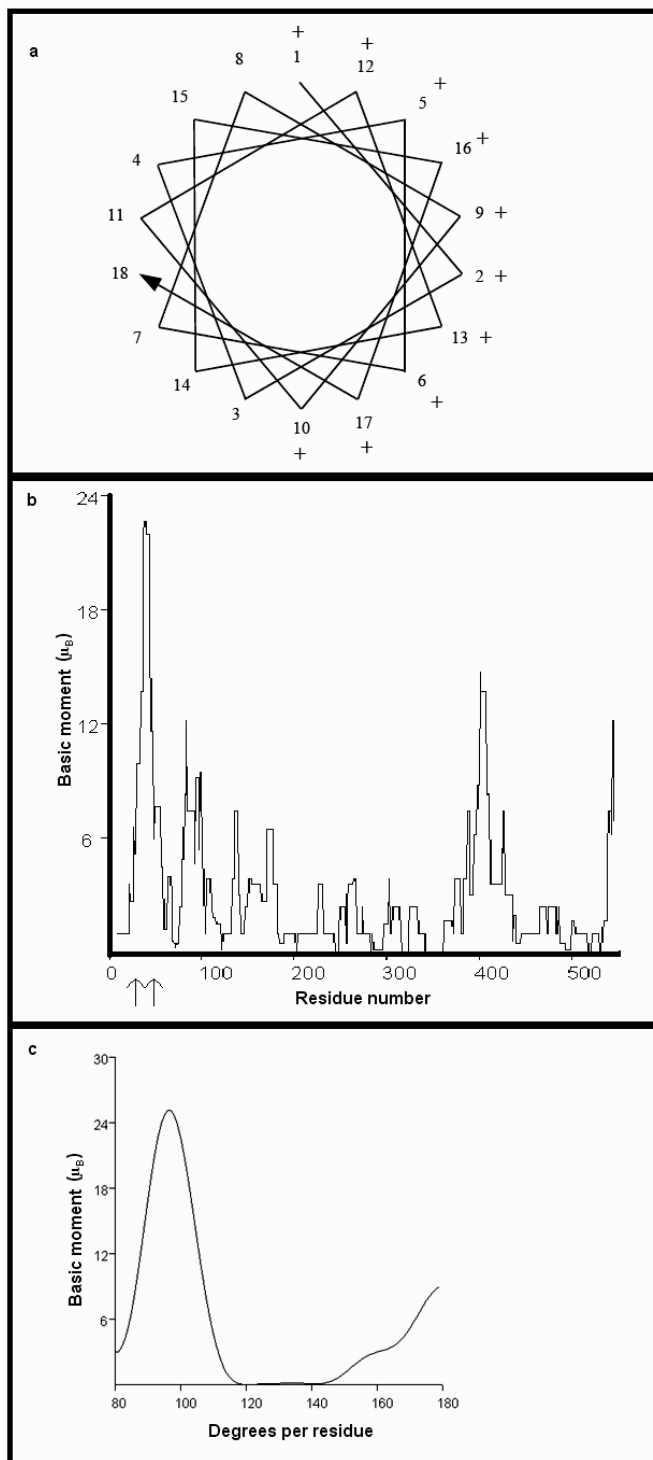
**Figure 1: (a)** A helical wheel representaion of an ideal BFAH. Basic residues, indicated by a + sign, are aligned on one side of the α-helix. Acidic, hydrophilic and hydrophobic residue properties are ignored; **(b)** Basic Moment plot of yeast cytoplasmic apartyl-tRNA synthetase, using a window of 18 residues. The BFAH motif is indicated by arrows (residues 30 to 47); **(c)** Basic Moment periodicity profile of yeast cytoplasmic aspartyl-tRNA synthetase, residues 30 to 47.

The Basic Moment is intended to be of utility in characterizing the nucleic acid binding sites of proteins, as these regions are usually basic in nature, and α-helices are important structures in such regions. The utility of the Basic Moment in characterizing BFAHs, other than those in the N-terminal extensions of the eukaryotic aminoacyl-tRNA synthetases has already been demonstrated, identifying the presence of a BFAH in the active site 'two-helix' pair of prokaryotic and eukaryotic alanyl-tRNA synthetases **[11, 12]**. Other examples of BFAHs remain to be discovered. ScanMoment is written in Java, and utilizes the Tomcat 6.0.2 server running on Red Hat Enterprise Linux Version 5.

### References
**[1]** L. Pauling, et al., *Proc. Natl. Acad. Sci.* (1951) **37:** 205.
**[2]** B. Chatton et al., *J. Biol. Chem.* (1988) **263:** 52.
**[3]** B. Lorber et al., *Eur. J. Biochem.* (1988) **174:** 155.
**[4]** M. Mirande & J.P. Waller, *J. Biol. Chem.* (1988) **263:** 18443.
**[5]** G. Souciet, et al., *Eur. J. Biochem.* (1999) **266:** 848.
**[6]** M. Francin & M. Mirande, *J. Biol. Chem.* (2003) **278:** 1472.
**[7]** F. Agou et al., *Biochem.* (1995) **34:** 569.
**[8]** M. Frugier, M., et al., *EMBO J.* (2000) **19:** 2371.
**[9]** M. Frugier & R. Giege, *J. Mol. Biol.* (2003) **331:** 375.
**[10]** J.L. Cornette et al., *J. Mol. Biol.* (1987) **195:** 659.
**[11]** L. Ribas de Pouplana, et al., EMBO J. (1998) **17:** 5449.
**[12]** S.E. Massey, *In Silico Biol.* (2006) **6:** 0025.

## Supplementary material

$$\mu = \sqrt{\left(\sum_{i=1}^{N} B_i \sin\theta_i\right)^2 + \left(\sum_{i=1}^{N} B_i \cos\theta_i\right)^2} \qquad (1)$$

$$\mu B = \left(\sum_{i=1}^{N} B_i \sin\theta_i\right)^2 + \left(\sum_{i=1}^{N} B_i \cos\theta_i\right)^2 \qquad (2)$$

$$\mu B(\theta) = \left(\sum_{i=1}^{N} B_i \sin\Phi_i\right)^2 + \left(\sum_{i=1}^{N} B_i \cos\Phi_i\right)^2 \qquad (3)$$

where N is number of residues, $B_i$ is the basicity of the ith residue (1 for lysine and arginine and a value of 0 for all other amino acids), and $\theta_i$ is the angle made to the perpendicular by the $i^{th}$ residue.